Increasing the Impact of Materials in and beyond Bio-Nano Science

Mattias Björnmal, Matthew Faria, and Frank Caruso*

ARC Centre of Excellence in Convergent Bio-Nano Science and Technology, and the Department of Chemical and Biomolecular Engineering, The University of Melbourne, Parkville, Victoria 3010, Australia.

ABSTRACT

This is an exciting time for the field of bio-nano science; enormous progress has been made in recent years, especially in academic research, and materials developed and studied in this area are poised to make a substantial impact in real-world applications. Herein, we discuss ways to leverage the strengths of the field, current limitations, and valuable lessons learned from neighboring fields that can be adopted to accelerate scientific discovery and translational research in bio-nano science. We identify and discuss five interconnected topics: (i) the advantages of cumulative research; (ii) the necessity of aligning projects with research priorities; (iii) the value of transparent science; (iv) the opportunities presented by “dark data”; and (v) the importance of establishing bio-nano standards.
INTRODUCTION

The field of bio-nano science is as exciting as it is diverse. Thought-provoking developments are reported almost daily from a growing number of research teams.\textsuperscript{1–5} As an area that both fascinates and invokes hope for new medical treatments, it generally enjoys both the public’s interest and relatively high levels of funding. However, there is growing concern that despite the academic impact of this field, translation into real-world applications and treatments has been lacking. For example, in the field of cancer nanomedicine the current approach has been called “broken”, capable of “curing mice” but unlikely to “translate to patient care”.\textsuperscript{6} Although seen as controversial statements by some,\textsuperscript{7} they highlight an existing frustration and underline that translational work remains challenging.\textsuperscript{8–10} Moving forward we should ask: will “research as usual” get us to where we want to be? Or are there aspects of our research that we can alter to accelerate this pace and increase the impact of materials in bio-nano science and beyond?

Herein, we discuss our views on current limitations in the field of bio-nano science, focusing on material-biological interactions, and directions we believe would accelerate the pace of scientific discovery and translational work. An important consideration is that the field is not unique in facing the challenges we raise. Many other fields, both related and unrelated, have experienced similar issues that can be informative and useful to consider. An important theme here has been labeled “convergence”: the integration of the physical sciences, engineering, and biomedicine into a unified whole, offering potentially revolutionary new possibilities.\textsuperscript{11} Within this framework there are many topics that are of interest and we will focus our discussion on five interconnected areas (Figure 1).
Figure 1. Accelerating scientific discovery and translational research at the intersection of materials science, engineering and biomedicine through convergent science.

CUMULATIVE RESEARCH

Opening the latest issue of your favorite journal will almost certainly yield several reports of impressively intricate solutions to complex problems. For example, it is not uncommon to read about nanoparticles designed to detect, image, diagnose, and treat tumors, all in a single nanoparticle system. Well-performed studies of this kind are important and deserve to be published in high-exposure journals, as they help push the boundaries of our understanding, both in what type of materials we can engineer and how they interact with biological environments. However, the addition of new functionalities to a bio- or nanomaterial make its biological behavior more convoluted, and highly complex multifunctional materials face virtually insurmountable regulatory hurdles. Therefore, although these materials are of great interest academically and new strategies that facilitate regulatory decision-making are being developed, they remain the most complicated to pursue for translational purposes. As an example, multifunctional long-circulating nanoparticles that detect, image, diagnose and treat cancer metastasis (the cause of much cancer-related morbidity and mortality) would be of great clinical benefit, but explicit pursuit of this goal alone is unlikely to quickly translate into improved patient outcomes.

