
Research Article

Theme: Nanotechnology in Complex Drug Products: Learning from the Past, Preparing for the Future
Guest Editors: Katherine Tyner, Sau (Larry) Lee, and Marc Wolfgang

Early Development Challenges for Drug Products Containing Nanomaterials

Jennifer H. Grossman,^{1,2} Rachael M. Crist,¹ and Jeffrey D. Clogston¹

Received 14 May 2016; accepted 19 August 2016; published online 9 September 2016

Abstract. The vast majority of drug product candidates in early development fail to progress to clinics. This is true for products containing nanomaterials just as for other types of pharmaceuticals. Early development pathways should therefore place high priority on experiments that help candidates fail faster and less expensively. Nanomedicines fail for many reasons, but some are more avoidable than others. Some of the points of failure are not considerations in the development of small molecules or biopharmaceuticals, and so may be unexpected, even to those with previous experience bringing drug products to the clinic. This article reviews experiments that have proven useful in providing “go/no-go” decision-making data for nanomedicines in early preclinical development. Of course, the specifics depend on the particulars of the drug product and the nanomaterial type, and not every product shares the same development pathway or the same potential points of failure. Here, we focus on challenges that differ from those in the development of traditional small molecule therapeutics, and on experiments that reveal deficiencies that can only be corrected by essentially starting over—altering the nanomedicine to an extent that all previous characterization and proof-of-concept testing must be repeated. Conducting these experiments early in the development process can save significant resources and time and allow developers to focus on derisked candidates with a greater likelihood of ultimate success.

KEY WORDS: drug delivery; drug development; nanomedicine; preclinical; targeted delivery.

INTRODUCTION

Developing a nanomedicine shares many similarities with developing a traditional small molecule therapeutic, but also presents additional challenges. Recent reviews of clinical translation of nanomedicines include products that contain both solid structures and liquid phases, small and large molecules, and biotechnology-derived and chemically synthesized moieties (1–3). They are complex, macromolecular, and heterogeneous and change with conditions and time. This has a significant impact on the strategies and activities required for their development. This article focuses on early development—during preclinical proof-of-concept testing, prototyping, and early feasibility testing of the drug product. Early development is at the nexus of basic research and preclinical R&D and often involves collaborations between academic investigators and industry scientists.

Studies at this stage may be the first tests of a drug product’s translational potential and will provide a foundation for further preclinical development, to eventually include tests geared at meeting regulatory requirements for an Investigational New Drug (IND) application (4).

There are many important questions to answer at the outset of development of a drug product containing nanomaterials: Is the benefit offered by the product sufficient to garner investment and/or interest from pharmaceutical companies? Will it be economically feasible to characterize and manufacture with sufficient quality controls? What are the product’s liabilities and how can they be overcome? Another goal of early development testing is to weed out candidates, to allow them to “fail early, fail cheap.” Many of the nanomedicine products in development are novel delivery systems which offer putative improvements in delivery of active pharmaceutical ingredients (APIs) that already have a history of use in patients. For such products, the focus of early development proof-of-concept testing is on determining the extent to which the nanoparticle alters the pharmacokinetics (PK), biodistribution, target-cell uptake, or toxicity profile, and provides an advantage over existing formulations (5,6). For products where the nanomaterial itself is the API, or the API is a new chemical entity (NCE), the criteria for successful

¹Nanotechnology Characterization Laboratory, Cancer Research Technology Program, Leidos Biomedical Research, Inc, Frederick National Laboratory for Cancer Research, 8560 Progress Drive, Wing D, Rm 1003, Frederick, Maryland 21702, USA.

²To whom correspondence should be addressed. (e-mail: jennifer.grossman@fnlcr.nih.gov)

proof-of-concept may differ substantially and be highly product specific.

Here, we review early development considerations for drug products containing nanomaterials, highlighting pitfalls and successful strategies, with particular emphasis on experiments that can provide “go/no-go” decision-making data. Recognizing that resources may be extremely limited during early stages of development, we have restricted the areas for priority evaluation to five: (1) characterization of the raw materials used to prepare the nanoformulation, (2) evaluation of ligand and coating attachment, (3) assessment of drug and nanomaterial stability, (4) evaluation of *in vivo* stability and drug release, and (5) *in vitro* immunological evaluation (actually a set of 11 tests, noting specific categories of nanoformulations where each test may be particularly informative). Where possible, we have cited examples of methods for these assessments and have noted which tests may be informative for specific drug products and nanomaterial types, since nanomedicines include a broad set of materials. This article also reviews challenges in establishing preclinical proof-of-concept for nanomedicine drug delivery systems, and in initial scale-up of drug products containing nanomaterials.

These areas for evaluation were selected based on the results of tests conducted at the Nanotechnology Characterization Lab (NCL). NCL collaborates with more than 100 labs working in nanomedicine R&D, including large pharma, small biotechs, and academic labs. NCL has tested more than 350 different nanomedicines, many of which were in early development. These have included many different nanotech platforms, including liposomes, micelles, emulsions, dendrimers, metallic, polymeric, and more. We have had the opportunity to see why, when, and how these products fail to advance into clinical trials. This article highlights experiments that NCL has used to identify deficiencies and that we have noted are often neglected during basic research. Conducting these experiments early can help developers select candidates with greater potential for eventual success.

ADEQUATE CHARACTERIZATION

Characterization during early development differs from characterization in later development, in that it is not primarily geared at meeting regulatory requirements and providing data for the chemistry manufacturing and controls (CMC) portion of an IND application. In early development, characterization is instead focused on providing a thorough understanding of the physicochemical properties of the product and how those change with variations in synthesis processes and conditions. Such characterization of early candidates will be extremely useful later on, as it will begin to establish the acceptable range of process parameters that will become the design space of the product. If adequately characterized, early candidates used in proof-of-concept research studies can provide an understanding of the limits of formulation parameters and nanoparticle physicochemical properties which produce safe and efficacious products.

Products do not fail because of lack of characterization *per se*, but such characterization is critical for avoiding many of the pitfalls of early stage nanomedicine development (7). One of the most basic, and often overlooked, characterization

requirements is the confirmation of the structure, molecular weight, purity, or other features of raw materials used to prepare the nanoformulation. Complex starting materials such as polymers and nanoparticle platforms purchased commercially are particularly susceptible to having physical and chemical properties deviate from their theoretical/nominal values. Poly(ethylene glycol) (PEG) is one of the most widely used reagents in nanomedicine. However, the molecular weight, purity, and degree of functionalization can vary widely from manufacturer to manufacturer, or even lot to lot from a single manufacturer. Figure 1 shows characterization of several lots of 20 kDa mPEG-thiol from various manufacturers, evaluating purity using RP-HPLC with charged aerosol detection (CAD), and degree of functionalization (*i.e.*, thiol content) using Ellman’s reagent. The purity levels differed between manufacturers, and the percentage of thiol functionalization varied from nearly zero to almost complete functionalization. In this example, a few simple characterization steps on the starting reagents helped to save tremendous time and resources in troubleshooting a failed or suboptimal formulation of the much more complex nanomedicine product.

