Alternative in vitro assays in nanomaterial toxicology

Article *in* Wiley Interdisciplinary Reviews Nanomedicine and Nanobiotechnology · November 2011

DOI: 10.1002/wnan.153 · Source: PubMed

2 authors:

Thomas Hartung
Johns Hopkins University
558 PUBLICATIONS **18,230** CITATIONS
SEE PROFILE

Enrico Sabbioni
Veneto Nanotech
320 PUBLICATIONS **6,213** CITATIONS
SEE PROFILE

Overview

Alternative *in vitro* assays in nanomaterial toxicology

Thomas Hartung[1,2]* and Enrico Sabbioni[3,4]

Nanomaterials are acclaimed for their novel properties, for which broad new uses are being discovered with increasing frequency. It is obvious that, as the properties change, unwanted properties (toxicity) are to be expected as well. Current toxicology, however, is already overwhelmed with the challenge of addressing new chemicals, not to mention the enormous number of old chemicals never properly assessed. Limitations of traditional approaches range from animal welfare issues, which were a strong driving force for alternative approaches (the 3Rs concept) over the last two decades, to aspects of throughput and accuracy of the predicted toxicities. The latter has prompted discussion about a revolutionary change in chemical safety assessment, now known as Toxicology for the 21st Century (Tox-21c). The multitude of possible formulations of nanomaterials to be assessed for novel toxic properties makes these alternative approaches especially attractive, given the well recognized limitations of traditional animal-based approaches—limitations that might be even more pronounced for nanomaterials, which have notably altered biokinetics.
WIREs Nanomed Nanobiotechnol 2011 DOI: 10.1002/wnan.153

INTRODUCTION

In 2005, at the 44th annual meeting of the American Society of Toxicology in New Orleans, J.V. Rodricks and A. Seaton debated about a putative risk posed by nanoparticles (NP). They debated again, with reversed roles, at the EuroTox meeting in Cracow the same year. On both occasions a majority of colleagues voted against considering NP a risk. Since then, articles, journals, and textbooks, and funding exploring this concern have proliferated. Not surprisingly, the topic was again on the agenda of the 2009 SOT/EuroTox debate; the shift in opinion was easily grasped.

*Correspondence to: THartung@jhsph.edu

[1]Department of Environmental Health Sciences, Center for Alternatives to Animal Testing (CAAT), Doerenkamp-Zbinden Chair for Evidence-based Toxicology, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD, USA

[2]Pharmacology and Toxicology, University of Konstanz, Konstanz, Germany

[3]ECSIN LAB—European Center for the Sustainable Impact of Nanotechnology, Veneto Nanotech Scpa, Viale Porta Adige, Rovigo, Italy

[4]Immunotoxicology and Allergy Unit, Aging Research Centre (CeSI), G.d'Annunzio University Foundation, Via Colle dell'Ara, Chieti, Italy

DOI: 10.1002/wnan.153

Fifty years after Nobel laureate Richard Feynman's famous lecture predicting the potential of the nanoscale,[1] nanomaterials have arrived in daily life. For the purpose of this article, we will use the term NP for all the different nanoproducts to be tested. While these concerns are certainly less applicable to nanothick coatings, we will not distinguish specifically for the variety of materials developed. Obviously, different nanomaterials have very different health concerns, for example, nanofibers are more likely linked to respiratory carcinogenicity similar to asbestos or size and shape will dramatically influence bioavailability for the organism as well as the cell determining test needs and test possibilities. For the scope of this article, only those aspects that translate to the majority of NP will be discussed.

It may be argued that there is no such thing as alternatives in nanotoxicology. Given its short history,[2] nanotoxicology is not established as such and actually has the opportunity (and probably the need, as we will show) to base itself on novel approaches anyway. Especially in Europe, there are some political incentives to avoid animal testing,[3] but we will show that there are also scientific incentives to limit the *in vivo* part of the necessary risk assessment. However, the field of 'alternative methods' has a lot to offer with regard to technologies,[4] quality assurance,[5,6]

validation,[7,8] and regulatory acceptance[9,10] relevant in the nanotoxicology discussion, which shall be explored here. The call for additional NP evaluation further demonstrates the need to move toward a renovation of approaches in risk assessment of consumer products and medicines in general.

This article will focus first on human-relevant health effects. Only minor attention will be given to ecotoxicological and veterinary effects. This is because of the background of the authors, not to the importance of the possible effects. Similarly, computational approaches, 'in silico toxicology', will be touched upon only lightly.

WHAT IS AN 'ALTERNATIVE' METHOD?

Although the term 'alternative method' only emerged in the 1970s, the concept can be traced back to the 1959 book by Russell and Burch, *The Principles of Humane Experimental Technique*.[11] Russell and Burch described what they called 'the 3Rs' principle: whenever possible, *reduce, refine,* and *replace* animal experimentation. This paradigm has since been endorsed by government agencies, scientific associations, and funding bodies. Most of it should go without saying, that is, no animal pain or distress should be imposed unnecessarily. Most science has enough drivers beside animal welfare considerations to move in this direction anyway. We have seen the success of *in vitro* and *in silico* approaches in complementing and substituting for *in vivo* approaches in most areas of the life sciences and agent discovery. Work with genetically modified mice may serve as a notable exception to this trend,[12] however, showing an increase in animal use rather than a decline over the last decade.

Regulatory toxicology, which was especially resistant to change, prompted in response most of the development of alternative approaches and their underlying concepts. Why was this so? First, it is a field of regulated, prescribed methods, with set standards to be followed by the regulated industries. Second, the well warranted desire for international harmonization has even 'set in stone' agreed testing guidelines. For example, at the level of OECD or (Veterinary) International Conference on Harmonization [(V)ICH], the harmonization effort, which via avoidance of duplicate testing has likely saved more animals than all alternative methods together, increasingly represents an obstacle for adaptation to technical progress.[13] Third, a lot of work in toxicology is never published and thus remains outside the normal scientific peer discussion because of both the proprietary nature of data and the lack of scientific interest in routine, often negative, data.[14] Fourth, especially in the field of safety testing, repeat experiments limiting animal use are often prohibited, further hiding the reproducibility issues and discrepancies of current approaches.

A discussion has emerged over the last two decades, especially in Europe, which is animal welfare driven. The hallmark pan-European animal welfare legislation of 1986 not only requires the use of alternatives whenever available but also commits the European Commission, the executive branch of the EU, and the member states to support the availability of alternative methods. This has brought about without doubt some successes.[15] As a consequence, several national competence centers and funding programs were created.[16] At the European level, more than €200 ($300) million has been made available so far. An action plan for animal welfare was put forward[17] and the first validation body was installed in 1991: the European Centre for the Validation of Alternative Methods (ECVAM),[18,19] part of the European Commission's Joint Research Centre in Ispra, Italy. (Both authors have worked for ECVAM in the past.) Similar institutions followed in the United States [The National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM) and the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM)], Japan (JaCVAM), and Korea (KoCVAM). In 2005, international consensus was reached at the OECD level regarding the process and need for validation for test guidelines; every new or amended test guideline introduced after 2005 must be based on validation.[20] However, the identification of a suitable point of reference for validation often is a challenge.[21]

A number of dedicated organizations (3Rs centers in Austria, Germany, the Netherlands, United Kingdom, Poland, etc.) and societies, such as the European Society for Alternatives to Animal Testing (EUSAAT), the European Consensus Platform for Alternatives (ECOPA), the European Society for Toxicology *In Vitro* (ESTIV), and, most recently, the Center for Alternatives to Animal Testing-Europe (CAAT-EU),[22] along with many other national organizations, make this a lively arena in Europe. Most remarkably, the European Partnership for Alternative Approaches to Animal Testing (EPAA) brings together the European Commission, more than 45 companies, and seven trade associations.[23,24] The Johns Hopkins Center for Alternatives to Animal Testing (CAAT) and the Institute for *In Vitro* Sciences (IIVS) have greatly furthered the US discussion over the last decades. Now an American Society for Cellular and Computational Toxicology is being formed (http://ascctox.org).

Recently, remarkable progress has been made at various levels. This includes EU test guideline regulation, various methods accepted by OECD and ICH, and the creation of the International Collaboration of Cosmetic Regulators, or ICCR.[23] International collaboration among validation bodies is also extremely helpful. At the World Congress on Alternatives and Animal Use in the Life Sciences in Tokyo (2007), one of the authors called for the creation of an International Council of Validation Bodies, which led within 2 years to the creation of the International Collaboration on Alternative Test Methods (ICATM).

Two major legal initiatives in Europe have further shaped the field of alternatives over the last decade, with varying impact on the safety assessment of NP. The Seventh Amendment to the Cosmetic's Directive has banned animal testing for finished products as of 2004 and for ingredients as of 2009; this is reinforced by a marketing ban with deadlines in 2009 and 2013. A major European industry[24] was challenged by this legislation, which resulted in large-scale efforts,[14] also in collaboration with the European Commission. Most recently, for example, a joint call for proposals for €50 ($75) million was put forward to develop alternatives for systemic toxicity testing. Significant progress has been made for topical and acute toxicities (see below) in response to the legislation. As cosmetic products use NP extensively, especially for sun creams, available methods should be considered for all areas of safety assessment. The second European legislation, REACH, addresses the regulation of industrial chemicals.[25,26] Both the United States and the EU started systematic evaluation of new chemicals in the early 1980s, leaving existing 'old' chemicals essentially untouched. After two decades of experience, it became clear that this covered only few percent of turnover with, and exposure to, chemicals. Europe first advanced this with the REACH legislation, but similar legal initiatives are underway in Canada, Japan, and the United States. REACH is the forerunner of these, and problems are unavoidable, as pointed out in the current discussion about the excessive demands for reproductive toxicity testing.[27–29] NP, at this moment, would normally fall under the REACH legislation. As existing substances they would have had to be preregistered by December 2008; otherwise they need to be fully registered immediately to continue to be marketed in Europe. REACH also applies to all new chemicals exceeding 1 ton of production or commerce in Europe. So far, no specific requirements for NP exist, but any concerns about additional safety issues would require the notifier to adapt the testing proposals and plans. The requirements for full life cycle exposure assessment in REACH are also quite extensive, so that downstream users also have obligations under this legislation. Most importantly, the discussion about the creation of REACH has led to the insight (followed by implementation) that a program of this kind requires the development of new alternative approaches. While not really mentioned in the preparatory White Paper in 2001, the development of alternative approaches became a principal goal of the legislation anchored in Article 1. The legislation is in many respects a revolution in the way chemicals are assessed; most importantly, it moves from a laundry list of (animal) test methods to the integrated use of various information sources, including existing data, *in vitro, in silico* and—as a declared last resort—animal test data. This has spurred the development of integrated (or intelligent) testing strategies, published as guidance to industry by the European Chemical Agency. One of the authors was responsible for coordinating the elaboration of this guidance by more than 200 experts on behalf of the European Commission. These strategies represent the first steps toward a novel paradigm for testing, and projects like NanoTest[30] aim to develop similar test strategies for NP. Despite introducing many novel tools and a tiered approach to testing, however, they do not fundamentally change the test methodology, as validated alternatives (i.e., methods designed and calibrated against existing approaches) are employed. No redesign of approach was undertaken, and the use of modern biotechnology is scarce, to say the least.

An unexplored aspect of the REACH legislation for nanotoxicology is the fact that REACH considerably lowered the testing demands for new and old low-production volume chemicals: not only was the start point shifted from 100 kg to 1 ton of annual production, but below 10 tons only two animal tests (acute toxicity and skin sensitization) are required.[25,31] Certainly most NP are below 10 tons per year of commerce. Few will believe that this is sufficient to address new entities such as NP. This alone might prompt an independent regulation of NP in Europe.

A completely different discussion resulted from the US efforts toward regulating existing chemicals: in preparation for a possible reauthorization of the Toxic Substance Control Act (TSCA), the US Environmental Protection Agency (EPA) renewed its toxicity testing paradigm in 2009.[32] This decision was based on an EPA-commissioned vision and strategy paper by the National Research Council/National Academy of Sciences.[33] The paper is remarkable, as the committee unanimously states the limitations of current approaches and the need for new approaches, while creating a blueprint for the way forward. This report created an atmosphere of departure in

toxicology,[4,34–39] leading not only to broad discussions in journals and conferences but also to agencies setting these changes into practice[40]: '*We propose a shift from primarily* in vivo *animal studies to* in vitro *assays,* in vivo *assays with lower organisms, and computational modeling for toxicity assessments*'. Francis Collins, now Director of the National Institutes of Health (NIH), commented in *USA Today*: '*[Toxicity testing] was expensive, time-consuming, used animals in large numbers, and didn't always work*'. In the same article, Elias Zerhouni, then Director of NIH, is cited: '*Animal testing won't disappear overnight, but the agencies' work signals the beginning of an end*'. The Tox-21c discussion is on its way to creating a Tox-21c movement calling for a Human Toxicology Project.[14,41] The increasing support for EPA's ToxCast project,[42] which implements the first steps toward the vision, is evidence of this. Though there is overlap in the 3Rs and the Tox-21c ideas and activities, there are also remarkable differences.[14] Most importantly, Tox-21c is based less on animal welfare considerations and more on throughput of testing and the quality of results, especially when extrapolating to humans. It also embraces, to varying degrees, new computational and biotechnological methods. Most 3Rs methods so far are relatively simple, with the obvious advantages of robustness, transferability, low costs, and reproducibility. However, we should be aware that there is still a long road to go,[43] but the changes to come promise to be revolutionary.[44]

The efforts of both the 3Rs and Tox-21c lack a crucial component—a critical review of current practices, which we have called for under the concept of evidence-based toxicology.[6,45,46] This means that only an objective assessment of current practices, acknowledging the shortcomings of all approaches—be they *in vivo, in vitro,* or *in silico*[47–50]—will allow us to overcome the limitations of current tests and accommodate new technologies in their own right.

LIMITATIONS OF ANIMAL MODELS PROMPTING THE USE OF ALTERNATIVE APPROACHES

Before addressing the need, opportunities, and limitations of alternative approaches for NP, we need to address the limitations of the animal-based approaches in general and for NP in particular. Animal models are fundamental tools in the life sciences. They have advantages and disadvantages compared with other approaches. In a few instances, they represent the only reasonable approach. The limitations listed here and summarized elsewhere[48] are typical of animal models. This does not mean that each and every limitation holds true for all animal models, however, nor that the alternatives have fewer limitations. No model is perfect. The term 'model' implies deviation from reality, usually by simplifying and reducing variables. This is unavoidable if any progress is to be made, but we tend to forget the compromises we made to end up with something feasible.

The limitations can be grouped into three categories: ethical (can we justify animal use and suffering?), cost/throughput (can we test enough?), and predictivity (do we get the right results?). The third question is not as evident as the first two and requires some explanation before we can discuss the alternative approaches.

Species differences. Most laboratory animal species are much more similar to each other than to humans. There is no reason to assume that any species predicts human health effects better than they predict effects in each other. If such comparisons are done for the more interesting complex endpoints such as cancer or chronic and reproductive toxicity, correlations are no better than 60–70%. It is not yet established whether NP show larger species differences than other test materials.

Pharmaco- and toxicokinetics of substances are different in humans. Humans are not 70 kg rats—we differ tremendously from all laboratory animal species in: metabolic rates; food and water intake; gut pH; body temperature; metabolism of xenobiotics with regard to capacity, velocity, and metabolites formed; body fat distribution; elimination routes; carrier proteins/protein binding; defense mechanisms; and life span. There are some indications that the kinetics of NP are especially altered compared with other test materials.

Inbred strains. The broad use of inbred strains has helped to reduce the variability of experimental results. However, working with identical twins implies that the variability of the animal population is precisely what is no longer reflected. This is often acknowledged by a safety factor (typically 10) when extrapolating to the human situation, which might not hold true for every situation. This should not be different for NP.