Cumulative research is about building from what we know and is crucial for translational research. Academic research prioritizes projects at the very boundary of our knowledge, at times far from established or well-known facts. The unexpected and unknown is sought, and greatly rewarded if found. For translational research, it is instead important to minimize unknowns and proceed in a step-wise manner. A research community that aims to have substantial impact inside
and outside academia requires both, which is challenging but likely to provide significant rewards. For example, in the field of drug delivery, considerable effort has been spent on the development of injectable long-circulating nanomaterials and particles, a highly complex challenge where many biological barriers need to be considered and negotiated. However, there are considerable opportunities in developing “non-circulating” nanomaterials for drug delivery. Examples include vaccination and immunomodulation;\textsuperscript{16} as reservoirs to treat metabolic disorders (e.g., insulin reservoirs in diabetes)\textsuperscript{17} and central nervous system disorders,\textsuperscript{18} and to treat localized diseases (e.g., lung-related diseases,\textsuperscript{19} ear-related diseases,\textsuperscript{20–23} and localized infections among many others).\textsuperscript{1,24,25} By starting with these types of focused therapeutic areas where circulating particles are not necessarily needed, the issues and biological barriers associated with circulating particles (e.g., to understand and control bio-nano interactions with tissues and organs that are in contact with the circulatory system—an almost overwhelming challenge)\textsuperscript{26} can be avoided. This would lead to more translation of designed materials into the clinic, increasing our understanding of how they work in real patients, which in turn can help guide and accelerate the development of new materials. By building a solid foundation using the knowledge gained from how different types of materials behave in these “focused therapies”, we will be better equipped to tackle more complex challenges consisting of many biological barriers. This would accelerate the translation and increase the impact of materials in both these focused areas and beyond.

**TRANSLATIONAL ALIGNMENT**

While blue-sky research is of great value, a significant body of work in the field of bio-nano science has the stated goal of clinical translation. Vital in accomplishing this goal is that we pursue ideas that have the potential to provide substantial clinical benefit.\textsuperscript{27} For instance,
developing treatments for diseases that already have excellent treatment options, or have low morbidity and minor effects, is unlikely to provide substantial clinical benefit. As another example, consider organ-level targeting, a topic that has attracted significant interest in the development of materials for drug delivery. For some organs, localized drug delivery is simple for a clinician: direct injection into the organ of interest. Regular contact with clinicians and others in the medical field is essential to ensure that the research reflects top patient priorities. A recent article in *PLOS Medicine* provides guidelines for ensuring that translational and clinical research is useful.

Another vital component of impactful research is ensuring that our models—whether *in vitro*, *in vivo*, or *in silico*—are good approximations of the systems we are investigating. To quote George E. P. Box, “all models are wrong, but some are useful.” Unfortunately, this bar is not always cleared. *In vitro*, adherent cell lines are the norm for investigating cellular response, but it is important to remember that these only model aspects of an organism’s biology and that there can be issues with inferring more general conclusions from such studies. For example, HeLa, one of the most studied cell lines, show strikingly aberrant characteristics compared to healthy cells. Working with more complex *in vitro* systems is often desirable—for instance, primary cell culture, three-dimensional cell culture, or microfluidic cell culture systems, but the vast majority of these techniques still require considerable training and specialist equipment. Additionally, substantial questions remain about how to translate what is learned in animal models—the “*in vitro-in vivo* gap” and the “*in vitro impasse*”. There has been a trend to move more quickly into *in vivo* studies to address these concerns, but there are also many open questions about the applicability of animal models to human diseases. We believe that the *in vitro-in vivo* gap is not insurmountable, but that it requires concentrated research effort—in
developing advanced techniques, improving their adoption, and proving their worth. Given the ease of performing *in vitro* experimentation relative to *in vivo* studies, we believe this effort is worth expending, and are encouraged by the efforts of the organ-on-a-chip community\textsuperscript{39,40} for example.

Finally, we should consider whether the metrics we use are appropriate. As is the case for models, most of the metrics in use in the field of bio-nano interactions have been inherited from other fields—for example, drug half-life from pharmacokinetics, or percent cell association or activation from cell biology. A metric that works in one domain may be inappropriate in another. Examples of this include the debate around the appropriate dose metric for nanoparticles.\textsuperscript{41–43} We believe that the development of new metrics, and re-examination of those we use, are essential for high-quality, impactful work. The blind optimization of metrics at the expense of what they are meant to represent, for instance “*p*-hacking” (changing the analysis method or the data eligibility requirements until significant results are found)\textsuperscript{44–47} and misuse of impact factor (e.g., to judge individual papers or scientists),\textsuperscript{48,49} is also to be avoided. Appropriate use of metrics facilitates communication and collaboration by providing a common vocabulary, evaluating performance, and simplifying complexity.