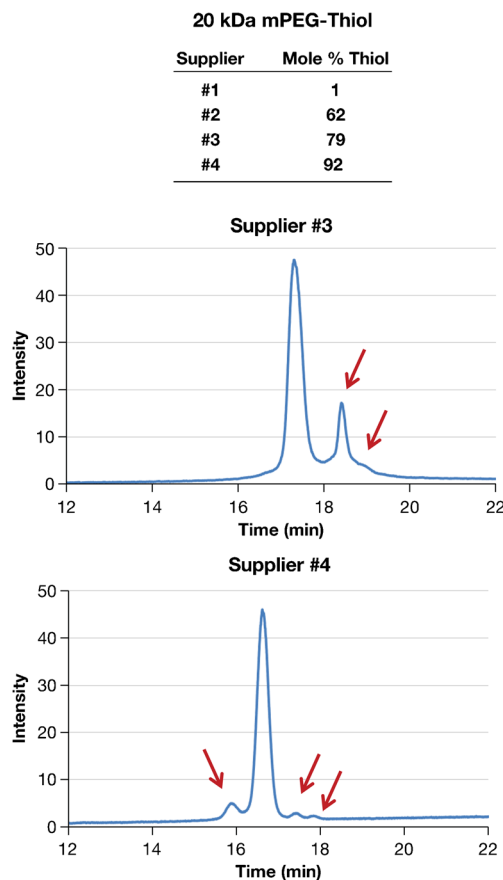


Fig. 1. Characterization of PEG starting materials; 20 kDa mPEG-thiol from several different manufacturers was characterized for thiol content using Ellman’s reagent and purity using RP-HPLC with CAD. No two batches of the functionalized polymer were identical. The four different lots evaluated for thiol content show a mole percentage ranging 1–92%, and two independent lots in the HPLC chromatograms show different levels of impurities. A quick screen of the starting materials can help save time in troubleshooting failed formulation of the nanomedicine product

For multicomponent nanomedicines, it is important to quantify the amounts of each component present in the nanomedicine, and ideally to determine the structure and interactions between those components. For liposomes and micelles, quantification of individual lipids and polymers is an important aspect of characterization, and analysis of the individual components over time can reveal degradation products generated during synthesis or during storage. However, it may not be enough to only measure the amounts of components. Nanomaterials are macromolecular systems that have to be made reproducibly, with functional attributes that also have to be reproducible. Because of this, it may also be important to include analysis of the physical state of those ingredients which impact functional aspects (*e.g.*, the release rate of drug *in vivo*). These functional attributes differentiate the lipids/polymers in a nanomaterial delivery system from compendial excipients. For example, differential scanning calorimetry (DSC) has proven to be useful for characterizing the physical state of lipids in liposome formulations and relating this to drug performance characteristics (8,9). Thermogravimetric analysis (TGA) is a similarly useful technique to assess the state of polymeric components.

Ligand and Coating Attachment

It is often extremely challenging to adequately characterize the surface properties of nanomedicines, especially those with active targeting ligands. Currently, there are no universal, or even generally applicable, techniques for quantifying targeting ligands on nanoparticles. Many targeted nanomedicines will fail to show an advantage over untargeted controls in initial testing (10–12). This may be for complex reasons associated with the interplay of passive and active targeting (13), but for nanomedicines in early development, it is often simply because the targeting ligand is not attached to the nanomaterial in sufficient quantities/densities, is inaccessible (*e.g.*, masked by polymer coating molecules (14)), or is in an inappropriate configuration for binding to receptors. It can save enormous resources to find out in early characterization experiments that a targeting molecule is not attached to the nanoparticle, rather than to make this discovery after multiple disappointing efficacy studies. This can be readily identified and fixed, more so than the complexities associated with active *vs.* passive targeting *in vivo*, and so should be given priority during early development.

The optimal ligand architecture and stoichiometry/density per particle are often unknown, yet may be critical to the product's performance. Furthermore, ligand distribution on the surface may have inherent heterogeneity due to the many potential sites of attachment on the particle, and small changes in conditions during the synthesis process may readily alter the distribution (15). The effectiveness of active targeting ligands is usually first assessed and optimized using *in vitro* biochemical and biological assays, such as competitive ELISA (16). Of course, increased cell uptake *in vitro* does not necessarily correlate with improved systemic delivery, so optimizing the per-particle ligand density based on *in vitro* experiments may not be probative of performance in animals or patients.

Surface properties are critical, even when no active targeting ligand is attached to the particle, since hydrophilic

coatings, such as PEG, can reduce interactions with plasma proteins (opsonization) and uptake from the systemic circulation by cells of the immune system (mononuclear phagocyte system, MPS) (17). Hydrophilic coatings can also reduce agglomeration of the nanoparticles, which can be important for avoiding potentially serious toxicities as will be discussed later. Subtle changes in coating quantity, density, or structure may not affect the batch-mode particle size distribution enough to be detected by routine techniques such as dynamic light scattering (DLS) or nanoparticle tracking analysis (NTA), and many hydrophilic polymers (including PEG) are not sufficiently electron dense to have contrast in transmission electron microscopy (TEM) images (Fig. 2). These coatings are deliberately chosen for their inertness, and so may not bind stains and dyes used for polymer quantitation. Phase analysis light scattering (PALS, *i.e.*, zeta potential) may be able to detect differences in surface-bound coating quantities, but will not detect unbound coating molecules, unadsorbed, or disassociated due to instability. Chromatographic separation methods using evaporative light scattering detection (ELSD) or charged aerosol detection (CAD) can displace and separate the coating from metal nanoparticles and detect/quantitate bound and unbound coating, but the particulars of the displacement and separation from the nanoparticle depend on the type of particle, its size, and the chemistry of the coating attachment (18). Knowing the limitations of the characterization tools used can help identify gaps in an overall understanding of the physical and chemical properties of the formulation.

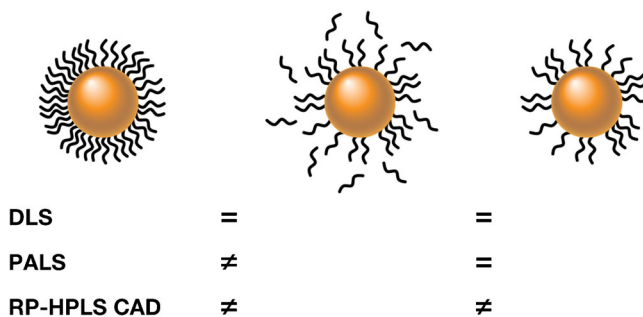


Fig. 2. Limitations in tools for surface characterization. Three PEGylated gold nanoparticles are depicted: one fully surface-functionalized (*left*); one depicting coating instability, affording a moderately functionalized particle and containing free unbound coating (*middle*); and a moderately functionalized particle with no free unbound coating (*right*). Batch-mode DLS would likely not be able to differentiate the three nanoformulations depicted, failing to detect differences in coating quantity or minor coating instability issues (the two equalities in the DLS row represent the formulations that would all appear equivalent in a DLS measurement). PALS could potentially detect differences in the quantity of surface coating but would not be affected by free, unbound surface coating in solution (there is only one equality in the PALS row, representing that the two formulations with equivalent coatings appear the same, despite differences in the amount of released coating in the buffer). RP-HPLC with CAD could be used to readily identify and quantitate differences in all three formulations (two inequalities in the RP-HPLC CAD row). It is important to know the limitations in each characterization tool. In many cases, a combination of techniques is the best approach to gaining a functional understanding of the formulation

Characterization using multiple techniques is the best way to ensure thorough understanding of the formulation.

Another approach for evaluating surface coatings is to use separation techniques such as asymmetric-flow field-flow fractionation (AF4) in line with DLS, multiangle light scattering (MALS), or other detection methods. Formulations can be separated into multiple bins with well-defined size ranges, allowing differentiation of populations with and without surface coatings. If the separation is sufficiently good, coating thicknesses and densities can be discriminated, offering a glimpse at the heterogeneity of the surface functionalization. AF4 separation can be run on multiple batches or synthesis variants to assess the reproducibility of the functionalization process and help optimize the procedure (19). In some cases, information on the coating orientation, and theoretical structure (*e.g.*, brush or mushroom (20)), can also be inferred. There are many reviews on nanoparticle surface characterization (21,22). All of them stress the importance of surface characterization but highlight challenges and the lack of broadly applicable methods.