Young animals and restriction to one sex. Because of cost and logistics, young animals of one sex are typically used, whereas most health effects are found in the elderly (which often display other diseases in parallel and in which sex differences are often noted). NP likely do not differ here.

Animals are stressed. There are many reasons for stress in the animals: transportation, insufficient time for adaptation after transport, cages without

enrichment (e.g., opportunities for species-specific behavior), handling, daytime procedures on night-active species, changes in groups for randomization, and too many male animals caged together. This general concern can be translated to NP.

Group sizes and number of replicates. The smallest possible group size is typically used, often without power analysis to establish whether a significant result can even be expected. The problem is aggravated because maximum information is often sought by studying multiple endpoints or time points without adjusting the statistics for multiple testing (e.g., standard test guidelines for repeat-dose toxicity include up to 40 evaluated endpoints and two-generation reproductive toxicity tests include 80 endpoints). The maintenance of adequate animal husbandry and the support of trained staff are extremely costly. For higher apes, for example, this can amount to tens and hundreds of thousands of Euros. As a consequence, the number of replications and repetitions is frequently limited. This holds true for NP as well.

The prevalence of effects is not considered. We are typically looking for rare properties of substances both in agent discovery and toxicology. The portion of either toxic or effective substances in a panel of substances tested is normally in the low percentage range. Thus, even tests with good specificity will produce some false positives, in a proportion similar to or higher than the real positives. We have pointed this out previously.[51] The typical combination of tests to increase sensitivity—for example, in order not to miss any toxic compound—further reduces the specificity and thus increases the number of false positives.

Tests are either sensitive or specific. Thus, we can either rely on positive or on negative results. However, we typically accept both outcomes of animal tests with the same confidence.

Unrealistic dosages and exposure durations. Toxicology applies high-dose treatments (often at the maximum tolerated dose, that is, with up to 10% of the animals dying), which hardly reflect human exposure. Experimental exposures to NP are especially difficult—agglomeration of NP and administration to animals are special challenges.

Many health effects are multifactorial. Disease models, for example, are mostly based on only one—often artificial—cause. This relates especially to nanomedicines but is no different from other remedies.

Acute models of chronic phenomena. There are enormous pressures to generate results quickly, leading us to develop animal models that establish symptoms quickly, even for modeling chronic health effects.

Adaptive responses of animals are underestimated. This holds true specifically for the fashionable knockout mice experiments. We often forget that these animals have had a lifetime to compensate for the gene defect, but we simplify our interpretation as if only the one gene is lacking. No difference for NP is apparent compared with other test materials.

Lethality as a side-effect. Mortality in animals is often the result of lack of food and water and not just an effect of the substance being studied. If small rodents are not capable of feeding, they will die within hours. It is highly likely that many substances would not be toxic if a simple sugar solution were injected. Similarly, loss of body heat probably contributes to the lethality of many treatments. This general concern can be translated to NP.

Post hoc definition of the experimental hypothesis. It is quite common, especially in basic research, to change a hypothesis according to the results obtained to 'tell a story' in a publication; few investigators and readers are aware of the impact this has on the statistics.

Standardization versus tradition. Although standardization is a prerequisite for mutual acceptance of data, especially in the regulatory context (often to accumulate sufficient knowledge of the model itself), it needs to be distinguished from the 'we have always made it like this' approach (tradition). This means that models should be modified if there is strong reason for it, but this modification needs to be documented and the information disseminated to create a new standard.

Proprietary data. Data on animal tests generally are not shared, especially in the regulatory context, seriously limiting the science. This lack of sharing is often due to commercial interests (transparency toward competitors, barriers to make use of the data for their developments), but sometimes it is just the hassle of and lack of incentive for publication.

Face-validity versus validation. We tend to believe in the relevance of animal tests if some features resemble our expectations; in contrast, validation means to systematically challenge the relevance of a model. The performance characteristics (reproducibility, limit of detection, relevance of results, and mechanistic basis) are often not subject to systematic assessment. Most animal models have never been validated, not even for the classic test materials. It is very unlikely that for NP, where even less human data are available for comparison, this should now happen.

So we are living in an imperfect world, working with imperfect tools. What are the prospects of producing relevant science, protecting the consumer, and identifying new agents with a high probability? First, we need to understand the limitations of the tools. Where possible, this requires optimization,

standardization, and validation, which can lead us to an evidence-based approach.[8] Second, we must not rely on one tool only; the combination of tools with complementary strengths, such as a sensitive and a specific test, can overcome weaknesses. Integrated testing, especially using different technologies, is most promising. However, the tools to compose and validate testing strategies are only now emerging. Third, we must stop believing and start doubting. There is a tendency to believe *a priori* in the relevance of animal models ('*in vivo veritas*'). This is further augmented by the desire to come to a definitive conclusion as a result of economic or publication pressures. Calling for *in vivo* studies[52] for NP alone will not solve our information needs—there is no reason to assume that they will perform better for NP than for traditional chemicals and, in fact, there is some reason to believe that they perform even worse for NP.

THE NEED FOR NANOTOXICOLOGY AND ALTERNATIVE APPROACHES

It is a fallacy to assume that one can gain beneficial effects in the human organism without increasing the unwanted collateral effects. First, any biologically active agent perturbs physiology, hopefully as a corrective for the patient but at least requiring compensatory reactions of the healthy. Second, few agents are specific enough to have only one effect, but it is rare that we want all the effects and in the given mix of strengths. Third, many agents show excess toxicity—even desired effects often become negative if excessively stimulated. This increase in negative effects is directly linked to more sensitive subpopulations (children, elderly, the diseased, those with genetic polymorphisms, etc.). Thus, the promise of NP's beneficial effects may be paid for in possible side-effects, such as toxicities.[53] For most manufactured NP toxicity data are unavailable, with some exceptions for carbon black, titanium dioxide, iron oxides, and amorphous silica.[54]

The question arises: Do we need to simply apply current toxicology to NP or do we need different approaches? We can subdivide this question into six parts:

1. Do We Have to Expect Hazards that are Irrelevant or Less Relevant for Other Test Substances and Thus Not Typically Addressed?

Yes. Lung toxicology, for example, lags far behind other areas of concern, while NP are especially likely to reach the alveoli of the lung and exert toxicity there.[55,56] Testing for respiratory irritation and sensitization is not standard for industrial chemicals and guideline tests have not even been developed. This is due not only to the effort (costs) involved but also to the poor reflection rodents give of human exposure. We should also be clear that the health effects of interest are continuously expanding: endocrine disruption, immunotoxicity, and developmental neurotoxicity are among the more recent additions not yet reflected in testing programs for industrial chemicals. It would just take one scandal, however, to change this. It may be unlikely that a completely new health effect is induced by NP (like 'the ears fall off'), but there are many human diseases where we do not suspect any chemical involvement where, in fact, it might exist. The chronic effects of chemicals, in particular, are so poorly understood that we have no idea whether we would get any relevant alert from the routine animal tests, which are inadequate even with well-known hazards. Possible examples of continuously increasing health problems include atherosclerosis, male fertility, autism, and diabetes. It is worth noting that air pollution involving natural NP led mainly to excess deaths associated with cardiovascular illness,[57] a hazard not generally addressed in toxicology. The pulmonary and systemic inflammatory hazards typically seen with NP[58] are not among the strengths of the toxicological toolbox.

2. Do We Have to Expect Hazards that Were Not Shown by the Parent Compound from Which the NP Were Made?

Yes— changed kinetics (bioavailability) alone[59] are sufficient to create such situations, as whenever higher plasma or tissue levels of the substance are obtained, thresholds of toxicity might be passed. We know very well from formulations of drugs that solubility after oral administration depends on particle size influencing peak plasma levels—a crucial determinant for toxicity. A number of studies show size-dependent effects.[60] At least for the antimicrobial properties of silver NP, shape dependence has been shown;[61] how this translates to toxic effects on human cells is not known, but cellular uptake of gold NP has been found to be shape-dependent,[62] while others reported no differences in a number of cell systems for silica NP.[63]

Similarly, reactive chemistry, a key feature of many toxicants, is strongly influenced by particle surface. Thus, we might see hazards with NP at lower concentrations than the maximally tested or testable doses currently applied to the parent compound.

We should not underestimate the hazards that are masked by current tests because the animals defend themselves successfully. More than 90% of substances that can exert genotoxicity in cells are not mutagenic in animal tests. This is not to say that the cell result was wrong, but it usually means we did not achieve the concentrations *in vivo* that we can apply *in vitro* or that other defenses were not reflected *in vitro*. The substance may still present a hazard, which might become relevant for humans or subpopulations. We do not know whether the defense mechanisms against some hazards are as effective when administered as NP. However, the new properties of NP can also lead to new biological interactions,[64] which could result in toxicities not shown by the parent compound.

3. Do We Have to Expect Changes in Dose–Response Relationships Relevant for Risk Assessment?

Yes—as argued above, kinetics will be affected, thus shifting concentration response curves, limits, and threshold concentrations, as well as no-effect levels. Thus a critical component of the risk assessment paradigm is altered, and we will need to explore whether this can be handled with, for example, safety or assessment factors. This means requiring, for example, an additional factor as safety margin for the use of NP in humans. Currently we often require one of 10 for interspecies differences and one of 10 for sensitive subpopulations to estimate safe-use doses from animal no-effect levels. Completely new ports of entry of substances have already been described, however, when NP sizes go below barrier cut-offs.[65] Note that the ease of cellular uptake may also result in bioaccumulation of NP.[2] Furthermore, the larger surface area per weight unit often makes NP more reactive. As many toxicities are mediated by chemical reactivity, for example, mutagenicity by formation of DNA adducts or sensitization by hapten binding, this raises special concerns of increased toxic potentials.

4. Are There Reasons to Make Current Tests Less Applicable to NP?

Yes—this holds true, unfortunately, for all approaches. The biokinetics of NP will affect *in vivo* and *in vitro* results very differently. NP aggregation and the difficulty of application to cells and animals affect the execution of routine tests. Databases allowing computational approaches are so rare that, for the immediate future, no major contribution can be expected.

5. Are There Reasons (in the Case of NP) to Question the Extrapolation of the Parent Compound to Humans?

Possibly—it is difficult to say whether animal results for NP are better or worse when extrapolated to humans. NP differ in exposure/kinetics, and with regard to hazards, even higher potentials for interspecies differences exist.

6. Is the Limit of Test Throughput More Relevant for NP?

Yes—first of all, because the main health concerns in particle toxicity hint at risks for the most complex endpoints (cancer, lung toxicity), which require the most test capacities. Rodent inhalation models are especially prohibitive in terms of time and expense.[66] Second, as any given substance can be at least theoretically formulated to particles of very different sizes, size distributions, shapes, and modifications, an almost unlimited testing demand could be considered. Choi et al.[67] calculated the costs for testing the NP already in the market to between $250 million and $1.2 billion and the time required at 34–53 years.

In addition, we should be aware that current regulatory toxicology was established for drugs under development. A very precautionary approach was taken to avoid putting volunteers and patients at risk.[37] While this might be appropriate for nanomedicine products, we have to ask ourselves what development opportunities we sacrifice if we apply the same precautionary methods to products with lesser exposure to humans or those that are not intended to be biologically active. The problem becomes most pronounced when precautionary methods (many false-positives) are used for substance groups with rare side-effects.[51]

This reasoning makes evident the need to explore the toxic profiles of a broad variety of NP to learn what must be control for. Furthermore, novel approaches need to be developed, as traditional approaches might have even greater limitations for NP than for other industrial chemicals and may not offer the throughput and velocity to cope with the dynamic developments in nanotechnologies.

GENERAL LIMITATIONS OF *IN VITRO* SYSTEMS

Cell culture technologies form the basis of most alternative methods. They have matured over the last decades but severe limitations remain.[47] The critical question is how well the cell in culture reflects the cell in the organism, especially with regard to its state of

differentiation and its response patterns in isolation in an artificial environment. Differentiation can be seen as the cellular equivalent of the phenotype at the level of an organism: Which genes are expressed at a given time point and what functionality do they allow? This critical appraisal of the problems of current cell culture techniques does not mean that *in vitro* systems are not useful, but it does stress the need to critically evaluate the predictive value of these models in each and every case. This emphasizes the need for validation, either in a formal sense for standardized tests or as a quality assurance process for the *in vitro* models in research and cell-based screening approaches.

Over-optimal growth conditions. An organism works highly efficiently. Cell mass that is not required is removed. Atrophy is the reduction of mass when no continuous request for function is sensed. There is some evidence in many toxicity tests that something similar occurs in cell culture: a common observation is that cell vitality increases at the beginning of a concentration/response curve, that is, a little bit of toxin keeps cell numbers higher than in untreated controls. This cell viability phenomenon might well translate to specific cell functions. The fact that cell functions are not requested during culture, such as metabolism of xenobiotics, might be much more a driving force for dedifferentiation than commonly acknowledged.

Driving cells into growth and division results in less competent cells. Shorter doubling times mean less time and work before the next experiment. By adding growth factors or fetal calf serum, we force cells into proliferation. How can we expect the cell to express its specific functions when signals drive it toward multiplication, not toward function?

Nonhomeostatic culture conditions. The organism maintains a tightly controlled, stable environment for most cells—the composition of blood and extracellular fluids is intensively regulated. In contrast, cell cultures experience dramatic changes—at the time of culture medium change, nutrients are exhausted and waste has accumulated. From one moment to the next, they are covered by fresh medium, often with associated temperature and pH changes. Unavoidably, this treatment favors cells that are able to adapt, not specialized cells maintaining sophisticated functions.

Lack of oxygen. Standard test systems supply only oxygen dissolved in the culture medium. At typical cell densities this is exhausted after a few hours. Further supply by equilibration with air in the incubator is low because of the limited surface area, the high water column of culture medium (which limits diffusion), and the lack of air stream. The result is anaerobic metabolism with accumulation of lactate—the main reason for acidification of the medium. Anaerobic cells cannot be expected to show normal physiological reactions.

Disproportion of cells versus culture media. Cells in culture achieve maximally 1% of the cell density in organs, because dissolved oxygen and nutrients need to be added. The need for nutrients is also increased dramatically by their inefficient, anaerobic use. Therefore, owing to the dilution in medium, cell contacts are limited and secreted substances, such as signal molecules, are strongly diluted. There is another problem linked to this, which we often overlook: not all effects of compounds depend on their concentration only; some also depend on absolute amounts relative to the cell number (especially when cells accumulate the substance). Thus, large quantities of substance added with the large amount of media necessary will overload the cells and mislead the interpretation of concentration/response curves.

The cancer origin of cell lines. The origin of cell lines is normally either a cancer tissue or primary cells that have undergone a similar transformation in culture. Tumor cells have been shown to have undergone tens of thousands of mutations compared with their progenitors (for review see Refs 68 and 69). They have drastically reduced the expression of organotypic function and favored cell growth and division over other cell functions. They often show chromosomal aberrations and losses. Karyotyping, that is, the control of the chromosomal setup of cell lines, regularly leads to big surprises. Few laboratories carry this out regularly and over many passages of cells. It is not possible to expect the cells to resemble normal tissue when their origin is already far from normal or when severe transformation and selection occurred.

