REVIEW

Nanoparticles, human health hazard and regulation

Anthony Seaton1,2,*, Lang Tran1, Robert Aitken1 and Kenneth Donaldson1,3

1Safety of Nanomaterials Interdisciplinary Research Centre, Institute of Occupational Medicine, Research Avenue North, Riccarton, Edinburgh EH14 4AP, UK
2Department of Environmental and Occupational Medicine, University of Aberdeen, Foresterhill, Aberdeen AB25 2ZP, UK
3ELEGI Colt Laboratory, MRC/University of Edinburgh Centre for Inflammation Research, Queen’s Medical Research Institute, 47 Little France Crescent, Edinburgh EH16 4TJ, UK

New developments in technology usually entail some hazard as well as advantage to a society. Hazard of a material translates into risk by exposure of humans and/or their environment to the agent in question, and risk is reduced by control of exposure, usually guided by regulation based on understanding of the mechanisms of harm. We illustrate risks relating to the causation of diseases associated with exposure to aerosols of combustion particles and asbestos, leading to paradigms of particle toxicity, and discuss analogies with potential exposure to manufactured nanoparticles (NPs). We review the current understanding of the hazard of NPs derived from the new science of nanotoxicology and the limited research to date into human exposure to these particles. We identify gaps in knowledge relating to the properties of NPs that might determine toxicity and in understanding the most appropriate ways both to measure this in the laboratory and to assess it in the workplace. Nevertheless, we point out that physical principles governing the behaviour of such particles allow determination of practical methods of protecting those potentially exposed. Finally, we discuss the early steps towards regulation and the difficulties facing regulators in controlling potentially harmful exposures in the absence of sufficient scientific evidence.

Keywords: nanoparticles; toxicological hazard; measurement; regulation

1. INTRODUCTION

It is probable that the introduction of all new technologies will have unexpected consequences, both beneficial and harmful. In a democratic society, the business of regulation is to reduce risks from the latter while not inhibiting the development of the former. Scientific evidence plays a major role, often misunderstood by the media and even by some scientists, in the process of regulation. Indeed ‘health and safety’ has become an easy excuse for the prevention of desirable activities by those unable or unwilling to understand the process.

Public awareness of nanotechnologies is now increasing rapidly, largely as a consequence of non-specifically regulated introduction of products containing nanomaterials into the marketplace. Self-cleaning windows, stain-resistant clothes and better tennis racquets are well known. The Internet provides opportunities for individuals to dose themselves with tonics containing nanosilver in the belief that this will prevent infections. The fear that nanotechnologies would be seriously hindered, as previously had those involving genetic modification, by the alarms sounded by pressure groups hinting at grey goo and human enhancement has now largely been set aside, and their entrepreneurial development is proceeding apace. Nevertheless, it is important to heed the advice given to government in 2005 by the working group of the Royal Society and the Royal Academy of Engineering (Royal Society and Royal Academy of Engineering 2004). This may be summarized as indicating that nanotechnologies present great opportunities but that it is also possible to foresee some hazards, which should be taken account of by research on which regulation might be based. The nature of such research must be interdisciplinary. In this paper, we discuss this research and summarize where we stand today with respect to translating this into regulation to ensure that those making, using and disposing of nanomaterials are protected as far as practicable from injury.

*Author and address for correspondence: 8 Avon Grove, Edinburgh EH4 6RF, UK (a.seaton@abdn.ac.uk).

One contribution to a Theme Supplement ‘NanoBioInterface: crossing borders’.

Published online 2 September 2009

Received 25 June 2009
Accepted 10 August 2009
2. HISTORY OF NANOHAZARDS

It is appropriate to start with a warning. Two toxic substances stand out above all others in the history of peacetime technology in terms of deaths caused: quartz and asbestos. In the former case, the harmfulness of inhaling fine particles was recognized in the eighteenth century and in the latter at the end of the nineteenth century. In both cases, physicians and scientists were intrigued by the possibility of understanding the mechanisms of toxicity, an area of investigation that attracted support from the industries concerned, driven by the medical mindset that if a mechanism can be found, then a cure may follow. In both cases, appropriate epidemiological research to associate exposure and effect was long delayed, effective regulation and enforcement were not introduced and workers continued to die (and still do). The industries concerned proved resistant to funding this research. The warning is this: it is not necessary to understand mechanisms before taking steps to prevent an occupational or environmental disease, and too great a focus on mechanisms alone, though scientifically interesting, may distract from applying preventive measures.