**TRANSPARENT SCIENCE**

Research reliability and reproducibility is an important and hotly debated topic.\textsuperscript{50,51} In a recent survey of active researchers across different disciplines and geographical areas, two-thirds said that current levels of reproducibility are a major concern,\textsuperscript{52,53} a sentiment that is shared by major funding bodies such as the US National Institutes of Health.\textsuperscript{54} Recent failures to confirm findings published in high-profile journals by organizations such as the Open Science Collaboration\textsuperscript{55} and large biotechnology companies\textsuperscript{56} further emphasizes the challenges associated with reliable and
reproducible research. The issue is further complicated by factors such as contextual sensitivity and so-called “hidden moderators” that are difficult to elucidate and account for, and the fact that there is widespread confusion of relevant terms, such as reproducibility, replicability, reliability, robustness, and generalizability. This concern and frustration is shared by related industries and commercial sectors, as exemplified by a recent proposal from a high-profile industry leader that academia and industry should form “incentive-based” agreements involving “money-back guarantees” if results cannot be reproduced, a proposal that quickly received strong criticism. We believe embracing a transparent research culture will form an important part of the solution for improving research reliability and reproducibility.

The core concept of transparency in reported research is to make it as easy as possible to understand exactly how the research was performed, how the data were analyzed, and how the conclusions were reached. In increasingly interconnected and convergent research environments, this is a multifaceted problem involving many stakeholders: researchers in your field, related fields, funding agencies, regulators and governing bodies, industries and commercial interests, the general public, and many others affected directly or indirectly by impactful translational research.

Increasingly open and transparent research environments accelerate discovery, innovation and translational work. Central to advancing this process is how we communicate research. Last year the Transparency and Openness Promotion (TOP) Committee published guidelines to reward and incentivize transparency and to move scientific communication toward greater openness. Since then over 700 journals and 60 organizations have expressed their support. The TOP guidelines provide eight standards—including areas such as data transparency, analytical method (code) transparency, and research materials transparency—with multiple
levels leading toward increasing openness and transparency. The idea is to increase awareness and help guide journals and researchers toward best practices.

The TOP guidelines facilitate understanding, replication, and application of published research, but their benefits go beyond this. They help establish a framework where not only the findings of a study are shared, but also the underlying data and methods. Imagine, for example, if researchers routinely recorded (e.g., using GoPro cameras) and made videos accessible to complement their publications—viewing and understanding exactly how an experiment or procedure was performed would be much easier. Similar thinking has already led to the creation of web portals and data systems for transparent sharing and open usage of genomic and gene expression data\textsuperscript{70–72} and mass spectrometry data,\textsuperscript{73} and a movement toward “open medicine” and the sharing of clinical trial and outbreak data,\textsuperscript{74–83} and more exotically, the creation of a public xenograft repository.\textsuperscript{84} There are even examples of institutes adopting open science policies across the board to ensure transparency and reproducibility, and to increase the pace of discovery and boost innovation.\textsuperscript{85} An important part of the examples listed above is increased transparency and accessibility of data, which is crucial for improving reproducibility.\textsuperscript{86} In addition to this, there is the tragic waste of data generated through well-performed studies that never see the light of day, or so-called “dark data”.

**SHINING A LIGHT ON DARK DATA**

Research is exploration. We may start at the well-known area \( A \) and try to find our way deep into uncharted territory \( B \), where we hope to discover an increased understanding of a system or the solution to a problem. Studies can last anywhere from weeks to many years during which there are numerous obstacles to be overcome, and the research direction can completely change. Typically, at the end of the journey large areas of uncharted territory have been traversed before
“finding the destination”. However, when we communicate research findings, they are typically focused on “the shortest way from $A$ to $B$” and other knowledge gathered—the dead ends encountered and challenges conquered—exist only in the researchers’ mind, or in laboratory notebooks gathering dust.