Drug and Nanoparticulate Stability

Drug stability is an early development concern for some nanomedicines. The process of encapsulating a drug in a nanoparticle can, in some cases, cause degradation of the drug. Sometimes this is due to pH extremes caused by electrochemical gradients used for loading the drug into the nanoparticle, but it can also be caused by dimerization reactions due to the high concentrations/close proximity of the drug molecules within the particle. In addition to the chemical stability of the drug, nanomaterial stability (also called colloidal stability or particulate stability) may affect the performance of the drug product. There are a variety of types of nanomaterial instabilities, including not only dissolution of the nanoparticle, but also aggregation, agglomeration, flocculation, and precipitation. These should be monitored carefully, as the particulate stability of some nanomaterials may be exquisitely sensitive to small changes in surface charge, which may change with dilution or local pH (4). Many intravenously administered nanomedicines are given by slow infusion over a period of minutes to hours and are therefore diluted with buffers (usually saline or dextrose solutions) prior to administration. It is important to evaluate the stability of the nanomaterial in these dosing solutions, and to vary the conditions of the infusion (buffer concentration, infusion time, temperature), to fully understand the conditions under which the nanomaterial may either fall apart or aggregate.

Challenges in assessing the aggregation/agglomeration of nanomedicines under clinically relevant conditions are highlighted in Fig. 3. A drug product containing a lipid-protein nanoparticle was administered in an animal study in which adverse reactions (dyspnea, blue coloration, respiratory distress, and animal death) were observed in approximately one out of ten administrations, with no obvious relation to dose or dose-rate (bolus *vs.* slow press injection). The respiratory events were hypothesized to be linked to nanoparticle aggregation during injection. Though DLS of the formulation showed a broad size distribution, ranging from ~20 nm to over a micron, it was difficult to interpret whether the large particle size population was meaningful (the Z-average size was small, ~62 nm), since

DLS requires dilution of the samples beyond that of the administered dosing solutions. Laser diffraction experiments provided confirmatory evidence of the large size population, possibly representing aggregates, but again did not reproduce exactly the dosing conditions. AF4-DLS provided further information on the heterogeneity of the formulation, but again was not a true measure of the actual dosing solution and did not show aggregates larger than a micron. Light microscopy images, aided by a lipophilic dye, provided confirmatory evidence of large aggregates present in the actual dosing solutions for which the adverse events occurred. However, it also revealed aggregates in the dosing solutions which had not caused reactions, making it impossible to rule out if this aggregation had occurred after the material had been administered. This example highlights the challenge of analyzing the nanomaterial under the exact conditions in which it will be administered *in vivo*, *e.g.*, concentration, buffer, pH, *etc.* Many characterization techniques require dilution of the sample which may alter aggregation/agglomeration state. If the exact dosing conditions cannot be mimicked, multiple techniques must be used (as illustrated in Fig. 3) to gain an understanding of how the size distribution changes with conditions.

Nanoparticle aggregation upon intravenous administration can have potentially serious adverse effects. If aggregates interact with blood components, they can cause thrombosis or coagulation (23), and/or they can become trapped in the lungs (24), potentially causing inflammatory lesions or even embolism. Biotherapeutic aggregation has been shown to have the potential to induce a serious immunogenic response in human patients (25). However, evaluation of aggregation in dosing solutions can be challenging since there is no good technique to measure particle size in solution accurately and precisely across the entire range of relevant aggregate sizes (tens of nanometers to multiple microns (26)). DLS, which relates rapid fluctuations in scattered light due to Brownian motion to particle size, provides an accurate measure in the 20 nm to 1 μ m range, but larger particles move too slowly for their Brownian motion to be accurately measured by this technique. Very large particulates may escape detection *via* DLS simply by sedimenting out of solution (samples cannot be stirred during measurement since the technique measures particle motion). Classical laser light scattering (LS) analyzes the spatial variation in scattered light over a wide range of scattering angles, but this variation is only significant when particle size is not small compared to the wavelength of laser light (typically 635 nm), so the technique is not precise for particulates with diameters in the sub 500 nm range. Light obscuration and light microscopy can reliably detect particles >5 μ m (27), but may not detect transparent particles. In particular, the combination of DLS and light obscuration may miss aggregates in the 1–5 μ m range, which is a size range in which even deformable particulates such as fat globules can be trapped in the lungs (24). Nanoparticle tracking analysis (NTA) and resistive pulse sensing (RPS) are also potentially useful techniques for measuring nanoparticle size in solution and monitoring aggregation and agglomeration. NTA has been used recently to detect subvisible particulates in

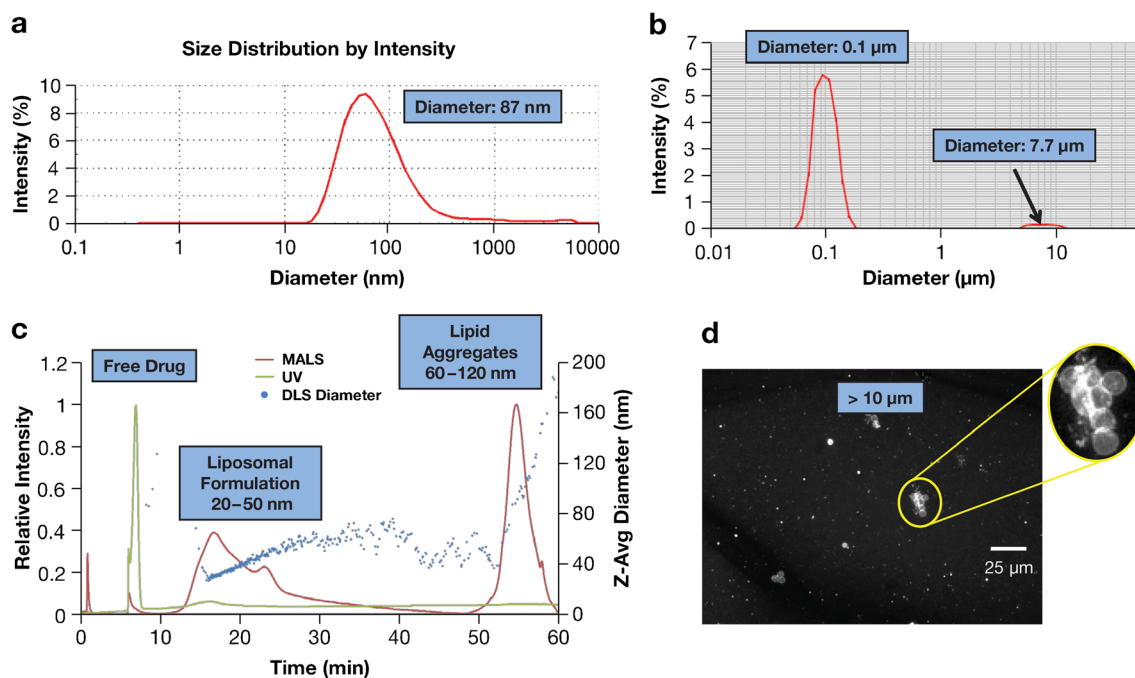


Fig. 3. Assessing aggregation/agglomeration in dosing solutions. **a** DLS measurement (10-fold dilution in PBS) of a lipid-protein nanoparticle showed a broad distribution with an intensity average diameter of 87 nm. Micron-sized populations were not definitively detected using DLS. **b** Laser diffraction (stock, no dilution) of the sample showed a 100-nm population, but readily detected a larger population $>7 \mu\text{m}$. **c** AF4 (5-fold dilution in PBS) separation further highlighted the polydispersity of the formulation, showing free drug, liposomal drug, and lipid aggregates present in the solution. **d** Light microscopy aided by a lipophilic dye confirmed the presence of $>10 \mu\text{m}$ aggregates in the dosing solution (3-fold dilution in saline). This example highlights the challenges of assessing aggregation under the exact conditions in which the nanomaterial is dosed

Peginesatide (Omontys; Affymax, Inc.) after it was voluntarily withdrawn from the market following an unexpected rise in severe adverse events (including anaphylaxis) upon exposure (28). The underlying biological mechanism for the hypersensitivity events remains under investigation (29).