Cell authenticity and cross-contamination. This problem has been addressed earlier in the context of developing a Good Cell Culture Practice.[5,70] It remains a grave concern, and has been put forward more recently.[71–74] There is compelling evidence that 15–20% of the cells used in experiments are not what they are supposed to be. A main reason for this is overgrowth with HeLa cells. Cross-contaminating any cell line with a single HeLa cell will result after a while in overgrowth by HeLa, which often takes place unnoticed. Buehring et al.[73] identified 220 scientific papers where cell line results were reported with cells known today to be nothing but HeLa cells. In a survey involving more than 400 researchers, 9% were unknowingly using HeLa contaminants. MacLeod et al.[72] analyzed cells submitted to cell culture banks—a situation where researchers will be most certain about the authenticity of their cells—and 18% of these were cross-contaminated, often even with cells from other species. These cell lines had

already been used for hundreds of papers and will most probably continue to be used. Action is required, as called for again recently by Nardone.[74]

DMSO and other additives. We often deal with substances that are not soluble in water, and we cannot mimic the uptake and transport of the substance in the organism. The amount of solvent to be used is often borderline toxic itself, with uncontrollable cocktail effects with the compound under study. Another concern with regard to additives is the prophylactic use of antibiotics in cell culture and their effects on cells. The influence of antibiotics on cell morphology, cellular degeneration, cell death, and cellular function was summarized by Kuhlmann.[75,76] We often forget that these compounds, added at high concentrations (up to 100 μg/mL) and interfering, for example, with protein biosynthesis also of mammalian cells, affect our test systems. Last, but not least, is the addition of serum—typically 10% and from a different species—hardly a standardizable, complex biological product and not even added at relevant concentrations. The binding of compounds to serum proteins and the fact that only the free portion in this equilibrium is available to act on its cellular target is common knowledge to every pharmacologist. However, how do we interpret that 10% binding capacity? Astonishingly, higher serum conditions are not well tolerated by most cells and, even after complement inactivation, human serum appears to be especially toxic even to human cells. A sufficient explanation for this phenomenon is not yet available.

Lack of activating metabolism. Technical solutions to increase metabolic competence (viral transfection, stably transfected cell lines, maintenance of metabolic capacity of primary cells, cells from stem cells, etc.) are advancing, albeit slowly. The contribution of metabolism to the toxicity of general chemicals is not really known; it is probably overestimated because of the predominance of drugs in the literature and a publication bias toward toxic compounds and complex mechanisms. Often the biological effect of the metabolite is stronger than that of the parent molecule but does not represent a completely new quality; thus, the hazard can also be picked up from the parent compound, though at misleadingly high concentrations. The species differences in metabolism are well known; in most cases metabolism is protective (which is why it represented a competitive evolutionary advantage). It is thus quite likely that we miss hazards by using animals, which do not show the activating metabolism of humans.

Still, metabolism is one of the key problems for alternative methods.[77] The role it plays for NP toxicity has not yet been established.

Lack of defense mechanisms. Baby cells, heavily mutated by the loss of many genes and not trained to react to a toxic attack (see above)—this is a scenario of vulnerability. Indeed, we often observe that cell lines are much more sensitive than primary cells. We can now take either of two positions, that is, praise this as precautionary or blame it as an overestimation of toxicity. We clearly favor the first view because the most common purpose of such tests is hazard identification and classification and labeling. If the cell studied could protect itself against the toxicant successfully within the organism but its descendant in culture cannot, this might indicate that the parent cell may be chronically overwhelmed or that there are other tissues and circumstances where the protective measures would fail. Better to know and flag an alert.

The standardization of culture conditions. Why do we culture murine cells at 37°C? Body temperatures of rodents are several degrees higher. Why do we force a pH of 7 while inflamed tissue can go down to a pH of 4? (The tight pH control mentioned above holds true for blood only.) Why do we use the osmotic strength of human blood in our cell culture media? Rodents have a higher salt concentration. It is all a matter of convenience and common practice. However, neither ensures the most physiological response or predictive test outcome.

Lack of quality assurance in cell culture. We have elaborated on this before[5,78] in the context of developing the Good Cell Culture Practice guidance. To some extent all previous points discussed here give further reason for quality assurance. In systems that are highly artificial and reductionistic, not only the quality of procedures but also the predictive value of the results needs to be assured. In basic research, at least, this is not common practice.

The lack of standard protocols. Most method sections are far too brief and heterogeneous. What we typically report in a scientific journal is not sufficient for another party to redo the cell culture experiment.

In conclusion, these reflections encourage considering the shortcomings of cell culture systems when applying them to NP. This should prompt more intense quality assurance, also in the context of an evidence-based toxicology.[45]

SPECIAL PROBLEMS FOR *IN VITRO* NANOTOXICOLOGY

A number of articles addressing problems of *in vitro* nanotoxicology are available.[66,79–82] Below a number of common themes are summarized.

Agglomeration and aggregation. Not everything called nano is actually nano. Aggregation or

agglomeration of NP (aggregates being very strongly boundwhile agglomerates can still be dissociated) is very common and difficult to prevent. NP can have complex aggregation behavior in aqueous solutions[59] with substantial impact on their toxicity. Many of the studies published so far did not exclude agglomeration/aggregation, but even as nonmonodispersed particles, the smaller particles are more potent in many respects.[2] Agglomeration/aggregation effects have also been recognized in the ecotoxicity of silica, titanium dioxide, and zinc oxide NP.[83] Some systematic approaches to dispersed NP have been proposed,[84] but the problem still needs to be addressed on a case-by-case basis.

Stability. The stability of NP is not often discussed, but the sheer surface area represents a problem, as it not only attracts substances offering binding sites but also lends itself to chemical reactions such as oxidation. Many NP might actually be coated. We know as little about the modification and degradation occurring over time as we do about the metabolic fate of NP.

Artifacts. Many artifacts in cell culture are because of the enormous surface of NP, which not only allows absorption of many components but also augments chemical reactivity. The abundant serum albumin from fetal calf serum of most cell cultures is likely coating many NP; similarly, in the lung surfactant proteins would cover NP or mucus in the gastrointestinal tract, but might not be able in the presence of albumin. Anecdotal observations have shown that NP can absorb many dyes, enzymes, or inflammatory mediators thereby influencing test results. They can carry pyrogenic contaminations, difficult to detect but biologically strongly active.[85] Single-walled carbon nanotubes have been shown to interact with MTT, commonly used to test cell viability.[86]

Dosimetry. In toxicology, we have seen a move from primarily weight-based doses (mg/kg or ppm) to (molar) concentrations, especially when kinetic measures (plasma concentrations, for example) could be assessed or when experiments were done *in vitro*. For NP, weight, particle number, and surface area are typical dose measures, but shape, coating, and electrophysical properties also can have a further impact. The chemical characterization[87] and dosimetry clearly require closer attention than for traditional chemicals.[64]

In some cases, toxicity correlated best with NP surface area,[88] but it has still to be established whether this is a more general rule. It makes sense for reactive chemistry, which is a leading mechanism of toxicological damage, and for oxidative processes; in fact, generation of reactive oxygen species has been a key mode of action associated with NP toxicity.

In vitro biokinetics. This term has been coined to indicate that test substances *in vitro* also exhibit kinetics: they are adsorbed (e.g., by plastic or the albumin of fetal calf serum), stay soluble or precipitate, are taken up by cells, are oxidized by air or metabolized by the cells, and we are interfering with their presence when changing cell culture media. The situation is not as complex as *in vivo* kinetics, but certainly the actual effective concentration reaching the cells is not the one we added. Paracelsus' paradigm holds true also *in vitro*—the presence or absence of toxicity can be understood only if the effective concentration of the dose applied is understood. Just as *in vivo* work has been augmented by introducing kinetics, we might likewise give consideration to these factors as we move the field of *in vitro* toxicology forward.[89] The situation is not less complex for NP, where aggregation and particle coating must be added. A major distinction are certainly soluble, partly soluble and nonsoluble NP; while the first introduce mainly an additional kinetic component of dissolution for cell exposure, the situation can become very complex when both NP and solved material are present in cell culture. Cell membranes, mitochondria, and nuclei are considered major compartments for NP toxicity.[88] Thus, uptake and intracellular trafficking must also be addressed. There is increasing literature on cellular uptake of NP[90] in part stimulated by drug targeting, that is, the targeted delivery of drugs to their desired site of action, using NP.[91] Uptake of NP into cells appears to be an effect specific for materials in the range of 50–200 nm. Surface modifications (positively or negatively charged side groups of the polymers, amino acids, or peptides/proteins) enhance this uptake. It is not clear to date, which types of particles are taken up by pinocytosis, receptor-mediated endoytosis or phagocytosis, respectively, to mention the most prominent uptake routes, which might cooperate at least in some instances. Research about factors influencing cellular uptake, like size, surface properties, cell type, and endocytotic pathways will be critical to interpret *in vitro* toxicity data and the suitability of test systems.

Cell contact of NP. Actual exposure of cells to NP needs to be assured, as NP might swim on the culture media. Also, NP is known to pass from cell to cell. Cell monolayers resemble pan-fried eggs lying next to each other, giving only minimal cell-to-cell contact areas. Furthermore, cell density in a typical culture is only 1% of normal tissue, which changes dosimetry and the likelihood of NP-to-cell contact.

EXISTING (VALIDATED) ALTERNATIVES AND THEIR SUITABILITY FOR NANOTOXICOLOGY

It is disturbing that nanotoxicology is reinventing alternative approaches, often without bridging the two decades of development and validation done for chemicals and cosmetic ingredients. Cytotoxicity and mutagenicity assays are broadly used[66,81] without necessarily bridging the validated methodologies. Others have highlighted the need to optimize validated toxicity and ecotoxicity tests for NP.[2] A variety of approaches lend themselves to adaptation for NP (Tables 1 and 2), but none has been formally validated for NP.

Skin corrosion. The skin corrosion endpoint for chemicals has been replaced, following ECVAM validation, by seven validated and accepted methods (transepithelial electrical resistance test with rat skin, Corrositex™, and five skin models). It is certainly critical that the US accepts only positive results and requires animal testing for the negative ones. A long-planned study to address false negative skin corrosives in the novel alternative skin irritant tests is essential to demonstrate the safety of the combined use for US acceptance. The methods primarily in use for skin corrosion require commercial, artificial human skin. To safeguard availability and free market, new products (similar methods, sometimes also referred to as 'me-too' methods) require assessment of equivalence. The endpoint seems to be of lesser importance for NP, but given the augmented reactivity of NP due to large surface area, NP from irritant substances might actually become corrosive.

Skin irritation. The endpoint has been replaced for chemicals by a successful validation in 2007 with inclusion in REACH test strategies. Test guideline acceptance succeeded in the EU and has just been completed at the OECD level. A number of me-too developments are ongoing. This is important to maintain competition in the market and to gain acceptance. Research activities using omics technologies are being considered to reduce the false positive rate in future test strategies. The methods appear to be suitable for NP but have not been formally evaluated.

Skin absorption. The method for this endpoint has been accepted by the OECD, as a political compromise, in parallel to the creation of an *in vivo* method. The methods are certainly key for adaptation to NP.

Eye irritation. This endpoint might, in fact, be considered the 'holy grail' of alternative methods because of its emotional impact on many animal lovers. Various assays are being evaluated, which, combined in a test strategy, might have the potential to replace the animal test for this endpoint.[92] The strategy aims at utilizing the strengths of specific *in vitro* assays to address required ranges of irritation and/or chemical classes (applicability domains) to classify test substances. Around 20 methods have been in industrial use for more than two decades, and six large validation trials have evaluated them. However, no single method was found to fully replace the complexity of the animal responses. ICCVAM and ECVAM have shared the work of collecting and compiling data for the ten most promising methods (four methods by ICCVAM, six methods by ECVAM). Test strategy development, based on the strategies identified in an ECVAM workshop held in 2005,[92] is supported by COLIPA and the Integrated Project OSIRIS (sponsored by the EU). The two most promising methods based on reconstituted human corneal epithelial cells are currently starting validation with the support of COLIPA. These models appear promising for the identification of nonirritant substances like the *in vivo* application with a large applicability domain. It appears that the various methods can be highly relevant for NP, but they have not been evaluated to larger extent.

Acute toxicity. This test could probably be considered obsolete—in the pharmaceutical arena, efforts to abandon the test (led by the UK National Centre for the Replacement, Refinement and Reduction of Animals in Research (NC3Rs)) are ongoing because other preliminary single-dose tolerability studies and/or dose range-finding data are sufficient. An analysis of the existing and new chemical databases to support such a strategy was recently published.[93] It demonstrates the potential use of NOAEL (no observable adverse effect level) data from repeated dose to extrapolate the LD_{50}. A number of activities are ongoing.[94] A validation study between ICCVAM and ECVAM to predict acute toxicity by cytotoxicity assays was completed in 2005. The study confirmed previous findings suggesting that the *in vitro* tests could identify nontoxic chemicals. These make up 80% of new chemicals (EU New Chemical Database) and about 70% of old chemicals (IUCLID). A follow-up study to assess this hypothesis was commissioned in 2007 to test about 60 additional chemicals, with promising results (to be published). Furthermore, the test was established on the robotized testing facility at the JRC (a collaboration between the Nanotechnology and Molecular Imaging unit and ECVAM), where additional substances are now being tested. The Integrated Project ACuteTox was initiated following an ECVAM workshop.[95] The project aimed to develop and pre-validate testing strategies and ended mid-2010. Two ECVAM-validated assays

TABLE 1 | Status of Scientific Work, Validation Activities, and Regulatory Acceptance of Alternative Methods for the Main Toxicological Endpoints Relevant for NP

Toxicological Endpoint	Ongoing R&D	Ongoing Validation	ESAC Statement	EU Regulatory Acceptance	OECD Test Guideline
Skin corrosion		Similar tests (human skin models)	Human skin models: Episkin (1998), EpiDerm (2000), SkinEthic (2006), EST-1000 (2009) TER (1998) 1997 Corrositex[1]	EU Test Method B.40 (2000)	TG 430 and TG 431 (additional data needed for the confirmation of noncorrosives)
Acute phototoxicity		Integrated test strategy with 3T3 NRU-PT and human skin model to reduce 'false' positive rates	1997 (3T3 NRU-PT)	EU Test Method B.41 (2000)	TG 432
Skin absorption/ penetration		Submission of dossier awaited, methods already in use without formal validation		EU Test Method B.45 (2008)	TG 428 (Excised pig or human skin in flow-through or static diffusion cells)
Skin irritation	Follow-up: omics approach to identify new markers for better discrimination of mild irritants (GHS)	Several similar tests	Human skin model: Episkin (2007), EpiDerm (2008), SkinEthic RHE (2008)	EU Test Method B.46 (2008)	TG 439
Eye irritation	Mainly COLIPA sponsored research for the development of more mechanistic assays	Two reconstituted human tissue models, Irritection, Slug Mucosal test	Organotypic models: BCOP and ICE to identify severe eye irritants (2007)[4] Additional work needed for the HET-CAM and IRE Cell-based assays: Fluorescein Leakage and Microphysiometer test (2009) Refinement: Low Volume Eye Test (2009)[3]	TG for BCOP and ICE accepted (2009)	TG 437 (BCOP) and 438 (ICE)
Genotoxicity/ mutagenicity	Reduction of the false positives in *in vitro* tests Omission of positive/negative controls in *in vivo* genotoxicity	Genotoxicity tests in 3D human skin models COMET assay (Japan)	Micronucleus test *in vitro* (2006)[2]	MNT *in vitro* is mentioned in Annex VIII of regulation (EC) 1907/2006 (REACH)	TG 487 on MNT *in vitro*
Acute oral toxicity		Acute-Tox IP Predictive capacity of 3T3 NRU cytotox. assay to classify nontoxic substances	ATC, UDP, FDP (2007)[4]	ATC, UDP, FDP (2001)	TG 420, TG 423, TG 425

TABLE 1 | Continued

Toxicological Endpoint	Ongoing R&D	Ongoing Validation	ESAC Statement	EU Regulatory Acceptance	OECD Test Guideline
Acute inhalation toxicity		ATC[3,4] and acute inhalation toxicity revised[3,4]		Foreseen 2010	TG 436 and 403
Carcinogenicity			Two Cell Transformation Assays under peer-review		TG 451
Sensitization	Various methods under development	Three methods: peptide binding assay, two monocyte activation tests	Local lymphnode assay (1999), reduced variant (2007)	EU Test Method B.42 (2004)	TG 429;
Ecotoxicty	Fish cell tests		Acute aquatic toxicity reduction method (2006)[4]	In REACH guidance, but rejected for EU test regulation	guidance document 126
Pyrogenicity			Five Monocyte Activation Tests (2005)	European Pharmacopoeia (2009)	Not applicable; FDA (2009)

NRU-PT, neutral red uptake phototoxicity test; TG, test guideline; BCOP, bovine corneal opacity and permeability assay; ICE, isolated chicken egg; ATC, acute toxic class method; FDP, fixed dose procedure; UDP, up and down procedure; IP, integrated project.
[1]Corrositex is only applicable to acids and basis and their derivatives.
[2]The Micronucleus test *in vitro* is part of a test strategy for genotoxicity/mutagenicity. Several nonvalidated *in vitro* tests area accepted as OECD TG.
[3]Only refinement methods.
[4]Only reduction methods.