The concept that nanoparticles (NPs) may have unique toxic properties is a recent one, though the thinking behind it is derived from classical particle toxicology and epidemiology, especially from studies of air pollution. The harmfulness of air pollution was suspected in the reign of Charles II but its real cost in lives lost became clear in the 1930s and 1940s with two episodes, one in the Meuse valley in Belgium and the other in Donora in Pennsylvania. Coal smoke and a winter temperature inversion were a lethal combination (McDonald et al. 1951). These episodes were followed by the infamous London smog of December 1952 in which some 4000 excess deaths occurred over a week. These deaths mostly appeared to be from lung disease, chronic bronchitis being then a highly prevalent condition in urban Britain, but it was apparent that a proportion was from heart attacks. Unusually, this is an example of regulation occurring promptly (cynics would argue that this was because the episode centred on the Houses of Parliament), and the Clean Air Acts were introduced and proved very effective in reducing pollution while argument continued as to which components of pollution, gases, acid aerosols or solid particles were responsible. In addition to effective regulation, research was funded by the UK Government at the MRC air pollution unit in St Bartholomew’s Medical School, and among the unit’s many achievements was the demonstration that over half of the particles in London smog were less than 0.1 μm in diameter (Lawther et al. 1968). The significance of this observation went largely unnoticed for 25 years, but in the 1990s two separate lines of research converged: epidemiological studies demonstrated associations between deaths and particulate air pollution even at extraordinarily low mass concentrations (Pope et al. 1992; Schwartz 1994) and studies of deposition of inhaled particles in the lungs of rats led to the observation that particles in the nanosize range were retained in the lungs and translocated to the interstitial tissues more readily than larger particles (Ferin et al. 1992). We pointed out that the majority of deaths associated with air pollution in the epidemiological studies were from cardiac rather than respiratory disease and attempted to explain the apparent fact that toxicologically tiny doses of particulate matter (PM), mainly carbon, to the lungs could cause death from failure of another organ by proposing the hypothesis that the effect was a consequence not of the mass of particles but of the number. We suggested that such particles evaded the lungs’ primary defence, the macrophage, and caused inflammation in the interstitial space leading to release of cytokines and alteration of the coagulability of the blood (Seaton et al. 1995). This rather complex idea, generally called the ultrafine hypothesis, triggered a change in the way toxicologists were to investigate air pollution, with an emphasis on the role of particles in the nanosize range.

Even more importantly in terms of the number of individuals affected, longitudinal epidemiological studies have shown that the intensity of population exposure to air pollution in the past is related to the risk of later death from heart disease in that population (Dockery et al. 1993; Pope et al. 1995), suggesting that air pollution, like cigarette smoking, is a risk factor for the development of coronary artery disease and thus by implication of atheroma, the disease process in arterial walls that results in their progressive obstruction. Atheromatous plaques rupture, releasing their contents and precipitating thrombosis and the acute coronary syndrome. As there is good evidence that inflammation is involved in atherogenesis, cross-talk between pulmonary inflammation caused by particles and inflammatory activity in the plaques leading to their rupture may be invoked to explain this association.

Over the same period that scientists have investigated air pollution, there has been an analogous interest in the toxicity of asbestos. From the early twentieth century, it has been recognized that workers exposed to this mineral ran excessive risks of the development of fibrosis of the lungs and, from the 1950s and 1960s, of lung cancer and the pleural tumour mesothelioma. The epidemiological research relating to the risks of mesothelioma led to the conclusion that it would be virtually impossible to prevent this disease unless the importation and use of asbestos were prohibited. However, asbestos was a very useful material and many materials with similar properties were necessary for industry. Thus, the toxicological research, initially carried out to understand why this chemically inert mineral had such devastating effects on the health of those inhaling it (Mossman et al. 1990), became relevant to the investigation of possible hazards from substitute fibrous minerals (Miller et al. 1999). This led to the development of what became known as the fibre paradigm.

3. PARADIGMS OF TOXICITY

3.1. The ultrafine hypothesis

There is not yet convincing epidemiological evidence that the toxic component of particulate air pollution
with stable coronary artery disease were exposed to 2009). Moving closer to the at-risk population, patients in inhalation (Mills capacity, was shown to be reduced after diesel exhaust activator, a key regulator of endogenous fibrinolytic exercise and the acute release of tissue plasminogen and a threefold greater increase in ischaemia was evidenced during exposure to diesel exhaust than during exposure to filtered air (Mills et al. 2007). This vascular dysfunction seems to involve nitric oxide pathways; reduced nitric oxide bioavailability as a consequence of oxidative stress has been advanced as a potential mechanism. In vitro studies have confirmed the release of superoxide from diesel particles as potential arbiters of adverse vascular effects and suggest that exposure to PM could contribute to a hypertensive phenotype (Auchincloss et al. 2008; Miller et al. 2009). Moving closer to the at-risk population, patients with stable coronary artery disease were exposed to dilute diesel exhaust (300 μg m\(^{-3}\)) during intermittent exercise and the acute release of tissue plasminogen activator, a key regulator of endogenous fibrinolytic capacity, was shown to be reduced after diesel exhaust inhalation (Mills et al. 2007). The extent of this reduction was comparable to that seen in cigarette smokers. The clinical effect of diesel exposure was further examined in patients with coronary heart disease. When patients were exposed to diesel exhaust, myocardial ischaemia was quantified by ST-segment analysis using continuous 12-lead electrocardiography and a threefold greater increase in ischaemia was evident during exposure to diesel exhaust than during exposure to filtered air (Mills et al. 2007). Ex vivo thrombus formation has also been assessed, using a Badimon chamber, after controlled exposures to dilute diesel exhaust. The Badimon chamber measures thrombus formation triggered by exposure to porcine aorta denuded of its endothelium in extracorporeal whole blood, under flow conditions that mimic those found in diseased coronary arteries. Within 2 h of dilute diesel exhaust exposure, thrombus formation was enhanced and associated with increased platelet activation (Lucking et al. 2008). In studies where the particles were filtered out and exposure was only to the gases, there was no effect, confirming the role of the particles. Taken together, these studies forge a clear link between exposure to combustion-derived particles and atherothrombosis. It is accepted that these studies include exposure to particles of greater than nanodimensions, but, on the reasonable assumption that toxicity depends on surface area and reactivity, the greatest toxicity will be found in the huge numbers of small particles.