Assessing the potential for aggregation upon exposure to blood is also critical, but is even more complicated. Blood clots within seconds of venipuncture, and so samples for analysis must be treated with anticoagulants that may impact nanoparticle aggregation. Analyzing samples in whole blood or plasma using routine techniques such as DLS can be very tricky, as protein signals can swamp and/or overlap with nanoparticle signals (30). One approach to this has been to incubate particles in plasma, then wash with water, PBS, or other buffer to remove any material which did not bind the nanoparticle. This approach works to provide a qualitative measure of protein binding, but may not mimic the state of the nanomaterial in a truly physiological environment. AF4 has also successfully been used to assess protein binding to nanomaterials (31). Unlike batch-mode DLS, AF4 can discern different populations that may have more or less protein binding, but again this approach does not mimic a true physiological environment. Quartz crystal microbalance by dissipation (QCM-D) and isothermal calorimetry (ITC) have also been used, but have similar limitations (32,33). Quantitative assessment of nanoparticle size and aggregation in a truly physiological environment remains one of nanomedicine's biggest characterization challenges. It is hoped that additional studies using the abovementioned techniques will lead to future advances in this area.

***In Vivo* Stability and Drug Release**

Another challenge is assessing *in vivo* stability of the nanomedicine, including the extent and time dependence of drug release from the nanoparticle in blood. We classify this as an early development consideration because so many promising nanomedicine formulations fall apart seconds after contact with blood, even when they exhibit excellent stability in buffer (7). If the drug escapes instantly from the nanomaterial, the nanomaterial is little more than a solubilizing agent, and complicated multistage delivery systems, targeting ligands, and surface chemistries are useless. It is a “back to the drawing board” moment, and the earlier in the development process this can be tested, the better.

Nanoparticle stability can be assessed *in vitro* using a biological matrix, mimicking *in vivo* conditions, as part of the early development process. Unfortunately, it is often difficult to separate released drug from both nanoparticle-bound and plasma-protein-bound drug without artefactually extracting the drug from the nanomedicine. *In vitro* methods are also challenged by the difficulty of mimicking the complex equilibrium of reversible nanocarrier binding of the released drug and reversible protein binding that exists *in vivo* (34). Dialysis, ultrafiltration, ultracentrifugation, size exclusion, ion exchange, solid-phase extraction, and liquid-liquid extraction have all been used (and are reviewed here (35)), but at the time of writing, there is no universal, or even generally applicable, method for nanomedicines.

Our laboratory has recently developed a technique that has improved accuracy and precision for quantitating *in vitro* drug release, and is potentially more broadly applicable than existing methods (36). The uniqueness of the technique not only allows for quantitation of drug release from the nanoparticle, but differentiation between free-unbound drug and protein-bound drug fractions. The method uses a stable isotopically labeled version of the free drug to correct for protein and formulation binding in plasma. The nanoformulation and the stable isotopically labeled drug are allowed to equilibrate in plasma, an aliquot is subjected to separation using ultrafiltration, then the normoisotopic and isotopically labeled drug concentrations in the retentate and filtrate are measured using LC-MS (36). Since the stable isotope control essentially corrects for the particulars of the nanomedicine and released drug interaction with plasma proteins and the ultrafiltration device, we believe this method will be broadly applicable to a variety of nanomedicines and can also be used to measure released drug in plasma samples from nanomedicine pharmacokinetic studies.

The utility of this method is highlighted in Fig. 4, where a liposomal docetaxel formulation was compared against two immediate-releasing formulations of the drug (Taxotere and acetonitrile-solubilized drug). Assessment of drug release was performed in a human plasma matrix to mimic physiological conditions. The liposomal formulation released 90–96% of the drug at the earliest measurement ($t=0$ min), whereas the Taxotere and acetonitrile-solubilized drug showed 100% release. The remaining drug in the liposomal formulation was released in less than 30 min, suggesting the formulation was not stable in human plasma matrix (*e.g.*, the release of doxorubicin from Doxil occurs much more slowly, indicating stability and controlled release). The liposomal formulation had shown no signs of instability in buffer.

Immunotoxicity

Though some drug products containing nanomaterials are made entirely from materials that are generally recognized as safe (GRAS), many incorporate novel, unique excipients that have not previously been tested in patients. Many contain both biotechnology-derived (37,38) and small molecule (39) components in new combinations that may pose unknown safety risks. It may be difficult to predict how the human immune system will respond to the novel product—if it will be immunostimulatory, inhibitory, or immunologically inert. Early consideration of the immunological responses elicited from nanomaterial interaction with blood components and immune cells can help avoid a late-stage failure of the product over serious immunological toxicities. Many *in vitro* assays have been developed for evaluation of nanoparticle immunotoxicity, and many have been shown to be reasonably predictive of *in vivo* responses (40–42).

Table I lists several key *in vitro* immunology tests that can be critical for derisking early development nanomedicines, as well as their potential associated *in vivo* consequences (40). A substantial fraction of nanomedicines in early preclinical development are contaminated with endotoxin above allowable levels for parenteral administration to humans (7). We recommend measuring endotoxin levels prior

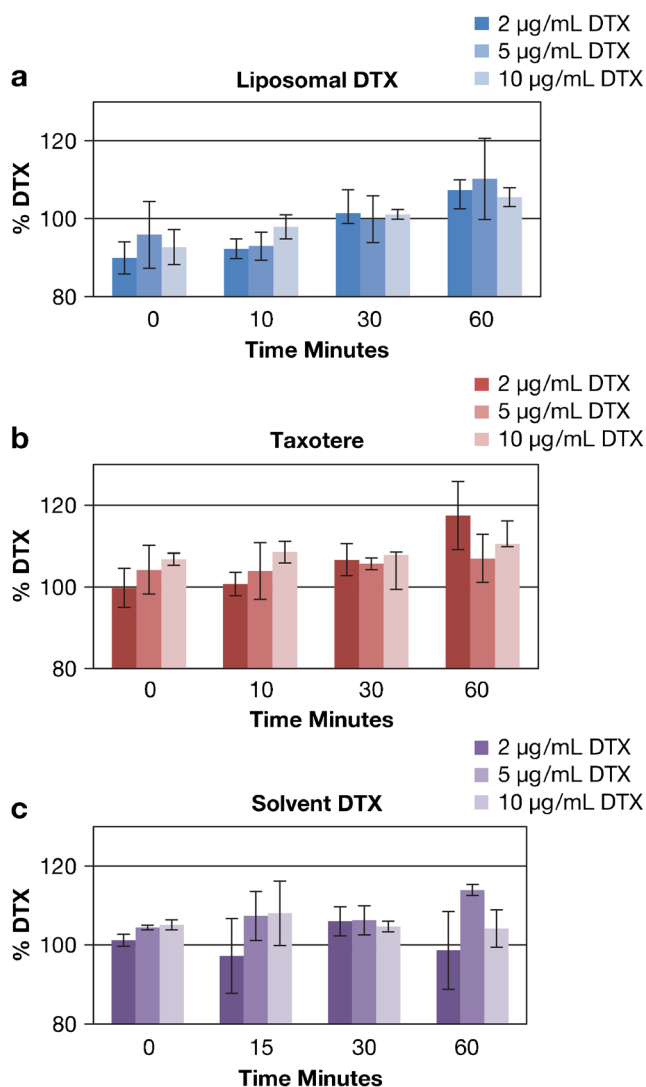


Fig. 4. *In vitro* drug release in plasma. The *in vitro* drug release in plasma for **a** a liposomal docetaxel (DTX) formulation was compared to two commercial unstable, immediate-releasing formulations of docetaxel, **b** Taxotere, and **c** acetonitrile-solubilized docetaxel. The liposomal formulation released 90–96% of the drug by the earliest time point measured (0 min), whereas the Taxotere and acetonitrile-solubilized formulations showed 100% release. The remaining $\leq 10\%$ drug in the liposomal formulation was released within 30 min, suggesting the formulation was not stable upon contact with plasma. Assays to measure drug release in plasma can be used to quickly screen for unstable formulations