(CFU-GM and pyrogen tests) within the ACuteTox project were shown to predict acute toxicity much better (>85%) than the earlier cytotoxicity tests. Additional testing of these assays is ongoing. Cytotoxicity assays will represent a key tool for screening NP and identifying novel toxic properties.

Genotoxicity. Several accepted *in vitro* tests exist (also as OECD test guidelines), but they often identify substances as false positive, which then need to be sorted out using animal tests. Various *in vitro* assays are available; the problem, however, is the low quality of these nonvalidated tests, particularly their over-sensitivity. When three such tests are combined, as is typical, only 3% of the positive results are correct while 97% are negative in the subsequent animal test.[96] ECVAM held a workshop on the reduction of false positives in *in vitro* genotoxicity tests[97] and the recommendations are being followed with funding from COLIPA, NC3Rs, and ECVAM. Validation of skin models for genotoxicity testing is co-sponsored by COLIPA and ECVAM.

The *in vitro* micronucleus test (MNT) was validated in 2006. This is a remarkable test case for various reasons: it was the first completely retrospective validation ever, as suggested by the Modular Approach to Validation[98] and it made use of a compilation of existing data from various sources, instead of designing a new validation study. In fact, the validation was completed in roughly half the time of a prospective study and with minimal costs. The MNT validation also represents the new close collaboration with OECD, as test guideline development was carried out in parallel to validation. Thus, consensus at the OECD level could already be reached in 2008. It is, however, only an *in vitro* test replacing another *in vitro* test of lower quality.

JaCVAM is leading the validation of the COMET assay with EU and US involvement. Both the MNT and the COMET assays promise to be more predictive. More importantly, however, they can be incorporated into the repeated-dose animal studies, making a further mutagenicity test unnecessary. Other options not formally validated include reducing the number of animals in *in vivo* genotoxicity testing, for example, the use of only one gender, the omission of positive controls. New DNA repair-based assays were submitted or identified and are currently under discussion.

The question might be raised, however, whether mutagenicity in human cells should be ruled out at all by an animal test. A genotoxic effect *in vitro* shows that the substance has a potentially hazardous property. Differences in the *in vivo* test can be either

TABLE 2 | ESAC Statements on Alternative Methods Relevant for NP

No.	Method	Date of ESAC Statement	ECVAM Prospective Validation Study?	Impact on 3Rs	ATLA Reference
1	3T3 NRU phototoxicity test	03/11/1997	Yes	Replacement OECD TG 432	ATLA 26 (1), 7–8
2	EpiSkin skin corrosivity test	03/04/1998	Yes	Replacement (EU) OECD TG 431	ATLA 26 (3), 275–280
3	Rat TER skin corrosivity test	03/04/1998	Yes	Replacement (EU) OECD TG 430	ATLA 26 (3), 275–280
4	Application of the 3T3 NRU phototoxicity test to UV filter chemicals	20/05/1998	Yes	Replacement OECD TG 432	ATLA 26 (4), 383–386
6	Local lymph node assay for skin sensitization	21/03/2000	No	Reduction / Refinement OECD TG 429	ATLA 28 (3), 365–367
7	EpiDerm skin corrosivity test	21/03/2000	Yes	Replacement (EU) OECD TG 431; Annex V TG B.42	ATLA 28 (3), 365–367
8	CORROSITEX skin corrosivity test	06/12/2000	Yes	Reduction OECD Draft TG 435	ATLA 29 (2), 93–97
11	Micromass embryotoxicity assay	01/05/2002	Yes	Reduction	ATLA 30 (3), 265–273
12	Whole rat embryotoxicity assay	01/05/2002	Yes	Reduction	ATLA 30 (3), 265–273
13	Embryonic stem cell test for embryotoxicity	01/05/2002	Yes	Reduction OECD Draft GD 43	ATLA 30 (3), 265–273
17	Upper Threshold Concentration (UTC) step-down approach for acute aquatic toxicity testing	21/03/2006	Yes	Reduction OECD guidance document 126	ATLA 35 (2), 199–208
18	CFU-GM assay for predicting acute neutropenia in humans	21/03/2006	Yes	Reduction Submitted to EMEA	ATLA 35 (2), 199–208
19	Human Whole Blood IL-1 for *in vitro* pyrogenicity testing	21/03/2006	Yes	Replacement European Pharmacopoeia monograph	ATLA 35 (2), 199–208
20	Human Whole Blood IL-6 for *in vitro* pyrogenicity testing	21/03/2006	Yes	Replacement European Pharmacopoeia monograph	ATLA 35 (2), 199–208
21	PBMC IL-6 for *in vitro* pyrogenicity testing	21/03/2006	Yes	Replacement European Pharmacopoeia monograph	ATLA 35 (2), 199–208
22	MM6 IL-6 for *in vitro* pyrogenicity testing	21/03/2006	Yes	Replacement European Pharmacopoeia monograph	ATLA 35 (2), 199–208
23	Human Cryopreserved Whole Blood IL-1 for *in vitro* pyrogenicity testing	21/03/2006	Yes	Replacement European Pharmacopoeia monograph	ATLA 35 (2), 199–208
24	*In vitro* micronucleus test as an alternative to the *in vitro* chromosome aberration assay for genotoxicity testing	17/11/2006	Yes	Enhancement of *in vitro* test battery, OECD TG 487 Guideline 827	ATLA 35 (2), 199–208

TABLE 2 | Continued

No.	Method	Date of ESAC Statement	ECVAM Prospective Validation Study?	Impact on 3Rs	ATLA Reference
25	Application of the SkinEthic human skin model for skin corrosivity testing	17/11/2006	No	Replacement (EU) OECD TG 431	ATLA 35 (2), 199–208
27	Bovine Corneal Opacity and Permeability (BCOP) test method	27/04/2007	No	Reduction OECD TG 437	ATLA 35 (3), 303–312
28	Isolated Chicken Eye (ICE) test method	27/04/2007	No	Reduction OECD TG 438	ATLA 35 (3), 303–312
29	Reduced Local Lymph Node Assay (rLLNA)	27/04/2007	Yes	Reduction OECD TG 429	ATLA 35 (3), 303–312
30	EpiDerm (with MTT reduction) for skin irritation	27/04/2007	Yes	Reduction OECD TG 439	ATLA 35 (3), 303–312
31	EPISKIN (with MTT reduction) for skin irritation	27/04/2007	Yes	Replacement OECD TG 439	ATLA 35 (3), 303–312
32	Fixed dose procedure (FDP)	31/10/2007	No	OECD TG 420	ATLA 36(1), 12–13
33	Acute Toxic Class Method (ATC)	31/10/2007	No	OECD TG 423	ATLA 36(1), 12–13
34	Up and Down procedure (UDP)	31/10/2007	No	OECD TG 425	ATLA 36(1), 12–13
35	EpiDerm SIT (with MTT reduction) for skin irritation	5/11/2008	No	OECD TG 439	
36	SkinEthic (with MTT reduction) for skin irritation	5/11/2008	No	OECD TG 439	
37	CellSystems human skin model EST-1000 for skin corrosivity testing	12/06/2009	Yes	Replacement (EU) OECD TG 431	
38	Cellsensor Microphysiometer for eye irritation	10/07/2009	No		
39	Fluorescence Leakage Assay for severe eye irritation	10/07/2009	No		

Noteworthy, methods 18–23 were not developed for the purpose of chemicals testing, but current validation activities explore their use for acute toxicity testing, which might lead to an extension of the applicability domain. Validity statements not listed are not relevant for NP.The statements 5, 9, 10, 14–16, 26 that are not relevant for NP are not listed. See Table 1 for abbreviations. TER, transepithelial resistance.

species-specific (rat vs. human) or due to kinetics (not reaching the tissue at sufficiently high concentrations). These do not necessarily rule out a hazard for humans, especially in chronic situations or for hypersensitive individuals. The animal experiment may possibly hide a human hazard.

For NP, mutagenicity tests are certainly first line tests and need proper adaptation. The putative carcinogenic potential of nanofibers is mainly attributed to inflammatory effects and is less likely to be covered by these genotoxicity tests.

Photogenotoxicity. This rarely tested endpoint could not be properly addressed yet, but promising approaches are being followed by COLIPA and others. It might be of special interest for NP used in sunscreens.

Phototoxicity. The endpoint has been replaced by a validated method for which minor follow-up activities are required. However, the validated method is considered overly sensitive. Work using skin models instead or together with cell cultures in a tiered testing strategy are on the way from companies and the Centre

for Documentation and Evaluation of Alternatives to Animal Experiments (ZEBET). It appears that the tests can be adapted to NP but have lesser priority for the risk assessment process.

Skin sensitization. The Local Lymph Node Assay, an ICCVAM-validated refinement method, was scientifically endorsed by the ECVAM Scientific Advisory Committee (ESAC) and is now an OECD-adopted Test Guideline (TG 429). Nonradioactive variants have been adopted.[99]

A number of relevant workshops have been organized[100–103] to identify the most promising methods and to give advise on how to progress toward their validation (peptide-binding assay, skin permeability, and dendritic cell based methods). Three promising methods, the peptide-binding assay, the U937 test, and the hCLAT test, are currently under validation with help from COLIPA.

The Integrated Project Sens-it-iv was set up (http://www.sens-it-iv.eu/) following an ECVAM workshop.[100] Four different cell systems and four different markers have shown promising results. These are to be standardized and assessed in an inter-laboratory trial. Beside Sens-it-iv, the general interest in developing relevant alternative methods for skin sensitization testing is supported by industry's large investment in this area.

For NP, the sensitization endpoint is of high interest as reactive chemistry is the first step in the sensitization process and surface expansion of NP might have an impact here.

Toxicokinetics. Toxicokinetics is best defined as an adjunct technique to support the interpretation of toxicological results in risk assessment. At present, animal-free techniques of prediction of pharmaco- or toxicokinetics are used as a screen in pharmaceutical research. They rely essentially upon the use of physiologically based pharmacokinetic (PBPK) modeling techniques, with input of quantitative parameter values generated by various techniques *in silico, in vitro, ex vivo,* and *in vivo*. This field is still under development, and the use and acceptance of PBPK in risk assessment is increasing. For NP, it appears that kinetics represent a key concern, as unusual behavior of NP have been documented. It is not clear yet how to address this, as modeling might be as misleading as animal models as a result of species differences.

The most promising approach for the moment are barrier models, for example, for the gastrointestinal barrier, skin penetration, blood–brain-barrier, placenta barrier. The exclusion of transfer through these barriers might help to reduce the concern about toxic effects in the respective compartment. This could mean simply no bioavailability, no systemic toxicity testing, or no blood–brain-barrier transfer for central neurotoxicity. For NP these should have highest priority.

Carcinogenicity. ECVAM carried out the validation of three cell transformation assays parallel to OECD test guideline development. The peer review has been recently completed with a positive outcome. The tests have the potential to partially replace the 2-year bioassay in the rat. At up to €1 million per substance, the assay is so expensive that it is rarely carried out (from almost 5,000 new chemicals notified in Europe over the last 25 years, 14 had a cancer bioassay in the new chemical database). Thus, any *in vitro* test would first enable a larger number of substances to be tested. However, the general concerns about the high false-positive rate of the cancer bioassay (some aspects summarized in Hoffmann and Hartung[45]) prompt a need for new approaches. New tests based on the loss of gap-junctions have been proposed, for which pre-validation needs to be initiated. For NP, the cell transformation assays represent a key opportunity to assess some aspects of the carcinogenic potential.

Reproductive toxicity. Because of the complexity of reproductive toxicity and the variety of toxicological mechanisms involved, only a (partial) replacement can be achieved in the meantime by focusing on toxicological targets with high prevalence. This requires setting priority on the relevant target cells and/or biological mechanisms, as the complete mammalian reproductive cycle cannot currently be mimicked *in vitro*. In order not to compromise consumer safety, the use of *in vitro* tests for assessing reproductive toxicity therefore depends on a detailed analysis of historical *in vivo* data of reproductive toxicants belonging to the various chemical groups. In this context, the histopathological data of reproductive organs in repeated dose experiments are of high interest, as this endpoint seems to be very sensitive.

Within FP6 the Integrated Project 'ReProTect'[104] has been established to develop a toolbox of *in vitro* tests assessing various cell functions relevant for mammalian reproduction, such as hormone production, germ cell maturation, uterine function. However, as not all possible target cells/mechanisms are covered within this project, further research activity is necessary.

For NP, this area is of concern, and placenta barrier models might help a lot here. However, as this is an area of massive costs and animal use, it is most important to define strategies for key endpoints such as embryonic stem cell tests for malformations or endocrine disruptor assays to cover key concerns in a screenable manner.

Repeated-dose toxicity. The Integrated Project Predict-IV (http://www.predict-iv.toxi.uni-wuerzburg.de/en/) addresses strategies to improve the assessment of drug safety in the early stage of development and in the late discovery phase by an intelligent combination of nonanimal-based test systems, cell biology, mechanistic toxicology, and *in silico* modeling in a rapid and cost-effective manner. As the assessment does not differ for NP, the project is also fully relevant here. The project will integrate new developments to improve and optimize cell culture models for toxicity testing and to characterize the dynamics and kinetics of cellular responses to toxic effects *in vitro*. It was based on two ECVAM workshops on chronic toxicity and physiology-based pharmacokinetic modeling (PBPK).[89,105]

The FP7 Predict-IV project will address one of the major problems of *in vitro* methods: the absence of data for extrapolations. From *in vitro* studies, NOAELs, serving as starting points for extrapolations, cannot be easily derived, as systemic doses cannot be directly compared to concentrations of a substance applied to cells. In cell systems that best represent *in vivo* target organs, the most predictive endpoints indicative of adverse effects are used to determine the no-observed-effect-concentrations (NOEC) *in vitro*. The NOEC are based on the estimation of measured *in vitro* intracellular concentrations of the drug and/or its relevant metabolites. These NOECs are then transformed to doses received using appropriate modeling techniques, in particular advanced PBPK modeling, including Monte-Carlo techniques. As the model systems are based on human cells and the PBPK-models incorporate human parameters and potential interindividual differences in humans, the need for extrapolations will be reduced and NOAELs can be predicted.

Repeated-dose systemic toxicity requires a new systematic approach in order to move from a traditional, animal-based assessment to a mechanism-based approach. In collaboration between the cosmetic industry and the European Commission, a call for proposals with up to €50 million funding has been launched. It will pioneer the field exploring various new technologies. These 5-year projects started in early 2011. NP are not considered, but translation of results could be most helpful. Repeat-dose toxicity is the cornerstone of risk assessment. It is the animal test most likely to be used for NP assessments. For NP as well as for general chemicals, however, alternative approaches are only emerging.

Ecotoxicology. Activities in ecotoxicology/environmental toxicity are currently focused on short-term (acute) aquatic toxicity and bioconcentration. Endpoints such as long-term (chronic) aquatic toxicity testing or toxicity to birds have not yet been tackled. Changes in legislation, such as environmental assessment of pharmaceuticals and REACH, have increased and will increase the numbers of fish used.

The Threshold Approach [=Upper Threshold Concentration (UTC) step-down approach[106]] for acute aquatic toxicity testing is based on a retrospective analysis of data in the New Chemicals Database carried out by ECVAM and the ECB. Its validity for the reduction of the numbers of fish was endorsed by ESAC in 2006. It is part of the intelligent testing strategy in the requirements for REACH. The Commission submitted a proposal to the OECD, which resulted in a guidance document.