Associations with exposure to air pollution particles in epidemiological studies of the clotting factors fibrinogen and factor VII have been equivocal, perhaps relating to the time of sampling of what is a dynamic system involving production and consumption of these factors. There is, however, in vitro evidence of expression of the clotting factor, tissue factor, and reduction in the clot disrupting factor, plasminogen activator, in response to air pollution particles (Gilmour et al. 2005).

3.2. Transgression of biological membranes: the role of size and shape

An NP may for practical purposes be characterized as having at least one dimension less than 100 nm. Thus, nanotubes and nanofibres may be many micrometres in length and only a few nanometres in diameter. Many asbestos fibres and fibrils conform to this definition. The fibrogenicity and carcinogenicity of asbestos is now very well known and much research has led to the conclusion that this human toxicity is related primarily to four factors:

(i) a length greater than about 15 μm, below which the fibre can be removed by lung macrophages;
(ii) a diameter less than about 3 μm, allowing the fibre to be inhaled to the gas-exchanging part of the lung;
(iii) insolubility or resistance to dissolution in the lung milieu;
(iv) a sufficient dose to the target organ.

This has become known as the fibre paradigm and is very likely to apply to any inhaled fibre. It seems that the chemical composition of the fibre is relatively unimportant other than as a determinant of its solubility and that macrophages may ingest many short fibres without coming to harm. There is evidence, for example, that the paradigm applies to fibrous erionite, to which individuals in Turkey suffering from asbestos-type diseases have been exposed environmentally (Artvinli & Baris 1982). The mechanism appears to be frustrated phagocytosis, whereby the macrophage is injured in attempting to engulf long fibres and releases cytokines, mitogens and oxidants that initiate the process of fibrosis and carcinogenesis. Experimentally, similar consequences have been shown to occur with other appropriately sized fibres of different compositions (Miller et al. 1999).

It is likely that the fact that human disease has not been associated with these other fibres reflects insufficient doses having reached the target organs of sufficient individuals rather than any intrinsic lack of toxicity. In part, this may be due to the awareness by manufacturers of the determinants of toxicity and deliberate design of fibres to take account of this. In the case of carbon nanotubes (CNTs) there is a lack
of knowledge on what proportion of airborne CNTs in workplaces satisfy the length criterion of the fibre paradigm, since CNTs may be long but they frequently curl up into tangled balls due to defects in the graphene structure and interception by other CNTs. The nature of the ‘particles’ that are formed means that they may be physically large (possibly larger than 15 \mu m), with high surface area but very low densities. Therefore, they could have very low aerodynamic diameter and could reach and deposit in the deep lung. It is not at present known what toxic effect such particles would have.

It is important to emphasize the difference between hazard and risk, as illustrated above, and that prevention of disease relies on prevention or reduction of exposure and hence of dose to the target organ. Asbestos has over many decades been inhaled by workers at concentrations of several million in every cubic metre of air breathed, is relatively insoluble and thus remains in the lung long enough for many fibres to migrate through to the lung lining and initiate the process that leads to mesothelioma. Dose, insolubility and fibre dimensions explain the epidemic of this cancer. In contrast, at present, industrial exposure to most other fibres of appropriate size and shape has been much more limited in both dose and duration (e.g. Miller et al. 2007). Continued reduction in exposure is the basis of action to prevent future problems.

Much debate has surrounded the possibility that roughly spherical NPs may exert their toxic effects by passage through epithelia and tissue membranes. This has led to consideration that the ability to translocate away from the site of deposition at the portal of entry leads to new hazards from NPs. The idea that inhaled NPs reach the blood, brain and other non-pulmonary target organs is the basis of much of the concern over manufactured NPs and the fear that existing risk assessment instruments are inadequate. Most hazard identification studies that are used for risk assessment do not include rigorous or sensitive analysis of cardiovascular or neurological endpoints that might be affected by chronic exposure leading to the accumulation of NPs in these sites.

The concept of translocation takes the ultrafine hypothesis a step further, to propose that the toxic effect of the particles is a consequence of a direct action on the distal target organ rather than the one mediated by cytokine release and secondary effects on the target. Thus, it proposes that inhaled NPs pass through the lungs’ alveolar membrane and go via the blood to the heart where they exert a direct toxic effect. This seems unlikely, ignoring as it does the known mechanisms of myocardial infarction, the amplification effect of provoking an inflammatory reaction, the dilutional effect of distribution through the blood and the likely uptake of any blood-borne particles by the reticulo-endothelial system. Nevertheless, some studies have searched for and found NPs in various organs including the heart following inhalation exposure (Semmler et al. 2004). Such translocation is small as a fraction of the deposited dose and has been focused on model and surrogate particles such as elemental carbon and iridium. So far, there are no data showing the translocation of environmental NPs such as those generated by diesel combustion.