to any of the other tests in Table I, since endotoxin can contribute to positive responses in some of the other assays (43,44). Endotoxin testing will be particularly important for drug products with components produced in *Escherichia coli*. Filtration, purification, or sterilization may not be possible without altering critical attributes of the formulation. *In vitro* analysis of hemolysis has been shown to be highly correlative to *in vivo* response. An *in vitro* hemolysis rate of 2–5% is considered mildly hemolytic, while $>5\%$ is considered strongly hemolytic (45). Hemolysis is often a concern for drug products with components that are cationic or contain surfactants. Even if the nanoformulation itself is neutrally charged, if it releases cationic components when it dissolves in

Table I. Immunological Considerations for Early Development

Early development <i>in vitro</i> tests	Immunological responses	Formulation properties of concern
Endotoxin	Inflammation, septic shock	Components produced in <i>E. coli</i>
Hemolysis	Anemia, hemorrhage	Cationic components; contain surfactants
Platelet aggregation	Thrombogenesis	Cationic components
Plasma coagulation	Thrombogenesis	Anionic components; nanoparticles for nucleic acid delivery
Complement activation	Anaphylaxis, CARPA	Nanoparticles for nucleic acid delivery
Total protein binding	Predictive of MPS uptake	Particles >300 nm; nonsurface coated
Phagocytosis	Predictive of MPS uptake	Particles >300 nm; nonsurface coated
Leukocyte procoagulant activity	Thrombogenesis	Cationic components
Leukocyte proliferation	Inhibits immunogenic recognition of pathogens	–
Cytokine induction	Cytokine storm, DIC	Nanoparticles for nucleic acid delivery; components produced in <i>E. coli</i>

These *in vitro* immunological tests are recommended as part of the early preclinical development process (40). Types of nanomaterials where each test is likely to be particularly informative (*i.e.*, reveal a potential concern) are listed in the third column. Many of these *in vitro* screens have strong correlation and predictability to *in vivo* outcomes (42)

CARPA complement activation-related pseudoallergy, *MPS* mononuclear phagocytic system, *DIC* disseminated intravascular coagulation

blood, these can cause significant hemolysis. *In vitro* tests for platelet aggregation, plasma coagulation times, and leukocyte procoagulant activity are good indicators of *in vivo* thrombogenic reactions. Leukocyte procoagulant activity may be an important screen for drug products with cationic components. Assays for complement activation, when using human or nonhuman primate blood, have also been good predictors of *in vivo* reactions such as complement activation-related pseudoallergy (*CARPA*). Biological matrices from other animals are not recommended for *in vitro* screening of *CARPA*. Complement activation is of particular concern for nanoparticles for nucleic acid delivery and for lipid excipients. Protein binding (opsonization) and phagocytosis assays are good indicators of *in vivo* biodistribution and accumulation in organs of the MPS. Total protein binding can also be used as a metric for optimizing polymer coating coverage and stability (46). Leukocyte proliferation assays have shown moderate *in vitro* to *in vivo* correlation, but still are considered a high priority assay for early immunological screening. Finally, *in vitro* evaluation of cytokine and interferon production are excellent indicators of cytokine storm reactions and disseminated intravascular coagulation (*DIC*) or *DIC*-like reactions. Cytokine/interferon screening may be important for drug products for nucleic acid delivery and those with components produced in *E. coli* (36). Most of these *in vitro* assays have been validated and verified to work for a variety of nanomaterials. That this *in vitro* screen can be performed relatively quickly, and it can help avoid deleterious *in vivo* outcomes, make it a critical part of the early development process of drug products containing nanomaterials.

Whenever possible, the individual components of a nanomedicine drug product should be screened alongside the final product. This is because the individual components may elicit separate responses that are then either additive (or even synergistic) or inhibitory. If this is discovered early, certain inessential reagents or raw materials that amplify negative responses to essential components may be swapped out for others. For example, the screening scheme in Table I might reveal complement activation by certain lipids in a

hypothetical nanomedicine for delivery of therapeutic oligonucleotides. If the oligo also activated complement (this is not far-fetched—certain lipid excipients have been shown to activate complement (47), as have certain therapeutic oligonucleotides (37)), developers could try alternate lipids and/or oligos. Such optimization of the nanomedicine will be less costly during early development, since fewer experiments have been conducted, and thus fewer must be repeated to conclude that the optimization has not altered the drug product's performance. If the developers decide to go ahead with the combination, at least they do so with awareness of the risks, and so are vigilant to possible future indications that their product may cause *CARPA*.

PRECLINICAL PROOF-OF-CONCEPT

No preclinical model can perfectly reproduce all the elements of human disease, and many have noted the limitations of preclinical models for predicting clinical efficacy (48,49). To maximize predictability, drug products containing nanomaterials should be evaluated in multiple models, in comparison to carefully selected controls (*e.g.*, an empty nanomaterial control and untargeted controls, as relevant). The nanomedicine should also be compared to the standard of care treatment, and the route of administration should mimic the clinical case as closely as possible. Many intravenous drug products are administered clinically *via* slow infusion and/or *via* multiple injections, which may be impractical in rodents, but which should be considered and tested if possible, as these may lessen the advantage of controlled release nanomedicines over the standard of care in patients. Preclinical oncology models should be selected carefully, since there may be structural differences between naturally occurring tumors and subcutaneous or orthotopic xenograft models which can impact nanoparticle delivery. For example, variations in tumor vascularity may be particularly important for proof-of-concept efficacy studies for cancer drug products containing nanomaterials (46).

In vitro experiments can be useful for establishing proof-of-concept, but must be regarded with cautious attention to their limitations. *In vitro* systems never allow full evaluation of the effects of distribution, or mimic exactly the conditions the nanomedicine will face *in vivo*. Aspects such as targeting can never be fully evaluated *in vitro* and will require *in vivo* proof-of-concept experiments. Nanoparticles loaded with fluorescent dyes are also frequently used to study the *in vivo* distribution of nanomedicines in preclinical animal models. This may be problematic, as many dyes leak from the nanoparticles, do not mimic the release rates of carried drugs, and are therefore misleading indicators of biodistribution, cellular uptake, and intracellular distribution (50).

It is important that proof-of-concept studies be conducted with thoroughly characterized material, even though the material will inevitably be optimized during later development, and even though the regulatory filing will only contain characterization data on the final, optimized formulation. Thorough characterization of early candidates will provide confidence that later candidates (formulations produced for safety and efficacy testing in later development and eventual GLP studies) reproduce critical attributes. This is important not only for physicochemical properties such as particle size distribution, charge, agglomeration state, composition, and purity, but also for performance characteristics such as drug release and target binding. For drug products containing nanomaterials, it is particularly risky to advance products based on single batch results. Instead, developers should generate and test multiple process/formulation variants to understand the tolerances of synthesis processes and determine specifications. Nanomedicine developers have cited this type of “combinatorial screening and optimization” analysis of a library of multiple variants as a best practice for nanomedicine development (16). These developers intentionally synthesized and tested pilot-scale batches of process variants by varying the manufacturing conditions (*e.g.*, different pH, temperature, different amount of lipids, polymers, *etc.*), then characterized and tested these *in vitro* and in pharmacokinetic studies to discriminate changes that were significant. Developers have also generated small-scale batches of specific structures (*e.g.*, with differences in particle size), or formulation variants, and characterized and tested these for *in vitro*, pharmacokinetic, and efficacy differences (51). This can be an effective way to determine the critical attributes of a nanomedicine and to link those attributes to process variables. This type of approach can greatly facilitate the scale-up of nanomedicine products.