The most promising replacement method is the Fish Embryo Test (FET), for which a lot of data are available and which is already used for effluent testing in Germany. Studies using fish cell lines showed that fish gill cell lines might be useful for acute aquatic toxicity testing.

The endpoint bioconcentration assesses whether a substance accumulates in fish. In a first step, bioconcentration models based on physicochemical properties and physiological parameters (uptake, metabolism, and excretion) are used. If the outcome is above a certain threshold a test in fish becomes mandatory. The current strategy followed by industry and also proposed in the intelligent testing strategy for data requirements under REACH is to improve the existing *in silico* methods by including *in vitro* data on metabolism. Substances that are metabolized do not accumulate in the organism.

For NP, to the best of our knowledge, the adaptation of the respective *in vitro* methods has not yet taken place.

Immunotoxicity. A number of *in vitro* tests have been identified[107] and have undergone some prevalidation,[108,109] Quite a few studies have been done *in vitro* on NP, but their predictive value is not clear.

Pyrogenicity. This might be considered a special form of immunotoxicity; however, it is typically not produced by the test material but by contaminations, mostly bacterial endotoxins. NP represent a most interesting test material because of their large surface area, putatively binding these endotoxins. The Limulus assay, the most prominent alternative pyrogen test, has problems with the testing of solid materials. It remains to be explored whether NP qualify for testing. However, a series of cellular methods based on the activation for cytokine release of blood monocytes or derived cells has been validated[110,111] and accepted for regulatory use. One author has been particularly involved in the development of a whole-blood pyrogen

test.[112,113] This test is unique, as it works with a cell suspension, which might be especially suited to test NP.

IN VITRO NANOTOXICOLOGY RESEARCH—WHICH APPROACH?

In vitro nanotoxicology finds its main justification in the fact that it will make the toxicology of nanomaterials a more scientifically based practice. However, severe gaps exist in our understanding of the cellular and molecular reasons for NP induced toxic effects. This is responsible for slow progress in establishing a scientifically sound and rational basis for the safety assessment of nanomaterials. In this context, mechanistically based information derived from *in vitro* systems have a high priority and certain research needs corresponding to the present status of the problem can be formulated:

Need for testing strategies. For those chemicals for which the toxicological properties of other physical forms are well established, a testing strategy (screening) is needed that will identify whether the NP form will or will not cause significantly different adverse effects. Recent research has established relationships between low exposure to some chemicals (e.g., trace metal compounds) and certain adverse effects, particularly at the level of reproductive, central nervous, cardiovascular and immune systems with the induction of long-term latent such as carcinogenic effects. There is a need for studies on NP on their human target tissues. Thus, *in vitro* studies on NP can involve an integrated approach using a battery of *in vitro* systems as models of the potential different toxic effects (toxicological profile). Moreover, as little information exists on toxic effects induced by NP in different tissues, there is a need for determining basic cytotoxicity by screening tests for main classes of NP. In this context, of great importance are the OECD guidelines concerning the priority of NP to be investigated (fullerenes (C60), single-walled carbon nanotubes (SWCNTs), multiwalled carbon nanotubes (MWCNTs), silver NP, iron NP, carbon black, titanium dioxide, aluminum oxide, cerium oxide, zinc oxide, silicon dioxide, polystyrene, dendrimers, nanoclays).[114]

Need of in vitro biokinetic and metabolic studies. This holds true and is increasingly recognized especially for NP in order to interpret the toxicity data obtained, and for extrapolating this to the *in vivo* situation. In particular, such mechanistic approach (Figure 1) requires the knowledge of the behavior and of the stability of NP in cell-free and complete culture media, along with an assessment of its transport to, and uptake by, the cells in culture, their subcellular distribution and binding to intracellular biomolecules. In addition, the timed-solubility in cellular compartments of accumulated NP is essential in understanding

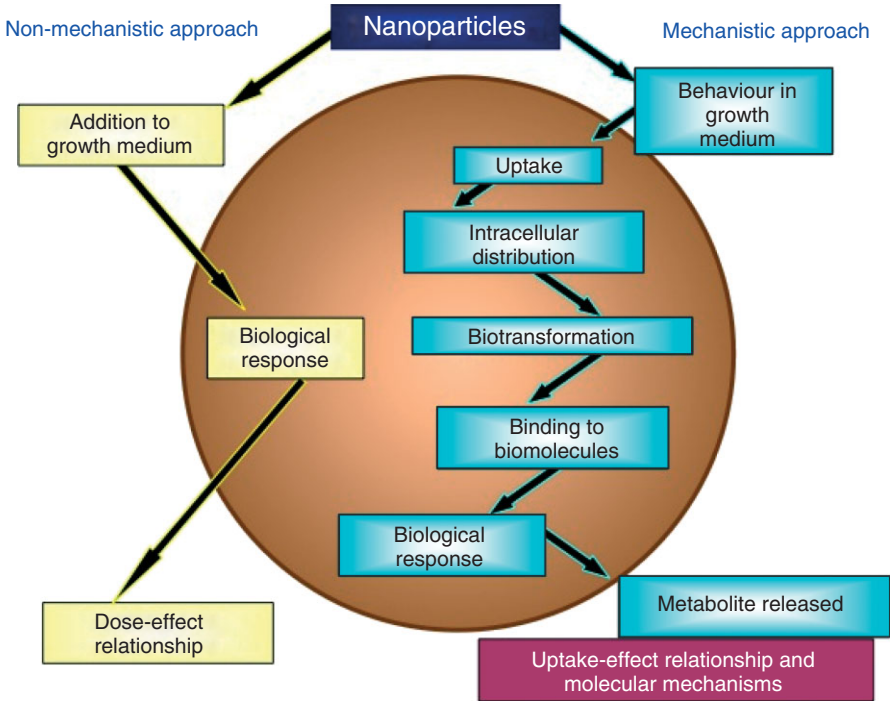


FIGURE 1 | The mechanistic approach to *in vitro* nanotoxicology.[115]

the long-term fate of NP in relation to their latent toxicological effects. Measurement of NP in cultured cells is critical, so that the target concentration needed to yield an effect can be determined. Such measurements are especially important if little or no effect is observed *in vitro* so that false negative results can be avoided. Great advantage can be taken from the possibility to use the *in vitro* systems covering topical and systemic toxicological areas as described above.

Need to study absorption via different routes. More data are needed on how NP enter the lung, GI tract,116,117 skin[118] and blood stream via direct injection for biomedical purpose. First studies are available especially in the context of drug delivery via NP.[119–121] There is a need to expand the studies to the behavior of NP in blood (time for dissolution, distribution between blood cells[122] and serum and binding of serum components[123] and determination of the potential immune toxic effects) and the impact on the reticuloendothelial system.

Figure 2 summarizes a proposed approach to *in vitro* nanotoxicology research.

Characterization of test materials. The underlying mechanisms in *in vitro* nanotoxicology need to be identified in order to anticipate and prevent widespread exposure to potentially harmful nanomaterials. For this to be achieved, detailed physical and chemical characterization of these nanomaterials is imperative. Complete characterization of NP includes measurements such as size and size distribution, shape and other morphological features (e.g., crystalinity, porosity, surface roughness), chemistry of the material, solubility, surface area, state of dispersion, surface chemistry, and other physicochemical properties. Exhaustive characterization of test materials is time consuming, expensive, and complex. To some extent, the characterization required depends on the objectives of the study. It is currently unclear whether existing physicochemical property testing methods are adequate to sufficiently characterize NP. Recently, it has been recalled, that in spite of recommendations from a couple of years ago concerning a minimum set of criteria for sufficient characterization of the investigated NP this requirement is at present not fulfilled.[125] This minimum set of characteristics that should be measured for test materials used in nanotoxicity studies has to be provided by the authors of a publication to give other groups the chance to repeat these experiments and to understand the underlying mechanisms. The list of properties

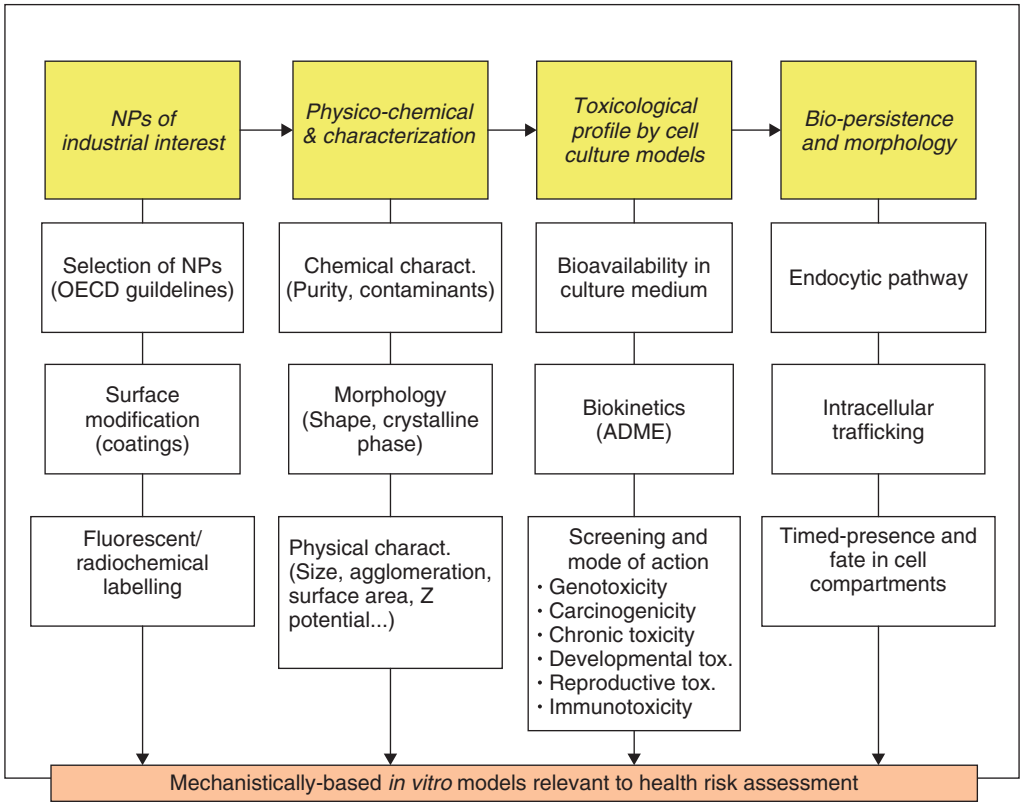


FIGURE 2 | Strategy for *in vitro* nanotoxicology research.[124]

that should be obligatory for the characterization of nanomaterials used in biological experiments includes: chemical composition (purity and contaminants), particle size and size distribution, specific surface area, morphology (crystalline phase, shape), surface chemistry/coatings, extent of agglomeration/aggregation, respective particle size distribution under experimental conditions, water solubility (discriminate between soluble, meta-stable and persistent particles), surface reactivity and/or surface charge (Zeta-potential). It would be a waste of time and resources if the minimum set of criteria or requirements is not available as it is not possible to discuss and repeat the same experiments.[125]

NP toxicity: ions, NP, or both? As many NP are metals or metal oxides following exposure to culture media they may release metal ions, which may then be a direct source of free-radicals, leading to lipid peroxidation and DNA damage. Measurement of the presence of released metal ions is extremely important in *in vitro* nanotoxicology in order to determine if the observed effects are due to the NP itself or the metal ion released by NP dissolution or both. Therefore, dissolution of NP has to be checked in storage medium and in culture media. A study on the effects of silver NP on the photosynthesis of *Chlamydomonas reinhardtii* indicates that the interaction of these NP with the algae influences their toxicity, which is mediated by Ag^+. Silver NP contribute to the toxicity as a source of Ag^+ (about 1% of the silver NP), which is formed in the presence of algae.[126]

Another *in vitro* study on the potential genotoxicity and carcinogenicity of cobalt NP in mouse fibroblasts Balb/3T3 cell line showed a time-dependent increase in Co^{2+} release into culture medium for the concentrations tested, reaching 44% for 100 μM at 72 h. On the basis of the cytotoxicity results obtained on individual cobalt NP and Co^{2+}, the concentration of Co^{2+} did not seem enough to justify the observed cobalt nanotoxicity, not excluding, however, a complementary and/or a synergic effect of the Co^{2+}.[127] Thus, *in vitro* nanotoxicity studies must take into account the exposure of the *in vitro* system not only to NP but in parallel to the corresponding ionic species that could be potentially released into the growth medium. In the same study, cobalt NP showed cytotoxicity and a dose-dependent statistically significant increase in the morphological neoplastic transformation in two different Balb/3T3 clones (Figure 3). Cobalt chloride induced cytotoxicity but did not show carcinogenic potential either for the dose dependency, or for the single concentration values versus control. Moreover genotoxicity assays (Comet and MN) show different

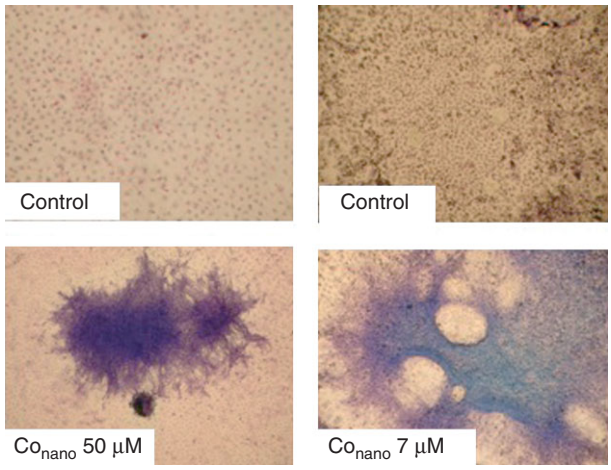


FIGURE 3 | Foci type III induced by Co nanopaticles into two clones of Balb/3T3 cell line. Concurrent cytotoxicity and neoplastic morphological transformation assays on two clones of Balb/3T3 cell line exposed to cobalt NP.[128] Left: clone HRI (Hatano Research Institute, Ochiai, Hadano, Kanagawa 257, Japan). Right: clone ECVAM (supplied by Laboratorio centro Substrati, Istituto Zooprofilattico Sperimentale della Lombardia e dell'Emilia (IZS), Brescia, Italy).

behavior between the two cobalt forms, suggesting that Co NP are potential genotoxic carcinogens (Table 3).

In *in vitro* nanotoxicology research the comparative study of metallic NP and of the corresponding potential released ions is also essential to avoid incorrect interpretation of the toxicological response. Confocal microscopy experiments on the effect of cobalt NP on the cytoskeleton and the chromatin of Balb/3T3 cells stained with annexin V and propidium iodide illustrates this aspect. Exposure to increasing concentrations of cobalt NP leads to a progressive cytoskeletal damage and degradation of chromatin (Figure 4(a)). However, a similar response is also obtained when cells are exposed to cobalt ions (Figure 4(b)). Again, we cannot ascribe the observed effects to cobalt NP per se. We are not aware if such effects are due to cobalt NP or cobalt ions or both.