More plausible have been the observations with respect to possible neurological effects of NPs. A higher concentration of NPs is deposited on the olfactory mucosa of the nose than elsewhere in the respiratory tract, where they are directly exposed to the nerve endings concerned with smell. Oberdörster et al. (2004) have reasoned that, as polio viruses gain access to the central nervous system, other nano-sized particles could be transmitted up the nerves into the brain, with potential pathological effects. Such migration has been suggested by studies on rats, but there is as yet no convincing epidemiological evidence associating air pollution with chronic brain diseases, and the relevance of this remains sub judice.

### 3.3. Inflammation and other mechanisms

Epidemiological studies have provided ample evidence, mainly in the form of rises in C-reactive protein and white blood cell count, that exposure to air pollution is associated with a generalized inflammatory reaction in the population being studied (Mills et al. 2009). Chamber studies using concentrated ambient particles have shown these particles to be mildly inflammagenic as measured by a leucocyte response in alveolar fluid obtained by bronchoscopy (Ghio et al. 2000). Human experimental studies have also shown increases in the release of vasoconstrictor substances with exposure to diesel exhaust, in this case without overt evidence of an inflammatory reaction. It seems to us likely that two or more mechanisms of toxicity are involved in mediating the association of heart disease and particulate air pollution: (i) indirect effects of pulmonary inflammation and oxidative stress on atherothrombosis and (ii) translocation of particles to blood vessels and a direct effect on their lining endothelial cells.

As noted above, longitudinal epidemiological studies have shown an association between historic exposures to pollution and later risk of developing a heart attack, strongly suggesting that pollution exposure is a risk factor for the development of arterial atheroma. Experimental support for this has been provided by studies of animals genetically predisposed to atheroma. The well-documented cardiovascular impact of ambient particles is considered to be a result of the combustion-derived NP component. The high fat diet rabbit and ApoE mouse model have been used extensively to demonstrate that pulmonary exposure to particles causes plaque expansion and increases in indices of plaque destabilisation (Suwa et al. 2002; Lippmann et al. 2005; Sun et al. 2005). Currently, there are only two studies of CNTs. One showed evidence of oxidative stress in the aorta and progression of atherosclerosis in ApoE mice after a single pulmonary aspiration of single wall CNT (Li et al. 2007). The other showed upregulation of gene expression in aortic tissue after deposition of CNT in mouse lungs (Erdely et al. 2009). For some genes, this was of orders of magnitude greater with CNT than with carbon black, suggesting that CNT may have a special action that affects the vessel wall. It is not clear whether the CNTs themselves reach the aortic
wall/atherosclerotic plaques or whether there is an indirect inflammatory/oxidative stress mechanism, although the latter is more plausible.

4. TOXICITY OF MANUFACTURED NANOPARTICLES

The predominant process underlying the pathological effects of particles in the lungs and cardiovascular system is inflammation, involved in atherothrombosis, asthma, chronic obstructive lung disease, pulmonary fibrosis and cancer (Donaldson & Tran 2002). Therefore, the ability of particles to initiate, prolong or worsen inflammation can be seen as a key property. Many studies have examined the pro-inflammatory effects of manufactured NPs, on the basis that their ability to cause inflammation is a major predictor of potential hazard in such particles. The first important finding was that NPs have a more pronounced effect on inflammation, cell damage and cell stimulation than an equal mass of particles of the same material of greater size (Donaldson et al. 2000). This appears to hold true for materials as varied as carbon black, titanium dioxide, various metals and polystyrene (Duffin et al. 2002). Surface area is the metric driving the pro-inflammatory effects and this is evident both in vitro (Donaldson & Tran 2002) and in vivo (Duffin et al. 2002), particles of various sizes producing inflammatory effects that are directly related to the surface area dose. This surface area dose-related inflammatory response is partly related to transition metals but is also found with low-toxicity materials and the cellular mechanism is not well understood. Even an apparently low toxicity surface has the ability to generate free radicals and oxidative stress in cells that has nothing to do with transition metals, as there is no soluble toxic component that could mediate the effect (Brown et al. 2000, 2001). The low toxicity surface free radical effect is evident in their ability to generate oxidants in cell-free chemical systems (Brown et al. 2001; Wilson et al. 2002). In cells, high surface area doses appear to initiate inflammation through a number of pathways but oxidative stress-responsive gene transcription is one of the most important. Oxidative stress is increased in cells exposed to some NP types (Donaldson & Stone 2007) as well as activation of oxidative stress-responsive transcription factors such as NF-κB and AP-1 (Mroz et al. 2008). However, if an NP has a reactive surface, soluble factors that are pro-inflammatory (e.g. transition metals, organics bioactivated to free radicals by the CyP450 system) or a high aspect ratio, there could be effects additional to that of simple total surface area (Donaldson et al. 2007; Poland et al. 2008; Wan et al. 2008; Stoeger et al. 2009). A wide range of NPs has been found to have the ability to cause pro-inflammatory effects (Donaldson & Stone 2007) and mechanisms other than the well-known ones described above may also be operative for some NPs (Lu et al. 2009). Such toxicological activities need to be resolved into an overarching structure–activity relationship for NPs, requiring collaboration between toxicology and surface science.