Stability and Compatibility with Scale-Up

Drug products containing nanomaterials, especially those functionalized with targeting ligands, are often expensive to produce and may be the result of low yield processes. Many early development investigators therefore opt to conduct proof-of-concept experiments with freshly made material, and have not performed studies to examine the stability of the material over time or in response to varied conditions (*e.g.*, upon freeze-thaw, varied storage temperatures, stored in different containers, *etc.*). Such practice limits the characterization that can be done on each batch of material, which may preclude a thorough understanding of the batch-to-batch consistency of the product. Our lab has previously published

recommendations for monitoring batch-to-batch consistency of drug products containing nanomaterials (7,52). While not all characterization parameters need to be remeasured for each batch, it is important to establish the most meaningful lot release assays for each individual formulation, for example size/polydispersity, drug release, *in vitro* biological activity, *etc.* Biological screening of multiple well-characterized batches will help to elucidate the most critical parameters to monitor for each individual formulation.

Another consideration: If it turns out later that the material has a limited shelf life stability (*e.g.*, is only stable for hours to days), this may greatly limit the prospects of the drug product. In these instances, developers must figure out and optimize lyophilization/reconstitution conditions that do not alter the drug product, which may be challenging, particularly for intricate and/or delicate nanomaterials. In addition to shelf life, there are an almost overwhelming number of interactions (*e.g.*, with filters, storage containers, syringes, in transit, *etc.*) which must be examined for a thorough stability analysis. Due to their high surface-to-volume ratios, nanomaterials are reactive, and physical interactions with surfaces may alter the formulation or reduce the administered dose of the drug product (4). In general, developers should aim for at least 3 months stability under conditions reasonably expected to be encountered in the lab, clinic, and in transit, and the stability should be optimized as a part of early development.

Nanomedicines are complex biophysical systems, where even small changes to synthesis processes frequently cause alterations to the drug product, and it may be impossible to predict if and/or how such alterations will impact performance. Efforts to move toward processes at clinical scale may therefore require repeating earlier characterization and proof-of-concept testing to ensure the impact of the changed processes is thoroughly understood. Researchers with backgrounds in small molecule development are often resistant to address stability and scale-up during early development, arguing that it does not make sense to spend time and resources optimizing the synthesis and evaluating stability, or in developing scaled-up processes, for material that has not been fully established to work. However, in many cases, the intricacy of nanomedicines, their sensitivity to conditions, and the difficulties associated with their characterization make it important to address stability and scale-up as early as possible. The more complicated the formulation, the earlier these issues should be examined and addressed.

Many nanomedicines fail during scale-up when characterization techniques used for lot release miss subtle variations in process parameters. We have previously discussed the importance of including biological assays for potency in quality assurance testing regimens during scale-up, since analytical and physicochemical characterization may not be adequate to detect differences in product efficacy (6). Quality by design (QbD) approaches, such as testing libraries of formulation and process variants, can be a significant help in avoiding failures during scale-up. Characterization during early development scale-up can also make use of research techniques that may be too unwieldy or time consuming for lot release during later stage development. Once the nanomaterial is well understood and its synthesis optimized and controlled, these techniques can be replaced by faster,

less resource-intensive methods that require less sample prep or methods development—provided the relationship between the measurements and their sensitivities has been thoroughly established.

For example, DLS is widely used for analysis of nanoparticle size in suspension as it provides a quick measure of hydrodynamic size and size distribution. However, batch-mode DLS may not be sufficiently sensitive to detect process

Batch #	Z-Avg nm	Pdl	Int-Peak nm	% Int-Peak
A1	28 ± 0	0.32 ± 0.01	40 ± 1	92 ± 0
A2	27 ± 1	0.48 ± 0.06	44 ± 1	90 ± 0
A3	19 ± 0	0.63 ± 0.00	45 ± 1	81 ± 0
A4	21 ± 4	0.28 ± 0.02	23 ± 1	91 ± 2

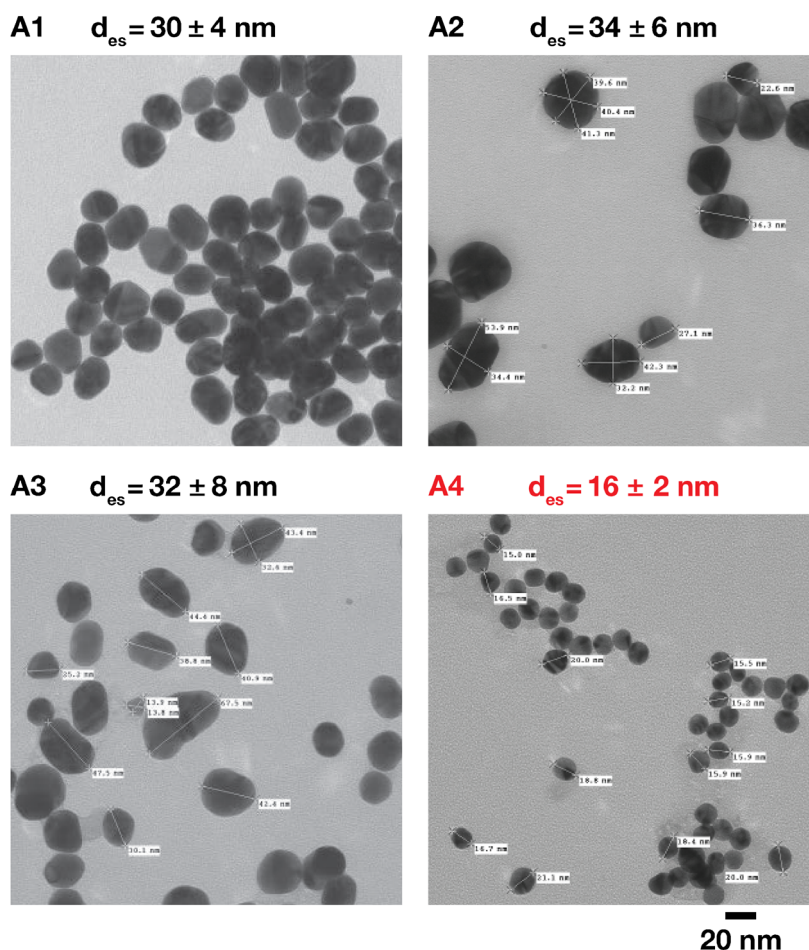


Fig. 5. DLS and TEM screening of process variants. This figure shows the elements of an extensive analysis of multiple batches of gold nanoparticles during scale-up of a drug product containing the particles. DLS and TEM measurements are shown for four batches of gold nanoparticles made using the same synthesis procedure, in which Na-citrate was added to a boiling aqueous solution of HAuCl_4 . Both DLS and TEM measurements identify batch A4 as an outlier, with substantially smaller particle size than the other batches, though only TEM allows analysis of differences in the shape distribution. The comparison of DLS and TEM for a larger set of process variants, where the reaction conditions were tested more extensively (*e.g.*, Na-citrate and HAuCl_4 were mixed at room temperature then heated to boiling, HAuCl_4 was added to boiling Na-citrate solution, or Na-citrate and HAuCl_4 were added to boiling H_2O) was previously published here (52). Testing during early development scale-up can bridge the toolsets of basic and preclinical research and begin to establish quality control testing criteria for the product

variations, and separation such as AF4 in line with DLS may be required to increase resolution (53), or alternate methods evaluated (e.g., NTA, RPS), until a combination of characterization techniques is arrived at that successfully detects significant process variations or changes to critical attributes. Transmission electron microscopy (TEM) image analysis can be used to characterize the size and shape distribution of the nanomaterial (54) while optimizing scaled-up synthesis procedures, and the results can be compared to establish the sensitivity of each technique to process changes. Figure 5 shows TEM and DLS analysis of multiple batches of gold nanoparticles during an evaluation of scale-up processes. All the batches were made by adding Na-citrate to an aqueous solution of boiling HAuCl₄ with slight differences in mixing conditions. Here, both DLS and TEM identified batch A4 as an outlier, with substantially smaller particle size than the other batches. The comparison of DLS and TEM for a larger set of process variants was previously published (52). Simultaneous analysis with orthogonal techniques (e.g., TEM, DLS, NTA, RPS) can be helpful while the synthesis is being optimized and can establish the ability of particular techniques to detect variations in conditions or processes.