Moreover, in order to ascribe any observed biological effect induced by exposure to the investigated NP parallel studies with the corresponding microparticle is also necessary. A comparative *in vitro* study on cobalt nano- and microparticles as well as $CoCl_2$ showed that cobalt NP modulate cytokine release through human mononuclear cells mimicking autoimmune disease.[130]

Behavior of nanoparticle and released ions in culture medium. When NP are introduced into culture medium at their surface, as at any solid surface, adsorption of molecules from serum components may occur.[131] The chemical composition of the

TABLE 3 | Alternative Testing in Nanotoxicology Research: The Example of Co NP Investigated by a Parallel Study on Cobalt Ions Potentially Released from the Particles into Culture Medium, and by a Battery of Testing Methods (Cell Transformation, Comet and Micronucleus Assays)

Step	Co NP	Cobalt Ions
Behavior in culture medium	Release of cobalt ions into culture medium (44% for 100 μM at 72 h)	—
Uptake	High	Low
Cytotoxicity	Cytotoxic effect (dose and time dependent). Synergistic effect of Co^{2+} ions?	Cytotoxic effect (dose and time dependent)
Carcinogenic potential	Positive (induction of foci type III)	Negative
Genotoxicity (Comet assay)	Induction of single strand and double strand breaks	Induction of single strand and double strand breaks
Genotoxicity (MN assay)	Induction of chromosomal aberration (not dose-dependent)	No chromosomal aberration observed

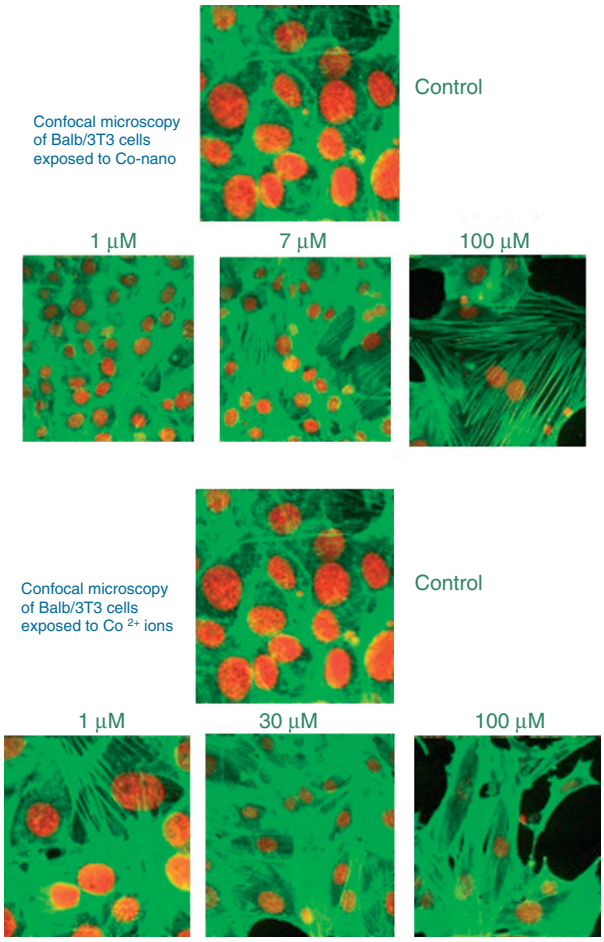


FIGURE 4 | Morphology of Balb/3T3 cells unexposed (control) or exposed to increasing concentrations of cobalt NP (Alexa-488-phalloidine and propidium iodide staining).[129]

surface and its topography determine the selectivity and strength of this adsorption. Proteins, as major constituent of the serum, are large amphipathic molecules which tend to adsorb onto surfaces, through noncovalent (e.g., hydrogen bonding, electrostatic and hydrophobic) interactions. While adsorbed, the protein can undergo various transformations, some of which are irreversible and result in denaturation. The retention or modification of the native structure of a protein upon adsorption is of paramount importance, because it determines its activity at the solid/living matter interface. Moreover, the physicochemical nature of the NP surface is expected to determine the kind of protein, which adsorbs preferentially, as well as the adsorption strength. As a general rule the more hydrophobic the surface, the greater is the extent of adsorption. Moreover, the metal ions potentially released into the medium can interact with culture medium components leading to metallobiocomplexes. Following exposure to NP the culture medium may contain NP and metal ions bound to medium components and, after saturation of such binding, 'free' metal ion. The determination of the contribution of each of such metal species to the observed toxic effect is of paramount importance in understanding the mode of action of NP toxicity.

Uptake and subcellular distribution. The determination of the kinetics of uptake and intracellular distribution recycling and excretion of NP versus the corresponding microparticles as well as the released metal ions at the cellular level is a key point for the extrapolation of *in vitro* data to *in vivo* situations and for a better understanding of the mechanisms underlying the interaction of NP with cells[90] and the possible corresponding toxicological effects. In order to carry out these studies the NP employed must be labeled, for example, by fluorescent or radioactive approaches. For example, the radiolabeling of cobalt NP with ^{60}Co radioisotope by neutron activation (^{60}Co-NP, Figure 5) and of Co^{2+} ions with ^{57}Co radiotracer allowed the sensitive and accurate determination of the uptake and

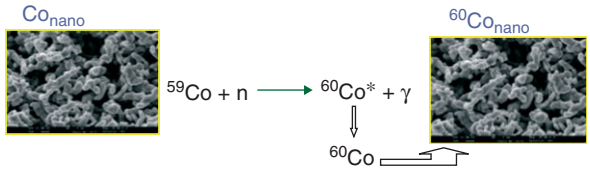


FIGURE 5 | The radiolabelling of solid cobalt NP by neutron activation in the nuclear reactor. Exposure of NP to a neutron flux of 10^{13} neutrons cm^{-2} s^{-1} (Triga Mark II reactor, University of Pavia, Italy) or of 2×10^{14} neutrons cm^{-2} s^{-1} (HFR reactor, petten, the Netherlands) led to Co [^{60}Co]-radiolabeled NP with different specific radioactivity to be used for uptake and intracellular distribution studies in Balb/3T3 cells.[135]

intracellular distribution of ^{60}Co (from cobalt NP) and of ^{57}Co (from cobalt ions) accumulated by mouse fibroblasts Balb/3T3 cell and its cellular subfractions. This was made possible because the two radioisotopes decay by emission of obviously different γ rays, which can be determined by high resolution computer-based γ ray spectrometry.[132] Results showed a 50 to 100-fold increase of uptake for ^{60}Co in comparison to $^{57}Co^{2+}$. This high quantity of Co NP uptake, in comparison with ions, is possible as a result of the chemical nature of Co NP, since the particles may interact with proteins present in culture medium on their surface, which can lead to easier uptake by the cells via a Trojan-horse type mechanism, that is, hiding under a cloak to gain entrance.[133] The same approach was used in a study on genotoxic effects induced by cobalt NP and cobalt ions in human peripheral blood leukocytes (PBLs). The uptake of cobalt from ^{60}Co NP was very efficient while the $^{57}Co^{2+}$ showed slight or no variations over the baseline levels.[134]

Single test or test battery? This *in vitro* study concerning the concurrent cytotoxicity and transforming capacity of Co NP in a Balb/3T3 cell line is an example of the application of a battery of toxicity testing that includes cell transformation assay (CTA) in combination with Comet and Micronucleus (MNT) assays to detect genotoxic carcinogenic and non genotoxic carcinogenic potential of NP.[127] Here, cobalt NP showed cytotoxicity and a dose-dependent statistically significant increase in the morphological neoplastic transformation in two different Balb/3T3 clones (Figure 3). Cobalt chloride induced cytotoxicity but did not show a carcinogenic potential neither with respect to dose dependency, nor for the single concentration values versus control. Moreover, genotoxicity assays (Comet assays, Figure 6(a) and (b), and micronucleus test) show different behaviors between the two cobalt forms, suggesting that Co NP are potential genotoxic carcinogens (Table 3) in the Balb/3T3 cell line.

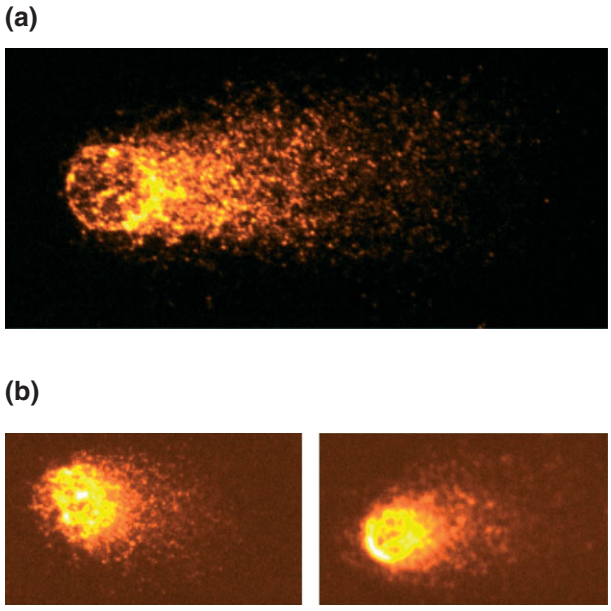
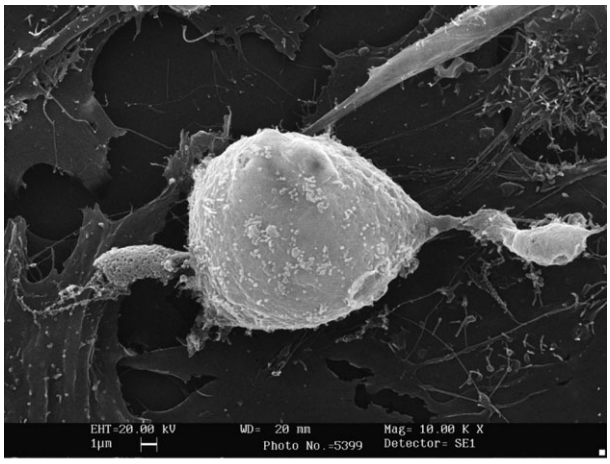


FIGURE 6 | Comet assay of Balb/3T3 cells exposed for 2 h to 1 μM of cobalt NP or equimolar concentrations of Co^{2+} ions. (a) Positive control (300 μM of H_2O_2, relevant DNA damage). (b) ModerateDNA damage induced by the two cobalt compounds (left: Co NP; right: Co^{2+} ions).[115]

Endocytic pathways and biopersistence. Scanning Electron Microscope (SEM) is important to study the morphology of unexposed and NP exposed cells. A microphotograph obtained of Balb/3T3 cells shows its typical morphology and a degrading light from the surface to inside the cell suggesting that particles are penetrating the Balb/3T3 cell (Figure 7). The penetration may occur by endocytosis as suggested by Figure 8 that gives evidence of an interaction of the particles with cells. At present we have no information about the fate of NP entering the cell. This is a difficult task and it represents a challenge for future nanotoxicology research.[136,137]

CONCLUSIONS

To quote George Box once again, 'All models are wrong, some are useful'. We are not 70 kg rats, but we also cannot reconstruct a human organism from a multitude of cell systems or model it in a computer. The more we know what we are looking for, the better we can target our testing. If we have no hypothesis, screening through several models and black-box types of animal tests might be the only way forward. NP are different but they are not so different that we should expect completely new hazards, although hazards not necessarily shown by the parent compound may be present.



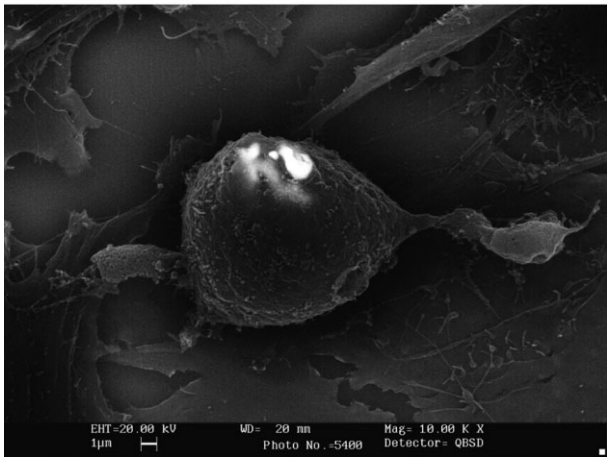


FIGURE 7 | SEM pictures of Balb/3T3 cells exposed to 100 µM Co-nano for 4 h. Preparation by Critical Point Dryer technique (with kind permission by Dr. T. Sasaki, Unit BMS, Joint Research Centre, Ispra, Italy). The observed degradation of the light suggests an effective penetration of NP into cells.[115]

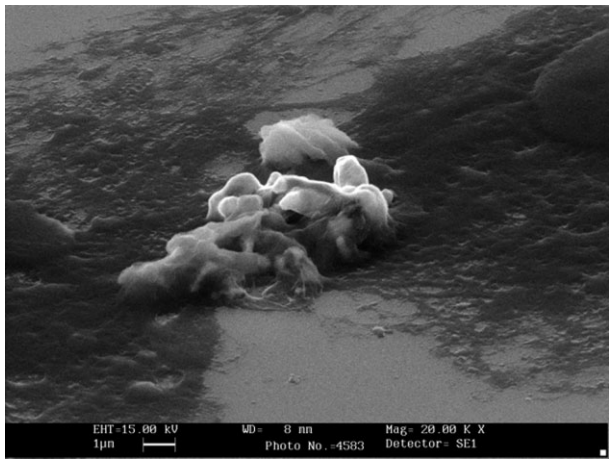


FIGURE 8 | SEM pictures of Balb/3T3 cells exposed to 100 µM Co-nano for 4 h. Preparation by Critical Point Dryer technique (with kind permission by Dr. T. Sasaki, Unit BMS, Joint Research Centre, Ispra, Italy). We observe an agglomeration of NP between two cells and the cell membrane 'engulfing' the particles.[115]

Hazard assessment, therefore, will often be possible with targeted tests (especially *in vitro* approaches), often referred to as alternative methods. They represent a reasonable compromise between effort and information gain, allowing direct comparison of various NP and their parent compounds. A broad animal-based screening approach is not feasible with regard to laboratory capacities and costs and is certainly not desirable from an animal welfare point of view.

A number of alternative approaches have undergone the optimization and validation process to make them suitable for regulatory purposes. It appears to be most promising to adapt these to NP in order to have a testing platform for broader characterization. When combined with somewhat more extensive physicochemical characterization than normally done for industrial chemicals, this will help us derive some more general rules regarding when to expect hazards from NP.

The field of alternative approaches has taken time and many resources to learn the importance of good practices and standardization for the success of validation and regulatory acceptance of methods. It is strongly advised that the respective guidance on Good Laboratory Practice for *in vitro* toxicity[138] and Good Cell Culture Practice[5] be followed from the beginning. It is promising that some good practices for how to test nanotoxicology emerged from expert workshops.[88,139–141] In the near future, the respective quality assurance for the execution of such tests will be integrated.

Because of the high number and heterogeneity of particle samples and experimental systems, it is still difficult to find common principles of NP toxicity.[142] However, it has been rightly warned[1] that we are witnessing the first generation of NP only; more sophisticated NP (active nanostructures, coated NP, integrated nanosystems, etc.) will make this even more complicated. Thus, it might well be that each and every NP formulation of a substance will have to be considered an individual entity requiring at least some risk assessment. K.C. Elliott[143] has characterized nanotoxicology as a pre-normal science, in which researchers have no widely accepted paradigm to guide their investigations.

We must not forget that not only the NP itself but also contaminations may have adverse effects; as an interesting example, carbon nanotubes were shown to include metals, amorphous carbon, and other compounds.[1] A special case of high relevance

is the contamination with pyrogens because of the large surface area and the high lipophilicity of these compounds.[144] It remains to be seen whether current pyrogenicity tests can retrieve such contamination before applying nanomedicines by injection.

The major problem for NP risk assessment is kinetics. Here, we have to expect differences from the parent compound due to size and shape, but we do not really know how to test for them. Species differences are not really well established. A key problem is that we still do not know how NP are metabolically processed.[52] The field of alternatives mainly has to offer some barrier models, which certainly represent a key priority.

Ultimately, the direct comparison of *in vitro* and *in vivo* findings will be necessary. This is best done in the context of a formal validation study, though some orientation in the optimization phase of assay development will be necessary. Today, however, mostly assays already validated for other materials, especially industrial chemicals, are sufficiently standardized to merit such comparisons. It is timely to evaluate their suitability for NP, that is, the extension of their applicability domain[98] to NP. Assays under development might be dropped if such *in vitro* to *in vivo* comparisons are done prematurely. It often takes many years until an assay is fit for formal validation.