5. EXPOSURE TO MANUFACTURED NANOPARTICLES

Hazardous materials will present risks to health only if people are exposed to them. The Royal Society/Royal Academy of Engineering report identified multiple scenarios through which humans could become exposed to engineered nanomaterials including occupational, environmental and consumer exposure (Royal Society and Royal Academy of Engineering 2004). In occupational settings, exposure to NPs is plausible at all phases of the material life cycle, although the nature, level of exposure and the number of people involved could differ greatly. During the development of a new material, it is probable that this will occur under laboratory conditions. Quantities produced and numbers involved are likely to be small, but accidental releases due to spills and accidents are a possibility. Later, in commercial production, exposures may occur during synthesis or in downstream activities such as recovery, packaging, transport and storage. The quantities of materials being handled will typically be much larger. Depending on the specific properties of the new material, it may be incorporated subsequently in a range of other products or may be used in other processes as a feed-stock material. All these scenarios have the potential to result in exposure of people. NPs may also be incorporated, for example, into a composite material which may be subsequently re-engineered or reprocessed, again with potential for exposure. Finally, exposure may occur at the end of product life when the material is disposed off. Hence, for a single material, there are multiple potential exposure scenarios depending on the details of its manufacture, use and disposal.

Already more than 800 consumer products containing NPs are listed on the Woodrow Wilson database (http://www.nanotechproject.org/inventories/consumer/). They include personal-care products such as sunscreen and skin lotions, food additives, cleaning products, sealants and paints. In many of these, deliberate exposure is intended, usually by ingestion or dermal application. In other cases, incidental exposure, including by inhalation, is clearly plausible (e.g. by spraying paint containing NPs). Many types of NPs are incorporated into these products; for example, zinc oxide and titanium dioxide are incorporated into sunscreens, often at mass concentrations greater than 10 per cent. The quantities of NPs in other products are not usually clearly identified, and identifying which products do contain NPs is challenging, as there is no requirement for specific labeling. Nanosilver is one of the most commonly used NPs in consumer products, particularly as a food supplement.

5.1. Exposure metrics

Assessment of risk of harm to individuals requires understanding of the inherent toxicity of the material to which they are exposed and the dose delivered to the target organ; dose is a function of exposure and time, and assessing exposure requires measurements to be made. In early studies of the health effects of inhaled particles, dust samples were collected by drawing air
through a filter or other medium and subsequently analysed off-line to define estimates of exposure, expressed as a concentration in air. For example, in early studies in the coal industry, the samples were analysed by counting particles collected on the filter under a light microscope. This resulted in an estimate of exposure in terms of particle number concentration, expressed as number of particles per cubic centimetre or per cubic metre of air. Epidemiological studies in that industry later showed a good correlation between pneumoconiosis and mass concentration, expressed typically as milligrams per metre cube. Assessment in terms of mass was less demanding and more accurate than manual counting under a microscope and so was the preferred choice. Subsequently, the use of workplace exposure limits (WELs), based on mass concentrations, has become the norm for measuring or regulating exposure for most hazardous chemicals or particles (Health and Safety Executive 2005).

As discussed above, there is a growing view based on the toxicological literature that possible health effects arising from exposure to NPs may be better correlated with surface area than with mass concentration. Particle number is also widely suggested as an appropriate metric, based on the associations between particulate air pollution and cardiovascular disease among older people (Seaton et al. 1995). A structured review of this issue has failed to find sufficient evidence to select one metric in preference to others; it was recommended that, where possible, mass, number and surface area should all be measured (Maynard & Aitken 2007).

5.2. Studies of exposure in the workplace

Few studies so far have attempted to assess human exposure to NPs in occupational situations. Exposure to CNTs has been assessed in five studies, which focussed on laboratory-scale activities (Maynard et al. 2004; Methner et al. 2007; Bello et al. 2008, 2009; Han et al. 2008). Two of these also looked at the potential of aerosol release from cutting composites containing CNTs (Methner et al. 2007; Bello et al. 2009). These authors used a combination of methods including real-time counting and systems to measure size, number, mass and surface area concentrations together with off-line image analysis of collected samples by electron microscopy. All found some evidence of exposure under some circumstances, though in all save one fibre release was not identified. Only one study used fibre-counting methods based on those required for the measurement of asbestos in occupational settings (Han et al. 2008), reporting concentrations of 194 and 173 fibres ml$^{-1}$ associated with blending and mixing activities, based on both personal and area samples. While these are three orders of magnitude greater than typical asbestos WELs of 0.1 fibres ml$^{-1}$, all the fibres were too short to be considered a fibre under the World Health Organization (WHO) criteria, the reported maximum length observed being 1500 nm, in contrast to the WHO minimum of 5000 nm (World Health Organization 1997). Enclosure of the blending activity reduced the fibre count by four orders of magnitude, indicating effective control of fibre release.

With respect to non-fibrous particles, one study at a pilot scale ‘nanostructured particle’ gas phase metal oxide production facility showed elevated number concentrations during production rising to an order of magnitude higher than background levels (Demou et al. 2008). Another assessed airborne NP exposures associated with manual handling of small quantities (15 and 100 g) of nano-alumina and nanosilver in fume hoods (Tsai et al. 2009). Using a conventional fume hood, particle number concentration in the operator’s breathing zone increased significantly during the 100 g handling, peak concentrations at 10 and 200 nm being reported, varying according to activity. A third study characterized airborne particle concentrations during the production of carbonaceous nanomaterials, including fullerenes and CNTs, in a commercial facility (Yeganeh et al. 2008). Mass concentrations, sub-micrometre size distributions and photoionization potential (an indicator of the particles’ carbonaceous content) were measured at three locations inside the facility. Rises in concentration were associated with physical handling of the particles.