We use the term “dark data” to refer to the information that researchers collect from well-designed and well-executed experiments that are never disseminated to the research community (e.g., unsuccessful and unfinished studies, or studies yielding negative results). An example is a recent high-profile article, which resulted from the use of previously unreported failed or unsuccessful syntheses (“dark reactions”) collected from archived laboratory notebooks that, combined with machine learning, was used for materials discovery.\textsuperscript{87} This example demonstrates that data on its own may not be interesting, but can be combined and aggregated to provide completely new insights and directions. Another example of dark data is the data generated by companies that can be difficult to access, but may be very valuable for guiding research and policy. An example of this is the controversy during the recent influenza pandemics surrounding antiviral drugs and their associated clinical trials.\textsuperscript{88–90}

Each piece of data generated from well-designed studies, even if they turn out uninteresting for the research question of that study, may add to the bigger picture of a research area. Reporting of dark data also provides guidance to other researchers by preventing or minimizing effort in areas already explored and shown not to work. It enables large and important patterns (that are invisible when studying small sets of data) to be discovered and can thus guide developments and provide valuable insight. This is especially true when working in complex fields such as the bio-nano area, with the combined complexity of the physical sciences,
engineering, and biomedicine. But improved dark data reporting will only happen if we create a culture and develop ways to encourage and reward the work required to disseminate such data.

Based on our experiences in working in and with different laboratories, fields, and countries, we believe that more than half of the total amount of data generated through academic research may be dark data. So for all data available today, there exists at least an equal amount of dark data. (This can be compared with the claim that more than 85% of research investment in the biomedical field is wasted.\textsuperscript{91,92}) If we expand the concept of dark data to consider data that is not easily accessible (e.g., in an open database or repository), the fraction of dark data would explode to include the vast majority of all data. As an example, consider a representative bio-nano study using flow cytometry to investigate the interactions between cells and a new type of particle. At least $10^5$ data points will be collected \[(\text{two samples and control}) \times 5 \ (\text{time points}) \times 3 \ (\text{triplicate}) \times 10 \ 000 \ (\text{flow cytometry}) = 450 \ 000 \ \text{data points}\]. This data set would then, typically, be reported using average values and values describing the spread showing how the cells responded over time to the particles. So the almost half a million data points would be reported using a handful of values. There is nothing inherently wrong with this, as it facilitates understanding of the data and can be fully appropriate for the research question being investigated. However, if the raw data is not made accessible, all the additional information that could be extracted is lost. If alternative methods to interpret these data arise—and history shows that we should expect them to, just consider the emergence of “big data” methods\textsuperscript{93,94}—these data will be unavailable for use. Raw data can, and should, be expected to have use for many other research questions, both now and in the future.

If we unlocked the potential of dark data, many new types of studies could be performed. To illustrate this, consider the field of medical research. Medicine is a highly complex and active
field where decisions (based on research) directly impact the lives of patients. To help advise clinicians and inform medical decisions, meta-analysis (literature data-mining/knowledge-extraction) has become a powerful and widespread tool.\textsuperscript{95–97} Meta-analysis is a procedure for combining the knowledge from many studies to give an unbiased and more complete understanding of a research topic, and is central in evidence-based medicine.\textsuperscript{98,99} The studies providing the source data do not have to have focused on the same topic as the meta-analysis, they only need to contain relevant data reported in an accessible way. This relevance might not even have been known at the time of the original studies, but only later in light of further evidence.