Last but not least, toxicity is not always bad news, as sometimes this can actually be exploited for therapeutic purposes.[2] The main difference between toxicology and pharmacology is whether an effect is desired. NP offer fascinating opportunities to interfere with the organism in new ways. We must take care to find the right balance between opportunities and safety concerns. *In vitro* approaches promise to provide an affordable database on the biological activities to help understand the risks and opportunities.

REFERENCES

1. Fadeel B, Kagan V, Krug H, Shvedova A, Svartengren M, Tran L, Wiklund L. There's plenty of room at the forum: potential risks and safety assessments of engineered nanomaterials. *Nanotoxicol* 2007, 1: 73–84.
2. Oberdörster G, Stone V, Donaldson K. Toxicology of nanoparticles: a historical perspective. *Nanotoxicol* 2007, 1:2–25.
3. Sauer UG. Political incentives towards replacing animal testing in nanotechnology? *ALTEX* 2009, 26:285–294.
4. Hartung T, Leist M. Food for thought... on the evolution of toxicology and phasing out of animal testing. *ALTEX* 2008, 25:91–96.
5. Coecke S, Balls M, Bowe G, Davis J, Gstraunthaler G, Hartung T, Hay R, Merten O-W, Price A, Schechtman L, et al. Guidance on good cell culture practice. *ATLA—Altern Lab Anim* 2005, 33:261–287.
6. Hartung T. Food for thought... on evidence-based toxicology. *ALTEX* 2009, 26:75–82.
7. Balls M, Amcoff P, Bremer S, Casati S, Coecke S, Clothier R, Combes R, Corvi R, Curren R, Eskes C, et al. The principles of weight of evidence validation of test methods and testing strategies. The report and recommendations of ECVAM workshop 58. *ATLA—Altern Lab Anim* 2006, 34:603–620.
8. Hartung T. Food for thought... on validation. *ALTEX* 2007, 24:67–72.
9. Schiffelers M-JWA, Blaauboer BJ, Martje Fentener van Vlissingen J, Kuil J, Remie R, Thuring JWGM, Vaal MA, Hendriksen CFM. Factors stimulating or obstructing the implementation of the 3Rs in the regulatory process. *ALTEX* 2007, 24:273–278.
10. Bottini AA, Alepee N, De Silva O, Hartung T, Hendriksen C, Kuil J, Pazos P, Philips B, Rhein C, Schiffelers M-J, et al. of the post-validation process. The report and recommendations of ECVAM workshop 67. *ATLA—Altern Lab Anim* 2008, 36:353–366.
11. Russell WMS, Burch RL. The Principles of Humane Experimental Technique. London: Methuen; 1959. Available at: http://altweb.jhsph.edu/pubs/books/humane_exp/het-toc. (Accessed July 11, 2011).
12. Hudson M. Why do the numbers of laboratory animal procedures conducted continue to rise? An analysis of the Home Office Statistics of Scientific Procedures on Living Animals: Great Britain 2005. *ATLA—Altern Lab Anim* 2007, 35:177–187.
13. Bottini AA, Amcoff P, Hartung T. Food for thought... on globalization of alternative methods. *ALTEX* 2007, 24:255–261.
14. Hartung T. Lessons learned from alternative methods and their validation for a new toxicology in the 21st century. *J Toxicol Environ Health* 2010, 13:277–290.
15. Leist M, Kadereit S, Schildknecht S. Food for thought... on the real success of 3R approaches. *ALTEX* 2008, 25:17–24.
16. Devolder T, Reid K, Rogiers V, Webb S, Wilkins D. A review of national public funding programmes in European countries. *ALTEX* 2008, 25:233–242.
17. European Commission. Fifth report from the Commission to the Council and the European Parliament on the Statistics on the number of animals used for experimental and other scientific purposes in the member

states of the European Union COM/2007/675 final, 2007. Available at: http://ec.europa.eu/environment/chemicals/lab_animals/reports_en.htm. (Accessed July 11, 2011).

18. Hartung T, Bremer S, Casati S, Coecke S, Corvi R, Fortaner S, Gribaldo L, Halder M, Roi AJ, Prieto P, et al. ECVAM's response to the changing political environment for alternatives: consequences of the European Union chemicals and cosmetics policy. *ATLA—Altern Lab Anim* 2003, 31:473–481.
19. Zuang V, Hartung T. Making validated alternatives available—the strategies and work of the European Centre for the Validation of Alternative Methods (ECVAM). *AATEX* 2005, 11:15–26.
20. OECD, Guidance Document on the Validation and International Acceptance of New or Updated Test Methods for Hazard Assessment. Environmental Health and Safety Monograph Series on Testing and Assessment, No. 34. 2005. Available at: http://www.oecd.org/document/30/0,3343,en_2649_34377_1916638_1_1_1_1,00.html. (Accessed July 11, 2011).
21. Hoffmann S, Edler L, Gardner I, Gribaldo L, Hartung T, Klein C, Liebsch M, Sauerland S, Schechtman L, Stammati A, et al. Points of reference in validation—the report and recommendations of ECVAM workshop. *ATLA—Altern Lab Anim* 2008, 36:343–352.
22. Daneshian M, Leist M, Hartung T. Center for alternatives to animal testing—Europe (CAAT-EU): a transatlantic bridge for the paradigm shift in toxicology. *ALTEX* 2010, 27:63–69.
23. Hartung T. Food for thought... on alternative methods for cosmetics safety testing. *ALTEX* 2008, 25:147–162.
24. Bottini AA, Hartung T. Food for thought... on economics of animal testing. *ALTEX* 2009, 26:3–16.
25. European Commission. Regulation (EC) No 1907/2006 of the European Parliament and of the Council of 18 December 2006 concerning the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH), establishing a European Chemicals Agency, amending Directive 1999/45/EC and repealing Council Regulation (EEC) No 793/93 and Commission Regulation (EC) No 1488/94 as well as Council Directive 76/769/EEC and Commission Directives 91/155/EEC, 93/67/EEC, 93/105/EC and 2000/21/EC. *Off J Eur Union* 2006, 396:1–849.
26. Hartung T. Food for thought... on alternative methods for chemical safety testing. *ALTEX* 2010, 27:3–14.
27. Rovida C, Hartung T. Re-evaluation of animal numbers and costs for in vivo tests to accomplish REACH legislation requirements for chemicals. *ALTEX* 2009, 26:187–208.
28. Hartung T, Rovida C. Chemical regulators have overreached. *Nature* 2009, 460:1080–1081.
29. Gilbert N. Streamlined chemical tests rebuffed. *Nature* 2010, 463:142–143.
30. Dusinska M, NanoTEST consortium. Testing strategies for the safety of nanoparticles used in medical applications. *Nanomed* 2009, 4:605–607.
31. Greim H, Arand M, Autrup H, Bolt HM, Bridges J, Dybing E, Glomot R, Foa V, Schulte-Hermann R. Toxicological comments to the discussion about REACH. *Arch Toxicol* 2006, 80:121–124.
32. EPA. 2009. Available at: http://www.epa.gov/osa/spc/toxicitytesting/. (Accessed July 11, 2011).
33. National Research Council. *Committee on Toxicity Testing and Assessment of Environmental Agents, National Research Council: Toxicity Testing in the 21st Century: A Vision and a Strategy*. Washington, DC: The National Academies Press; 2007, 196.
34. Blaauboer BJ, Andersen ME. The need for a new toxicity testing and risk analysis paradigm to implement REACH or any other large scale testing initiative. *Arch Toxicol* 2007, 81:385–387.
35. Gibb S. A review of the National Research Council report Toxicity testing in the 21st century: a vision and a strategy. *Reprod Toxicol* 2008, 25:136–138.
36. Leist M, Hartung T, Nicotera P. The dawning of a new age of toxicology. *ALTEX* 2008, 25:103–114.
37. Hartung T. Toxicology for the twenty-first century. *Nature* 2009, 460:208–212.
38. Kavlock RJ, Austin CP, Tice RR. Commentary—toxicity testing in the 21st century: implications for human health risk assessment. *Risk Anal* 2009, 29:485–487.
39. Andersen ME, Krewski D. Toxicity testing in the 21st century: bringing the vision to life. *Toxicol Sci* 2009, 107:324–330.
40. Collins FS, Gray GM, Bucher JR. Toxicology: transforming environmental health protection. *Science* 2008, 319:906–907.
41. Seidle T, Stephens ML. Bringing toxicology into the 21st century: a global call to action. *Toxicol in Vitro* 2009, 23:1576–1579.
42. Kavlock RJ. The future of toxicity testing—the NRC vision and the EPA's ToxCast program national center for computational toxicology. *Neurotoxicol Teratol* 2009, 31:237.
43. Hartung T. A toxicology for the 21st century: mapping the road ahead. *Toxicol Sci* 2009, 109:18–23.
44. Hartung T. Towards a new toxicology—evolution or revolution? *ATLA—Altern Lab Anim* 2008, 36:635–639.
45. Hoffmann S, Hartung T. Towards an evidence-based toxicology. *Human Exp Toxicol* 2006, 25:497–513.
46. Griesinger C, Hoffmann S, Kinsner-Ovaskainen A, Coecke S, Hartung T. Proceedings of the first international forum towards evidence-based toxicology. Conference Centre Spazio Villa Erba, Como, Italy, 15-18 October 2007. *Human Exp Toxicol* 2009,

28:81–163, Special Issue: Evidence-Based Toxicology (EBT).

47. Hartung T. Food for thought... on cell culture. *ALTEX* 2007, 24:143–147.
48. Hartung T. Food for thought... on animal tests. *ALTEX* 2008, 25:3–9.
49. Hartung T, Daston G. Are *in vitro* tests suitable for regulatory use? *Toxicol Sci* 2009, 111:233–237.
50. Hartung T, Hoffmann S. Food for thought on... in silico methods in toxicology. *ALTEX* 2009, 26: 155–166.
51. Hoffmann S, Hartung T. Diagnosis: Toxic!—trying to apply approaches of clinical diagnostics and prevalence in toxicology considerations. *Toxicol Sci* 2005, 85:422–428.
52. Fisher HC, Chan WCW. Nanotoxicity: the growing need for in vivo study. *Curr Opin Biotech* 2007, 18:565–571.
53. Garnett MC, Kallinterni P. Nanomedicine and nanotoxicology: some physiological principles. *Occupat Med* 2006, 56:307–311.
54. Di Gioacchino M, Verna N, Gornati R, Sabbioni E, Bernadini G. Metal nanoparticle health risk assessment. In: Sahu SC, Casciano DA, eds. *Nanotoxicity—From In Vitro Models to Health Risks*. Chichester: John Wiley & Sons; 2009, 519–541.
55. Donaldson K, Stone V, Tran CL, Kreyling W, Borm PJA. Nanotoxicology. *Occup Environ Med* 2004, 61:727–728.
56. Kagan VE, Bayir H, Shvedova AA. Nanomedicine and nanotoxicology: two sides of the same coin. *Nanomed Nanotech Biol Med* 2005, 1:313–316.
57. Seaton A, Donaldson K. Nanoscience, nanotoxicology, and the need to think small. *Lancet* 2005, 365:923–924.
58. Kipen HM, Laskin DL. Smaller is not always better: nanotechnology yields nanotoxicology. *Am J Physiol Lung Cell Mol Physiol* 2005, 289:L696–L697.
59. Holl MMB. Nanotoxicology: a personal perspective. *WIREs Nanomed Nanobiotechnol* 2009, 1:353–359.
60. Gornati R, Papis E, Di Gioacchino M, Sabbioni E, Dalle Donne I, Milzani A, Bernadini G. In vivo and *in vitro* models for nanotoxciology testing. In: Sahu SC, Casciano DA, eds. *Nanotoxicity—From In Vitro Models to Health Risks*. Chichester, UK: John Wiley & Sons; 2009, 279–302.
61. Pal S, Tak YK, Song JM. Does the antibacterial activity of silver nanoparticles depend on the shape of the nanoparticle? A study of the gram-negative bacterium *Escherichia coli*. *Appl Environ Microbiol* 2007, 73:1712–1720.
62. Chen M, Goodman DW. Catalytically active gold: From nanoparticles to ultrathin films. *Acc Chem Res* 2006, 39:739–746.
63. Cha EK, Myung H. Cytotoxic effects of nanoparticles assessed *in vitro* and in vivo. *J Microbiol Biotechnol* 2007, 17:1573–1578.
64. Walker NJ, Bucher JR. A 21st century paradigm for evaluating the health hazards of nanoscale materials? *Toxicol Sci* 2009, 110:251–254.
65. Hillyer JF, Albrecht RM. Gastrointestinal persorption and tissue distribution of differently sized colloidal gold nanoparticles. *J Pharma Sci* 2001, 90: 1927–1936.
66. Hillegass JM, Shukla A, Lathrop SA, MacPherson MB, Fukagawa NF, Mossman BT. Assessing nanotoxicity in cells *in vitro*. *WIREs Nanomed Nanobiotechnol* 2010, 2:219–231. doi:10.1002/wnan.054.
67. Choi J-Y, Ramachandrran G, Kandlikar M. The impact of toxicity testing costs on nanomaterial regulation. *Environ Sci Technol* 2009, 43:3030–3034.
68. Ponten J. Cell biology of precancer. *Eur J Cancer* 2001, 37(Suppl. 8):97–113.
69. Frank SA, Nowak MA. Problems of somatic mutation and cancer. *Bioessays* 2004, 26:291–299.
70. Stacey GN, Hartung T. Availability, standardization and safety of human cells and tissues for drug screening and testing. In: Marx U, Sandig V, eds. *Drug Testing In Vitro: Breakthroughs and Trends in Cell Culture Technology*. Weinheim: Wiley-VCH Verlag; 2007, 231–250.
71. Markovic O, Markovic N. Cell cross-contamination in cell cultures: the silent and neglected danger. *In Vitro Cell Dev Biol Anim* 1998, 34:1–8.
72. MacLeod RA, Dirks WG, Matsuo Y, Kaufmann M, Milch H, Drexler HG. Widespread intraspecies cross-contamination of human tumor cell lines arising at source. *Int J Cancer* 1999, 83:555–563.
73. Buehring GC, Eby EA, Eby MJ. Cell line cross-contamination: how aware are Mammalian cell culturists of the problem and how to monitor it? *In Vitro Cell Dev Biol Anim* 2004, 40:211–215.
74. Nardone RM. Eradication of cross-contaminated cell lines: a call for action. *Cell Biol Toxicol* 2007, 23:367–372.
75. Kuhlmann I. Zur Problematik des prophylaktischen Antibiotikaeinsatzes in Zellkultur (Problems of the prophylactic use of antibiotics in cell culture). *ALTEX* 1993, 2:27–49.
76. Kuhlmann I. The prophylactic use of antibiotics in cell culture. *Cytotechnology* 1995, 19:95–105.
77. Coecke S, Ahr H, Blaauboer BJ, Bremer S, Casati S, Castell J, Combes R, Corvi R, Crespi CL, Cunningham ML, et al. Metabolism: a bottleneck in *in vitro* toxicological test development. *ATLA—Altern Lab Anim* 2006, 34:49–84.
78. Hartung T, Balls M, Bardouille C, Blanck O, Coecke S, Gstraunthaler G, Lewis D. Report of

ECVAM task force on good cell culture practice (GCCP). *ATLA—Altern Lab Animal* 2002, 30:407–414.