A number of common features may be identified. For air velocities prevailing in workplaces, airborne NPs can be considered as having no inertia. They will therefore behave like a gas, will diffuse rapidly and remain airborne for a long time. Because of their high diffusion velocity, these particles will readily find leakage paths in systems in which the containment is not complete. Engineering control systems for NPs, such as enclosures, local ventilation or general ventilation, therefore need to be of similar quality and specification to those normally used for gases rather than for particulate matter. As with all such systems, effective performance is dependent on the appropriate use and maintenance. Several of the studies discussed above have considered the effectiveness of the control systems. In general, these have been shown to be broadly effective, exposure occurring where the systems are not used properly, are defective or have been deliberately disabled (Han et al. 2008).

Filtration theory indicates that commercial respirator filters should be efficient collectors of NPs owing to capture by diffusion (Aitken et al. 2004). Conventional theory suggests that the most penetrating particle size is likely to be of the order of 300–500 nm, which represents a minimum between the mechanisms of diffusion (greater for smaller particles) and impaction (greater for larger particles). A study investigating filtration performance of standard filtering facepiece respirators against monodisperse silver aerosol particles in the range of 4–30 nm diameter showed a decrease in percentage penetration with decrease in particle diameter down to 4 nm (Rengasamy et al. 2008). This efficient performance of the filter does not, however, take account of leakage around the facepiece, which is likely to be the primary issue in relation to respirator performance.

It is obvious that fugitive emissions from processes in which nanomaterials are produced could lead to increased air concentration of these nanomaterials. As well as environmental exposure in these circumstances,
it is plausible that the general public could be exposed. One application that could lead to the exposure of the general population is the use of cerium oxide as a fuel additive. Potential concentrations in air have been estimated, based on the assumptions about the quantity of cerium oxide present in fuel and the uptake of fuel containing cerium oxide, using dispersion models developed and validated by the Highways Agency (Boxall et al. 2007). Using a mix of traffic type with traffic flow at 40 km h$^{-1}$ and at 1000 vehicles per day, they estimated a cerium oxide concentration at a distance of 5 m from the road of 0.0006 mg m$^{-3}$, a very low mass concentration. They estimate, however, did not take into account the number concentration or standing traffic in congested city centre streets. They also used a simple modelling approach to estimate possible consumer exposure by inhalation and dermal exposure from spray-on sunscreen. They estimated that an inhalation NP exposure concentration of 35 mg m$^{-3}$ was plausible. Even in an occupational context, this would be a high concentration, although the duration of exposure would be very short. They also estimated that approximately 3 g might be applied daily, which would give a skin coverage of about 1 mg cm$^{-2}$. Of this, approximately 5 per cent is the NP ingredient. Although very few materials have an occupational skin exposure limit, this again would be a non-trivial exposure, even in mass terms.

6. A THEORETICAL APPROACH TO RISK ASSESSMENT

6.1. Hazard identification

From thermodynamic principles, it is logical to conclude that surface reactivity may change with particle size for any material and that the unit mass and chemical reactivity (as well as thermodynamic instability) of a compound increases as particle size decreases. However, the possible pathogenic mechanisms induced by particle exposure are very complex, depending on the route of exposure, dose, host response, susceptibility and the specific physico-chemical properties of the individual particles. The primary NP exposure can be to the lungs, skin or gut but translocation to other target organs raises the possibility that different mechanisms of toxicity, dependent on the target organ, can operate. For example, redox conditions vary from highly oxidizing on the skin or in the lung to highly reducing in the intestine or interstitial sites; this could have an important impact on oxidative stress-driven mechanisms. Thus, for a proper assessment of engineered NP hazard, it is necessary to focus on all the body systems that are potential targets for NPs.

6.2. Tools for hazard assessment

The current quantitative support tools for investigation are specified in the Organization for Economic Cooperation and Development (OECD) guideline and the new European Union regulatory framework REACH (Registration, Evaluation and Authorisation of Chemicals). They are (i) standard regulatory toxicity tests, (ii) quantitative structure–activity relationship (QSAR), and (iii) physiologically based pharmacokinetics models (PBPK).

6.2.1. Regulatory toxicology tests. The OECD guideline for the testing of chemicals has been implemented for many toxicological endpoints. Of relevance to NPs are the acute and subchronic inhalation toxicity tests. The main limitations of these tests are: some toxic endpoints are not relevant to nanotoxicology (e.g. LC50); the difficulty in aerosolizing NPs owing to their fast rate of agglomeration and the extensive use of animals for testing. The OECD is currently considering an alternative testing strategy for nanomaterials and non-inhalation toxicity tests in the spirit of reducing the need for animal experiments.

6.2.2. Quantitative structure–activity relationship. The aim of a QSAR model is to understand the properties of a chemical that influence its biological activity and to be able to predict the activity of previously untested structures/compounds. The use of a toxicity-based QSAR is a well-established approach for predicting the toxicity of chemicals for a wide variety of endpoints. The growing importance of in silico methods such as QSARs for providing information about toxicity is reflected in a number of regulatory frameworks (e.g. REACH) where these approaches are considered acceptable methods under certain conditions for filling in knowledge gaps for untested chemicals. The increasing production of novel formulations of NPs by industry poses an immediate problem for hazard and risk assessment, as many of them remain untested and thus QSARs, and in silico tools, in general, are highly desirable methods to predict their toxicity.