CONCLUSIONS

Pharmaceutical drug product candidates should be tested in experiments which expose deficiencies that may be “show stoppers” or “back to the drawing board moments” as early as possible. “Fail early, fail cheap” is a generally accepted means of improving the efficiency of pharmaceutical R&D and is not unique to nanomedicine (55). However, the tests which identify common points of failure for nanomedicines are different than those for small molecule drugs or biopharmaceuticals. Here, we have reviewed five go/no-go evaluation sets for nanomedicines in early development: (1) characterization of starting materials, (2) evaluation of ligand and surface coatings, (3) analysis of drug and nanomaterial stability, (4) assessment of *in vivo* stability and drug release, and (5) evaluation of *in vitro* immunological responses.

Testing in these five areas is obviously no guarantee of the eventual success of a nanomedicine product. They are recommended for early development because they are often overlooked during basic research, but potentially reveal critical deficiencies that will be expensive to correct later on. These tests can also help establish a preliminary understanding of the relationship between the physicochemical characteristics of the product and its performance and safety. This is a first step toward what will eventually become the critical quality attributes and quality control testing criteria for the product.

ACKNOWLEDGMENTS

This project has been funded in whole or in part with Federal funds from the National Cancer Institute, National Institutes of Health, under Contract No. HHSN261200800001E. The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government. The authors thank the scientists and staff of the

Nanotechnology Characterization Lab (NCL), with special thanks to Scott McNeil, Stephan Stern, Marina Dobrovolskaia, and Magdalena Scully for helpful discussions and Allen Kane for assistance with figures.

REFERENCES

1. Min Y, Caster JM, Eblan MJ, Wang AZ. Clinical translation of nanomedicine. *Chem Rev.* 2015;115(19):11147–90.
2. Noorlander CW, Kooi MW, Oomen AG, Park MV, Vandebriel RJ, Geertsma RE. Horizon scan of nanomedicinal products. *Nanomedicine (Lond).* 2015;10(10):1599–608.
3. Luo D, Carter KA, Lovell JF. Nanomedical engineering: shaping future nanomedicines. *Wiley Interdiscip Rev Nanomed Nanobiotechnol.* 2015;7(2):169–88.
4. Tyner KM, Zou P, Yang X, Zhang H, Cruz CN, Lee SL. Product quality for nanomaterials: current U.S. experience and perspective. *Wiley Interdiscip Rev Nanomed Nanobiotechnol.* 2015;7(5):640–54.
5. Bae YH, Park K. Targeted drug delivery to tumors: myths, reality and possibility. *J Control Release.* 2011;153(3):198–205.
6. Stern ST, Hall JB, Yu LL, Wood LJ, Paciotti GF, Tamarkin L, *et al.* Translational considerations for cancer nanomedicine. *J Controlled Release.* 2010;146(2):164–74.
7. Crist RM, Grossman JH, Patri AK, Stern ST, Dobrovolskaia MA, Adisheshaiah PP, *et al.* Common pitfalls in nanotechnology: lessons learned from NCI's Nanotechnology Characterization Laboratory. *Integr Biol (Camb).* 2013;5(1):66–73.
8. Dicko A, Kwak S, Frazier AA, Mayer LD, Liboiron BD. Biophysical characterization of a liposomal formulation of cytarabine and daunorubicin. *Int J Pharm.* 2010;391(1–2):248–59.
9. Wei X, Cohen R, Barenholz Y. Insights into composition/structure/function relationships of Doxil gained from “high-sensitivity” differential scanning calorimetry. *Eur J Pharm Biopharm.* 2016.
10. Huang X, Peng X, Wang Y, Wang Y, Shin DM, El-Sayed MA, *et al.* A reexamination of active and passive tumor targeting by using rod-shaped gold nanocrystals and covalently conjugated peptide ligands. *ACS Nano.* 2010;4(10):5887–96.
11. Kunjachan S, Pola R, Gremse F, Theek B, Ehling J, Moeckel D, *et al.* Passive versus active tumor targeting using RGD- and NGR-modified polymeric nanomedicines. *Nano Lett.* 2014;14(2):972–81.
12. Duncan R, Gaspar R. Nanomedicine(s) under the microscope. *Mol Pharm.* 2011;8(6):2101–41.
13. Sykes EA, Chen J, Zheng G, Chan WC. Investigating the impact of nanoparticle size on active and passive tumor targeting efficiency. *ACS Nano.* 2014;8(6):5696–706.
14. Klibanov AL, Maruyama K, Beckerleg AM, Torchilin VP, Huang L. Activity of amphipathic poly(ethylene glycol) 5000 to prolong the circulation time of liposomes depends on the liposome size and is unfavorable for immunoliposome binding to target. *Biochim Biophys Acta.* 1991;1062(2):142–8.
15. Mullen DG, Fang M, Desai A, Baker JR, Orr BG, Banaszak Holl MM. A quantitative assessment of nanoparticle-ligand distributions: implications for targeted drug and imaging delivery in dendrimer conjugates. *ACS Nano.* 2010;4(2):657–70.
16. Hrkach J, Von Hoff D, Mukkaram Ali M, Andrianova E, Auer J, Campbell T, *et al.* Preclinical development and clinical translation of a PSMA-targeted docetaxel nanoparticle with a differentiated pharmacological profile. *Sci Transl Med.* 2012;4(128):128ra39.
17. Jekerst JV, Lobovkina T, Zare RN, Gambhir SS. Nanoparticle PEGylation for imaging and therapy. *Nanomedicine (Lond).* 2011;6(4):715–28.
18. Smith MC, Crist RM, Clogston JD, McNeil SE. Quantitative analysis of PEG-functionalized colloidal gold nanoparticles using charged aerosol detection. *Anal Bioanal Chem.* 2015;407(13):3705–16.
19. Makan AC, Spallek MJ, du Toit M, Klein T, Pasch H. Advanced analysis of polymer emulsions: particle size and particle size distribution by field-flow fractionation and dynamic light scattering. *J Chromatogr A.* 2016;1442:94–106.
20. Perry JL, Reuter KG, Kai MP, Herlihy KP, Jones SW, Luft JC, *et al.* PEGylated PRINT nanoparticles: the impact of PEG density on