79. Monteiro-Riviere NA, Inman AO. Challenges for assessing carbon nanomaterial toxicity to the skin. *Carbon* 2006, 44:1070–1078.
80. Monteiro-Riviere NA, Inman AO, Zhang LW. Limitations and relative utility of screening assays to assess engineered nanoparticle toxicity in a human cell line. *Toxicol Appl Pharmacol* 2009, 234:222–235.
81. Kroll A, Pillukat MH, Hahn D, Schnekenburger J. Current *in vitro* methods in nanoparticle risk assessment: limitations and challenges. *Eur J Pharmaceut Biopharmaceut* 2009, 72:370–377.
82. Stone V, Johnston H, Schins RP. Development of *in vitro* systems for nanotoxicology: methodological considerations. *Crit Rev Toxicol* 2009, 39:613–626.
83. Adams LK, Lyon DY, Alvarez PJJ. Comparative ecotoxicity of nanoscale TiO_2, SiO_2 and ZnO water suspensions. *Water Res* 2006, 40:3527–3532.
84. Sager TM, Porter DW, Robinson VA, Lindsley WG, Schwegler-Berry DE, Castranova V. Improved method to disperse nanoparticles for *in vitro* and in vivo investigation of toxicity. *Nanotoxicol* 2007, 1:118–129.
85. Jones CF, Grainger DW. *In vitro* assessments of nanomaterial toxicity. *Adv Drug Deliv Rev* 2009, 61:438–456.
86. Woerle-Knirsch JM, Pulskamp K, Krug HF. Oops they did it again! Carbon nanotubes hoax scientists in viability assays. *Nano Lett* 2006, 6:1261–1268.
87. Powers KW, Palazuelos M, Moudgil BM, Roberts SM. Characterization of the size, shape, and state of dispersion of nanoparticles for toxicological studies. *Nanotoxicol* 2007, 1:42–51.
88. Unfried K, Albrecht C, Klotz L-O, von Mikecz A, Grether-Beck S, Schins RFF. Cellular responses to nanoparticles: target structure and mechanisms. *Nanotoxicol* 2007, 1:52–71.
89. Bouvier d'Yvoire M, Prieto P, Blaauboer BJ, Bois FY, Boobis A, Brochot C, Coecke S, Freidig A, Gundert-Remy U, Hartung T, et al. Physiologically-based Kinetic Modelling (PBK Modelling): meeting the 3Rs Agenda. The report and recommendations of ECVAM workshop 63. *ATLA—Altern Lab Anim* 2007, 35:661–671.
90. Mailander V, Landfester K. Interaction of nanoparticles with cells. *Biomacromolecules* 2009, 10:2379–2400.
91. Caldorera-Moore M, Guimard N, Shi L, Roy K. Designer nanoparticles: incorporating size, shape and triggered release into nanoscale drug carriers. *Expert Opin Drug Deliv* 2010, 7:479–495.
92. Scott L, Eskes C, Hoffman S, Adriaens E, Alepee N, Bufo M, Clothier R, Facchini D, Faller C, Guest R, et al. A proposed eye irritation testing strategy to reduce and replace in vivo studies using bottom-up and top-down approaches. *Toxicol In Vitro* 2010, 24:1–9.
93. Bulgheroni A, Kinsner-Ovaskainen A, Hoffmann S, Hartung T, Prieto P. Estimation of acute oral toxicity using no adverse effect level (NOAEL) from the 28-day repeated dose toxicity studies in rats. *Reg Tox Pharmacol* 2009, 53:16–19.
94. Kinsner-Ovaskainen A, Bulgheroni A, Hartung T, Prieto P. ECVAM's ongoing activities in the area of acute oral toxicity. *Toxicol In vitro* 2009, 23:1535–1540.
95. Gennari A, Berghe C, Casati S, Castell J, Clemedson C, Coecke S, Colombo A, Curren R, Negro GD, Goldberg A, et al. Strategies to replace in vivo acute systemic toxicity testing. *ATLA—Altern Lab Anim* 2004, 32:437–459.
96. Kirkland D, Aardema M, Henderson L, Muller L. Evaluation of the ability of a battery of three *in vitro* genotoxicity tests to discriminate rodent carcinogens and noncarcinogens I. Sensitivity, specificity and relative predictivity. *Mutat Res* 2005, 584:1–256.
97. Kirkland D, Pfuhler S, Tweats D, Aardema M, Corvi R, Darroudi F, Elhajouji A, Glatt H, Hastwell P, Hayashi M, et al. How to reduce false positive results when undertaking *in vitro* genotoxicity testing and thus avoid unnecessary follow-up animal tests: report of an ECVAM workshop. *Mutat Res* 2007, 628:31–55.
98. Hartung T, Bremer S, Casati S, Coecke S, Corvi R, Fortaner S, Gribaldo L, Halder M, Hoffmann S, Roi AJ, et al. A modular approach to the ECVAM principles on test validity. *ATLA—Altern Lab Anim* 2004, 32:467–472.
99. Basketter D, Cockshott A, Corsini E, Gerberick GF, Idehara K, Kimber I, van Loveren H, Matheson J, Mehling A, Omori T, et al. An evaluation of performance standards and non-radioactive endpoints for the local lymph node assay. The report and recommendations of ECVAM workshop 65. *ATLA—Altern Lab Anim* 2008, 36:243–257.
100. Casati S, Aeby P, Basketter DA, Cavani A, Gennari A, Gerberick GF, Griem P, Hartung T, Kimber I, Lepoittevin J-P, et al. Dendritic cells as a tool for the predictive identification of skin sensitisation hazard. *ATLA—Altern Lab Anim* 2005, 33:47–62.
101. Kimber I, Agius R, Basketter D, Corsini E, Cullinan P, Dearman R, Gimenez-Arnau E, Greenwell L, Hartung T, Kuper F, et al. Chemical respiratory allergy: opportunities for hazard identification and characterisation. The report and recommendations of ECVAM workshop 60. *ATLA—Altern Lab Anim* 2007, 35:243–265.
102. Basketter D, Pease C, Kasting G, Kimber I, Casati S, Cronin M, Diembeck W, Gerberick F, Hadgraft J, Hartung T, et al. Skin sensitisation and epidermal disposition: the relevance of epidermal disposition

for sensitisation hazard identification and risk assessment. The report and recommendations of ECVAM workshop 59. *ATLA—Altern Lab Anim* 2007, 35:137–154.

103. Gerberick F, Aleksic M, Basketter D, Casati S, Karlberg A-T, Kern P, Kimber I, Lepoittevin JP, Natsch A, Ovigne JM, et al. Chemical reactivity measurement and the predictive identification of skin sensitisers—the report and recommendations of ECVAM workshop 64. *ATLA—Altern Lab Anim* 2008, 36:215–242.
104. Hareng L, Pellizzer C, Bremer S, Schwarz M, Hartung T. The integrated project ReProTect: a novel approach in reproductive toxicity hazard assessment. *Reprod Toxicol* 2005, 20:441–452.
105. Prieto P, Baird AW, Blaauboer BJ, Vicente J, Ripoll C, Corvi R, Dekant W, Dietl P, Gennari A, Gribaldo L, et al. The assessment of repeated dose toxicity *in vitro*: a proposed approach. *ATLA—Altern Lab Anim* 2006, 34:315–341.
106. Jeram S, Riego Sintes JM, Halder M, Baraibar Fentanes J, Sokull-Klüttgen B, Hutchinson TH. A strategy to reduce the use of fish in acute ecotoxicity testing of new chemical substances notified in the European Union. *Reg Tox Pharmacol* 2005, 42:218–224.
107. Gennari A, Ban M, Braun A, Casati S, Corsini E, Dastych J, Descotes J, Hartung T, Hooghe-Peters House R, et al. The use of *in vitro* systems for evaluating immunotoxicity: the report and recommendations of an ECVAM workshop. *J Immunotox* 2005, 2:61–83.
108. Langezaal I, Hoffmann S, Hartung T, Coecke S. Evaluation and prevalidation of an immunotoxicity test based on human whole blood cytokine release. *ATLA—Altern Lab Anim* 2002, 30:581–595.
109. Carfi M, Gennari A, Malerba I, Corsini E, Pallardy M, Pieters M, Van Loveren H, Vohr HW, Hartung T, Gribaldo L. *In vitro* tests to evaluate immunotoxicity: a preliminary study. *Toxicol* 2007, 229:11–22.
110. Hoffmann S, Peterbauer A, Schindler S, Fennrich S, Poole S, Mistry Y, Montag-Lessing T, Spreitzer I, Loschner B, vam Aalderen M, et al. International validation of novel pyrogen tests based on the human fever reaction. *J Immunol Meth* 2005, 298:161–173.
111. Schindler S, Spreitzer I, Loschner B, Hoffmann S, Hennes K, Halder M, Brügger P, Frey E, Hartung T, Montag T. International validation of pyrogen tests based on cryopreserved human primary blood cells. *J Immunol Meth* 2006, 316:42–51.
112. Daneshian M, von Aulock S, Hartung T. Assessment of pyrogenic contaminations with the validated human whole blood assay. *Nature Protocols* 2009, 12:1709–1721.
113. Schindler S, von Aulock S, Daneshian M, Hartung T. Development, validation and applications of the monocyte activation test for pyrogens based on human whole blood. *ALTEX* 2009, 26:293–305.
114. OECD, Environment, Health and Safety Publications Series on the Safety of Manufactured Nanomaterials No. 6. Working party on manufactured nanomaterials: list of manufactured nanomaterials and list of endpoints for phase one of the OECD testing programme, 2008. Available at: http://appli1.oecd.org/olis/2008doc.nsf/linkto/env-jm-mono(2008)13-rev.
115. Sabbioni E. *In vitro* nanotoxicology research. A great potential for nuclear and radiochemical techniques, 8th International Conference on Nuclear Activation Techniques in the Life Sciences (NAMLS-8), Rio de Janeiro, Brazil, 17–22 April 2005. Available at: http://www.cena.usp.br/namls8/.
116. des Rieux A, Fievez V, Garinot M, Schneider YJ, Preat V. Nanoparticles as potential oral delivery systems of proteins and vaccines: a mechanistic approach. *J Control Release* 2006, 116:1–27.
117. Emerich DF, Thanos CG. Targeted nanoparticle-based drug delivery and diagnosis. *J Drug Target* 2007, 15:163–183.
118. Elder A, Vidyasagar S, De Louise L. Physicochemical factors that affect metal and metal oxide nanoparticle passage across epithelial barriers. *Wiley Interdiscip Rev Nanomed Nanobiotechnol* 2009, 1:434–450.
119. Yang W, Peters JI, Williams RO. Inhaled nanoparticles—a current review. *Int J Pharm* 2008, 356:239–47.
120. Card JW, Zeldin DC, Bonner JC, Nestmann ER. Pulmonary applications and toxicity of engineered nanoparticles. *Am J Physiol Lung Cell Mol Physiol* 2008, 295:L400–L411.
121. Muhlfeld C, Gehr P, Rothen-Rutishauser B. Translocation and cellular entering mechanisms of nanoparticles in the respiratory tract. *Swiss Med Wkly* 2008, 138:387–391.
122. Dobrovolskaia MA, Aggarwal P, Hall JB, McNeil SE. Preclinical studies to understand nanoparticle interaction with the immune system and its potential effects on nanoparticle biodistribution. *Mol Pharm* 2008, 5:487–95.
123. Aggarwal P, Hall JB, McLeland CB, Dobrovolskaia MA, McNeil SE. Nanoparticle interaction with plasma proteins as it relates to particle biodistribution, biocompatibility and therapeutic efficacy. *Adv Drug Deliv Rev* 2009, 61:428–437.
124. Sabbioni E. Nanoparticles toxicology research at JRC and its approach, Informal meeting on Nanoparticles and Nanotechnologies hosted by Austrian Presidency, Vienna, Austria, 2006. Available at: http://cordis.europa.eu./austria/ebents_09062006_en.htlm.
125. Behra R, Krug H. Nanoparticles at large. *Nature Nanotech* 2008, 3:253–254.
126. Lubick N. Nanosilver toxicity: ions, nanoparticles or both? *Environ Sci Technol* 2008, 42:8617.

127. Ponti J, Sabbioni E, Munaro B, Broggi F, Marmorato P, Franchini F, Colognato R, Rossi F. Genotoxicity and morphological transformation induced by cobalt nanoparticles and cobalt chloride: an *in vitro* study in Balb/3T3 mouse fibroblasts. *Mutagenesis* 2009, 24:439–445.

128. Broggi F. 'Quadri metabolici, genotossicita' e potenziale cancerogeno *in vitro* di cobalto nano particelle, Thesis, Faculty of Matemathics, Physical and Natural Sciences, University of the Insubria (Varese), Italy, Academic year 2005–2006.

129. Sabbioni E. Nanotossicologia, ricerca e metodi alternativi: la strategia europea, Simposio Italiano di Nanotossicologia, Università dell'Insubria, 28 February 2007. Available at: www.uninsubria.it/uninsubria/allegati/pagine/4229/nanotox.pdf. (Accessed June 23, 2011).

130. Papis E, Gornati R, Ponti J, Prati M, Sabbioni E, Bernardini G. Gene expression in nanotoxicology: a search for biomarkers of exposure to cobalt particles and ions. *Nanotoxicol* 2007, 1:198–203.

131. Linse S, Cabaleiro-Lago C, Xue WF, Lynch I, Lindman S, Thulin E, Radford SE, Dawson KA. Nucleation of protein fibrillation by nanoparticles. *Proc Natl Acad Sci U S A* 200:8691–8696.

132. Sabbioni E, Ponti J, Del Torchio R, Farina M, Fortaner S, Munaro B, Sasaki T, Rossi F. Recherche *in vitro* sur la toxicologie des nanoparticules au Joint Research Center. *Méd Nucl - Imagerie fonctionnelle métabolique* 2006, 30:15–23.

133. Limbach LK, Wick P, Manser P, Grass RN, Bruinink A, Stark WJ. Exposure of engineered nanoparticles to human lung epithelial cells: influence of chemical composition and catalytic activity on oxidative stress. *Environ Sci Technol* 2007, 41:3791–3792.

134. Colagnato R, Bonelli A, Poni J, Farina M, Bergamaschi E, Sabbioni E, Migliore L. Comparative genotoxicity of cobalt nanoparticles and ions on human peripheral leukocytes *in vitro*. *Mutagenesis* 2008, 23:377–382.

135. Sabbioni E, Gatti AM, Hartung T. Pathology of new diseases induced by nanomaterials and *in vitro* toxicology research, XXV Congress of the International Academy of Pathology Brisbane, 10–15 October 2004. *Pathol Int* 2004, 50:S141–S148.

136. Chavanpatil MD, Khdair A, Panyam J. Nanoparticles for cellular drug delivery: mechanisms and factors influencing delivery. *J Nanosci Nanotechnol* 2006, 6:2651–2663.

137. Harush-Frenkel O, Altschuler Y, Benita S. Nanoparticle-cell interactions: drug delivery implications. *Crit Rev Ther Drug Carrier Syst* 2008, 25:485–544.

138. OECD, Series on principles of good laboratory practice and compliance monitoring, Number 14, Advisory Document of the Working Group on Good Laboratory Practice, The Application of the Principles of GLP to *in vitro* Studies, 2004. Available at: http://appli1.oecd.org/olis/2004doc.nsf/linkto/env-jmmono(2004)26.

139. Maynard AD, Aitken RJ, Butz T, Colvin V, Donaldson K, Oberdörster G, Philbert MA, Ryan J, Seaton A, Stone V, et al. Safe handling of nanotechnology. *Nature* 2006, 444:267–269.

140. Balbus JM, Maynard AD, Colvin VL, Castranova V, Daston GP, Denison RA, Dreher KL, Goering PL, Goldberg AM, Kulinowski KM, et al. Meeting report: hazard assessment for nanoparticles—report from an interdisciplinary workshop. *Environ Health Perspect* 2007, 115:1654–1659.

141. Warheit DB, Borm PJ, Lademann HCJ. Testing strategies to establish the safety of nanomaterials: conclusions of an ECETOC workshop. *Inhal Toxicol* 2007, 19:631–643.

142. Hoet P, Boczkowski J. What's new in nanotoxicology? Brief review of the 2007 literature. *Nanotoxicol* 2008, 2:171–182.

143. Elliott KC. Varieties of exploratory experimentation in nanotoxicology. *Hist Phil Life Sci* 2007, 29:313–336.

144. Ashwood P, Thompson RP, Powell JJ. Fine particles that adsorb lipopolysaccharide via bridging calcium cations may mimic bacterial pathogenicity towards cells. *Exp Biol Med* 2007, 232:107–117.