6.2.3. Pharmacokinetic models. There is currently no established PBPK model for the distribution of NPs in the body. NPs are larger than most molecules and the standard pharmacokinetics model transport equations need to be re-examined to assess their validity for particles. An NP model is essential for describing the exposure–dose–response relationship and extrapolation of this relationship between species.

6.2.4. In vitro–in vivo extrapolation. Information on the toxicity of chemicals can be obtained more efficiently in in vitro experiments than in vivo but translation of the results is a major issue. Dose–response modelling is required both for quantitative comparisons of in vitro with in vivo studies and to compare different in vitro studies.

6.3. Exposure assessment

The very limited data on exposure to nanomaterials in an occupational or consumer context are discussed above. The most obvious difficulties relate to identification of small mass quantities of emission and of measuring number or surface area of emissions in a background of normal ambient particles which commonly may be hundreds of thousands of particles per millilitre. Likely expenses of monitoring such emissions
in workplaces could well outweigh the benefits unless the particles prove unusually toxic, and a similar argument may apply to methods of containment and worker protection. Nevertheless, there are important scientific challenges in devising instruments that could be used both in toxicology and in environmental monitoring for measuring NPs. These challenges exist both in measuring NPs in media such as air to which people are exposed and in measuring the appropriate metric for dose when examining target organs in toxicology. Response to these challenges depends on collaboration between designers, industrial hygienists and toxicologists if the biologically relevant metric is to be measured, because this is at present still unknown.

6.4. Risk assessment

Different approaches to assess the health risk for manufactured NPs have been proposed, such as a tool for risk-level assessment and control of NP exposure. However, a quantitative risk assessment of NPs is not yet possible. The traditional risk assessment process outlined by REACH involves the use of toxicology data to obtain the derived no-effect level (DNEL) of exposure. Risk is assessed by comparing the DNEL with the exposure levels obtained from the different scenarios in the exposure assessment.

7. REGULATION: CODES OF CONDUCT AND LEGISLATION

Despite the wide range of potential exposure situations and therefore potential risks, specific regulation has been slow to emerge. One of the issues here is that because of the breadth of applications, from paints to cosmetics to medicines, many different regulatory frameworks apply. For example, in Europe, this would include regulations on worker safety, chemicals, general products, cosmetics, food, pollution, biocides, water, waste and labelling. However, reviews of regulations, for example the BRASS report in the UK (Frater et al. 2006), have consistently reported that, while the regulatory framework is capable of dealing with nano issues, the regulations are not well attuned to dealing with the specifics of nano-sized materials. Few refer to particle size as a factor to be considered. In many cases, there is nothing within the regulations that would trigger any need for change of notification requirements if a macro-sized material is replaced by a nano-sized one of the same chemical. In some cases, for example, the REACH regulations, the ‘trigger points’, in terms of production mass which would necessitate additional action, are high relative to the probable quantities of NPs which would be produced. For many types of consumer products, almost no testing is necessary before they enter the market.

Whether this situation continues remains to be seen. A recent motion passed by the European Parliament has been critical of the current situation and has called on the Commission to review all relevant legislation within 2 years to ensure safety for all applications of nanomaterials in products with potential health, environmental or safety impacts over their life cycle and to ensure that legislative provisions and instruments of implementation reflect the particular features of nanomaterials to which workers, consumers and/or the environment may be exposed. On a more positive note, there has been an increase in the guidance made available, and codes of conduct have emerged intended to provide a basis by which companies and others manufacturing or using NPs can align themselves with the principles of good practice. The European Commission has published a recommendation aimed at all those involved in nanotechnology research including funding organizations, industry, academic groups and researchers themselves. The code recognizes many of the uncertainties involved and comprises a general exhortation towards good practice. It is based on a set of general principles that call for actions aimed at guaranteeing their respect by all stakeholders. Among these is the principle that ‘research activities should be conducted in accordance with the precautionary principle, anticipating potential environmental, health and safety impacts of nanoscience and nanotechnology outcomes and taking due precautions, proportional to the level of protection’. The code also recommends that ‘As long as risk assessment studies on long-term safety are not available, research involving deliberate intrusion of nano-objects into the human body, their inclusion in food (especially in food for babies), feed, toys, cosmetics and other products that may lead to exposure to humans and the environment, should be avoided’. And, finally, the code states that ‘Students, researchers and research organisations involved in nanoscience and nanotechnology research should take specific health, safety and environmental measures adapted to the particularities of the nano-objects manipulated. Specific guidelines on the prevention of pathologies induced by nano-objects should be developed in line with the Community Strategy 2007–2014 on Health and Safety at Work’.

It remains to be seen how effective this code will be or how many national or industry organizations will adhere to its principles.