- protein binding, macrophage association, biodistribution, and pharmacokinetics. *Nano Lett.* 2012;12(10):5304–10.
21. Baer DR, Engelhard MH, Johnson GE, Laskin J, Lai J, Mueller K, *et al.* Surface characterization of nanomaterials and nanoparticles: important needs and challenging opportunities. *J Vac Sci Technol A.* 2013;31(5):50820.
 22. Rabanel JM, Hildgen P, Banquy X. Assessment of PEG on polymeric particles surface, a key step in drug carrier translation. *J Controlled Release.* 2014;185:71–87.
 23. Ilinskaya AN, Dobrovolskaia MA. Nanoparticles and the blood coagulation system. Part II: safety concerns. *Nanomedicine (Lond).* 2013;8(6):969–81.
 24. (USP) USP. <729> Globule size distribution in lipid injectable emulsions. Physical tests. Rockville, MD 2013.
 25. Joubert MK, Hokom M, Eakin C, Zhou L, Deshpande M, Baker MP, *et al.* Highly aggregated antibody therapeutics can enhance the in vitro innate and late-stage T-cell immune responses. *J Biol Chem.* 2012;287(30):25266–79.
 26. den Engelsman J, Garidel P, Smulders R, Koll H, Smith B, Bassarab S, *et al.* Strategies for the assessment of protein aggregates in pharmaceutical biotech product development. *Pharm Res.* 2011;28(4):920–33.
 27. Driscoll DF, Etzler F, Barber TA, Nehne J, Niemann W, Bistrrian BR. Physicochemical assessments of parenteral lipid emulsions: light obscuration versus laser diffraction. *Int J Pharm.* 2001;219(1–2):21–37.
 28. Kotarek J, Stuart C, De Paoli SH, Simak J, Lin TL, Gao Y, *et al.* Subvisible particle content, formulation, and dose of an erythropoietin peptide mimetic product are associated with severe adverse postmarketing events. *J Pharm Sci.* 2016;105(3):1023–7.
 29. Weaver JL, Boyne M, Pang E, Chimalakonda K, Howard KE. Nonclinical evaluation of the potential for mast cell activation by an erythropoietin analog. *Toxicol Appl Pharmacol.* 2015;287(3):246–52.
 30. Wolfram J, Suri K, Yang Y, Shen J, Celia C, Fresta M, *et al.* Shrinkage of pegylated and non-pegylated liposomes in serum. *Colloids Surf B Biointerfaces.* 2014;114:294–300.
 31. Capomaccio R, Jimenez IO, Colpo P, Gilliland D, Ceccone G, Rossi F, *et al.* Determination of the structure and morphology of gold nanoparticle-HSA protein complexes. *Nanoscale.* 2015;7(42):17653–7.
 32. Aramesh M, Shimoni O, Ostrikov K, Prawer S, Cervenka J. Surface charge effects in protein adsorption on nanodiamonds. *Nanoscale.* 2015;7(13):5726–36.
 33. Cedervall T, Lynch I, Lindman S, Berggard T, Thulin E, Nilsson H, *et al.* Understanding the nanoparticle-protein corona using methods to quantify exchange rates and affinities of proteins for nanoparticles. *Proc Natl Acad Sci U S A.* 2007;104(7):2050–5.
 34. Modi S, Anderson BD. Determination of drug release kinetics from nanoparticles: overcoming pitfalls of the dynamic dialysis method. *Mol Pharm.* 2013;10(8):3076–89.
 35. Ambardekar VV, Stern ST. NBCE pharmacokinetics and drug release methods. In: Crommelin DJA, de Viegler JSB, editors. *Non-biological complex drugs: the science and the regulatory landscape.* Springer; 2015. p. 261–87.
 36. Skoczen S, McNeil SE, Stern ST. Stable isotope method to measure drug release from nanomedicines. *J Controlled Release.* 2015;220(Pt A):169–74.
 37. Levin AA. A review of the issues in the pharmacokinetics and toxicology of phosphorothioate antisense oligonucleotides. *Biochim Biophys Acta.* 1999;1489(1):69–84.
 38. Tamilvanan S, Raja NL, Sa B, Basu SK. Clinical concerns of immunogenicity produced at cellular levels by biopharmaceuticals following their parenteral administration into human body. *J Drug Target.* 2010;18(7):489–98.
 39. Weiszhar Z, Czucz J, Revesz C, Rosivall L, Szebeni J, Rozsnyay Z. Complement activation by polyethoxylated pharmaceutical surfactants: Cremophor-EL, Tween-80 and Tween-20. *Eur J Pharm Sci.* 2012;45(4):492–8.
 40. Dobrovolskaia MA. Pre-clinical immunotoxicity studies of nanotechnology-formulated drugs: challenges, considerations and strategy. *J Controlled Release.* 2015;220(Pt B):571–83.
 41. Dobrovolskaia MA, McNeil SE. Immunological properties of engineered nanomaterials. *Nat Nano.* 2007;2(8):469–78.
 42. Dobrovolskaia MA, McNeil SE. Understanding the correlation between in vitro and in vivo immunotoxicity tests for nanomedicines. *J Controlled Release.* 2013;172(2):456–66.
 43. Dobrovolskaia MA, Neun BW, Clogston JD, Ding H, Ljubimova J, McNeil SE. Ambiguities in applying traditional Limulus amoebocyte lysate tests to quantify endotoxin in nanoparticle formulations. *Nanomedicine (Lond).* 2010;5(4):555–62.
 44. Dobrovolskaia MA, Neun BW, Clogston JD, Grossman JH, McNeil SE. Choice of method for endotoxin detection depends on nanoformulation. *Nanomedicine (Lond).* 2014;9(12):1847–56.
 45. International A. ASTM F756 - 13 standard practice for assessment of hemolytic properties of materials. West Conshohocken, PA2013
 46. Adiseshaiah PP, Hall JB, McNeil SE. Nanomaterial standards for efficacy and toxicity assessment. *Wiley Interdiscip Rev Nanomed Nanobiotechnol.* 2010;2(1):99–112.
 47. Szebeni J, Muggia F, Gabizon A, Barenholz Y. Activation of complement by therapeutic liposomes and other lipid excipient-based therapeutic products: prediction and prevention. *Adv Drug Deliv Rev.* 2011;63(12):1020–30.
 48. Johnson JI, Decker S, Zaharevitz D, Rubinstein LV, Venditti JM, Schepartz S, *et al.* Relationships between drug activity in NCI preclinical in vitro and in vivo models and early clinical trials. *Br J Cancer.* 2001;84(10):1424–31.
 49. Gould SE, Junttila MR, de Sauvage FJ. Translational value of mouse models in oncology drug development. *Nat Med.* 2015;21(5):431–9.
 50. Snipstad S, Hak S, Baghirov H, Sulheim E, Morch Y, Lelu S, *et al.* Labeling nanoparticles: dye leakage and altered cellular uptake. *Cytometry A.* 2016.
 51. Mamidi RN, Weng S, Stellar S, Wang C, Yu N, Huang T, *et al.* Pharmacokinetics, efficacy and toxicity of different pegylated liposomal doxorubicin formulations in preclinical models: is a conventional bioequivalence approach sufficient to ensure therapeutic equivalence of pegylated liposomal doxorubicin products? *Cancer Chemother Pharmacol.* 2010;66(6):1173–84.
 52. Wei A, Mehtala JG, Patri AK. Challenges and opportunities in the advancement of nanomedicines. *J Controlled Release.* 2012;164(2):236–46.
 53. Hansen M, Smith MC, Crist RM, Clogston JD, McNeil SE. Analyzing the influence of PEG molecular weight on the separation of PEGylated gold nanoparticles by asymmetric-flow field-flow fractionation. *Anal Bioanal Chem.* 2015;407(29):8661–72.
 54. Hinterwirth H, Wiedmer SK, Moilanen M, Lehner A, Allmaier G, Waitz T, *et al.* Comparative method evaluation for size and size-distribution analysis of gold nanoparticles. *J Sep Sci.* 2013;36(17):2952–61.
 55. Paul SM, Mytelka DS, Dunwiddie CT, Persinger CC, Munos BH, Lindborg SR, *et al.* How to improve R&D productivity: the pharmaceutical industry's grand challenge. *Nat Rev Drug Discov.* 2010;9(3):203–14.