In the UK, the Health and Safety Executive (HSE) has recently published guidance on safe handling of CNTs. The HSE views CNTs as being substances of very high concern and have stated that a ‘precautionary approach should be taken to the risk management of all CNTs, unless sound documented evidence is available on the hazards from breathing in CNTs. If their use cannot be avoided, the HSE expects a high level of control to be used including a recommendation to control exposure at source by carrying out all tasks, including packaging for disposal, in a ducted fume cupboard with a high efficiency particulate air (HEPA) filter, or by using other suitable effective local exhaust ventilation with a HEPA filter (www.hse.gov.uk/pubns/5038.pdf). Other more general guidance has been provided by the British Standards Institute that has published the first extensive guide to safe handling of nanomaterials (British Standards Institute 2009). This document provides step-by-step guidance to the general approach to management of risks, information needs, hazard assessment, measurement of exposure, methods of control and disposal. It is intended to help manufacturers and users work with...
nanomaterials in a safe and responsible way and can be downloaded free from www.bsi.com/nano. Other web-based resources such as SAFENANO (www.safenano.org) are also available to help guide those unfamiliar with these issues.

8. LESSONS FOR NANOTECHNOLOGY AND NANOTOXICOLOGY RESEARCH

Consideration of the above issues leads to the conclusion that NPs (which after all are produced because their small size confers certain advantageous chemical or physical properties) might have different actions in biological systems from those of the same material at larger dimensions and that in some cases these may lead to unexpected toxic effects. It should, however, be borne in mind that workers with industrial diseases have usually been exposed to high doses of the materials over years; dose is critical in determining toxicity and dose is the product of concentration and duration of exposure. A fundamental method of avoiding harm is thus reduction or complete prevention of exposure, and this should be the aim of regulation of any NP considered likely to be toxic. We have outlined the fundamental elements of an integrated risk assessment approach. This approach originated from the assessment of health risks for chemicals and is currently adopted for manufactured NPs. An important limitation of this approach is that it is likely to produce estimates of no-effect levels that are too low to be distinguishable from the background level of exposure to airborne particles. Where it is not feasible to control exposure to an appropriate level, it is sensible to explore the possibility of hazard reduction. Current nanotoxicological research aims to identify the physico-chemical characteristics of NPs responsible for the observed health effects. These results could be incorporated in the design of new engineered NPs. The challenge is to produce new nanomaterials that are without adverse characteristics and still fulfill the industrial requirements. This approach would have the advantage of initiating a sustainable and safe nanotechnology.

Thus, two parallel streams of research are desirable: to understand the toxic properties of a range of NPs and to develop methods of assessing this toxicity in vitro and to investigate exposure scenarios in the likely life cycles of the materials being developed for the market. Both are complex, and some of the issues surrounding life cycle analysis are discussed above (§5). With respect to understanding the toxicity of NPs, the field of nanotechnologies involving particles is expanding rapidly and funding for the desirable range of research (Maynard et al. 2006) is likely to remain limited. Thus, research should aim at understanding fundamental issues relating surface activity and interaction with biological structures and systems, with the objective of determining structure–activity relationships that may be generalized for the design of new particles. Specific issues of toxicity of an individual NP should remain the responsibility of those who manufacture them. From this arises a second difficulty: what toxicological tests are appropriate for assessing new NPs? For nanotubes, it is probably sufficient to assume that they conform to the fibre paradigm (Poland et al. 2008), and thus if single long fibres are likely to be released into the air, they should be treated as asbestos without the need for further toxicology. If it can be shown that they do not release single fibres, they may be treated as non-fibrous particles. In this respect, as noted before, bundles of nanotubes will have a much larger surface area and lower aerodynamic diameter than similar-sized solid particles and thus pose particular difficulties to phagocytic cells, with as yet undetermined toxicological consequences. For other engineered NPs, standard toxicological testing, with its emphasis on carcinogenicity and foetal toxicity, is probably irrational and appropriate tests of inflammatory potential and endothelial activation need to be developed in order to detect those that might entail hazard. A third issue is the complex problem of biologically relevant measurement as mentioned below, and which is desirable both to understanding the toxicology of NPs and to assessing them in the workplace.

The most encouraging results of research in this field to date relate to those on containment and control of exposure. No one suffers adverse effects in the absence of exposure and measures to prevent exposure are neither more nor less complex than those required to prevent exposure to gases. Methods for ensuring this in research and production are available and may require to be used in cases where toxic effects might reasonably be anticipated. However, the widespread use of NPs in consumer products does give cause for concern, especially when these may be inhaled, ingested or applied to body surfaces. Definitions of what is an NP, what proportion is allowable in any mixture, how to measure and label it and how to enforce regulation present formidable problems for regulators. It would not surprise us if unexpected toxic effects, perhaps in relation to inhalation accidents, skin allergies or intestinal problems, were to come to light as a result of unregulated introduction of such particles into consumer products.

A final area of research need relates to measurement. There does not at present seem to be an alternative to counting fibres of appropriate dimensions for control of exposure to CNT, using the internationally agreed methods for asbestos, with electron microscopy. With respect to other non-fibrous particles, although the evidence is at present insufficient, on general principles it seems likely that a measure of surface area of an aerosol will prove the most useful. An instrument that measures this in real time that could be deployed as a personal sampler in workplaces is perhaps as good as could be achieved in the short term. As a longer term objective, an instrument might be developed that is able also to measure the toxic potential of the aerosol directly. This is a challenge to collaborate across the areas of instrument design, surface chemistry and toxicology.

REFERENCES