Two recent high profile publications illustrate the power of using meta-analysis techniques in the bio-nano field.\textsuperscript{6,100} In one of them 1741 cell viability-related data samples were obtained from 307 publications to analyze cellular toxicity of quantum dots.\textsuperscript{100} Over 1100 papers were initially identified but only 307 satisfied all selection criteria, and from these relevant data were manually examined and extracted. In the other, results from over one hundred publications were combined to analyze the efficiency of nanoparticle delivery to tumors.\textsuperscript{6} Over 200 publications were identified as potentially of interest and of these 117 publications fulfilled all the selection criteria. In many cases the authors of the original publications had to be contacted to gather all the data needed for the analysis, as it was not accessible through any other way. While the results produced are of great interest, one also realizes the herculean effort that must have been required to manually process hundreds or thousands of publications, extract and sort all the data, and contacting original authors to gather the additional information needed. Imagine instead that these types of data sets were made routinely available within the bio-nano field. Meta-analysis studies could then be regularly performed across the whole area, informing
research and directions by providing perspectives and overviews that today are somewhere between difficult and impossible to obtain.

The solution to dark data requires changes in both research methodology and culture. Dissemination of dark data is a classic, public good conundrum in which the community as a whole (greatly) benefits in the long run, but individual researchers must perform extra work in the short term. Part of the solution will be to ensure that this effort is acknowledged and rewarded so that contributions (e.g., adding raw data to repositories) are incentivized. Fortunately, substantial progress has been made in other fields that are also moving toward this goal. For instance, in the field of molecular biology there are numerous databases and repositories available, many of which have specific guidelines for how to acknowledge and cite data.\textsuperscript{101} Examples include the Protein Data Bank\textsuperscript{102} and the Gene Expression Omnibus.\textsuperscript{103} An example from the chemical sciences is the Cambridge Structural Database\textsuperscript{104} containing over 800 000 small molecule crystal structures. In addition to this, it is important to acknowledge the cultural change that will also be needed. Here, inspiration can be found in medical science and biomedical research, where recently there has been a focus on increasing value and reducing waste.\textsuperscript{92,105} One lesson is that if a study is well-planned and executed, the norm should be that the data is always disseminated and made accessible. This means that unless there are specific reasons not to disseminate data (e.g., privacy issues when working with patient-derived samples) we should strive toward making all research output available to the research community through means such as databases and repositories, preprint servers,\textsuperscript{106,107} and journal publications. For increased data dissemination to have greater impact on scientific discovery and translational work in bio-nano science, it is vital to agree on and establish research and reporting standards.
BIO-NANO STANDARDS

Undermining our desires for cumulative research, data dissemination, and systematic investigation is the lack of research standards for reporting on interactions between materials and biological systems. All of these issues are made significantly more difficult by a lack of consensus on what aspects of an experiment to report, and standardized data formats for doing so. As discussed earlier, even the reproduction of a published experiment can be a significant challenge, especially given the specialized equipment and experience that material fabrication and biological experimentation require.

While the literature contains a multitude of manuscripts of the form “The effect of $X$ on biological response”, typically $X$ is one variable for one particular material. Given the complex interplay of biological systems with materials, it is unlikely that the conclusions suggested by titles of this form are the full story. For example, consider that the field does not even have comparison metrics for disparate nanomaterial systems, short of highly downstream biological responses like cytotoxicity. There is no established standardized way of comparing, say, the targeting avidity of two designed drug delivery systems to a cell type, or their drug release capabilities. Instead, a variety of ad-hoc, usually qualitative methodologies are used, and materials are typically only compared to similar controls within the same study. After decades of work, basic questions remain unanswered—for instance, what are the tradeoffs between using material $A$ or $B$ for the development of a new bio- or nanomaterial? Which is better? Without standards—defining what “better” means—we cannot hope to readily address these questions.

Research standardization has four components: (i) a common language (an ontology), (ii) data formats for exchanging information, (iii) agreements on which information is relevant, and (iv) curated databases to store previously reported data in the standardized format. For the field
of nanomaterials, component (i) is being addressed through efforts such as the Nanoparticle Ontology,\textsuperscript{108} while (ii) is served by formats such as ISA-TAB-Nano.\textsuperscript{109} There are several databases for reporting bio-nano research, including the nanomaterial registry,\textsuperscript{110} caNanoLab,\textsuperscript{111} and Nano\textsuperscript{112} which begin to address (iv). In our view, (iii) needs additional attention. Minimal information about nanomaterials (MIAN)\textsuperscript{110} is an elegant solution for capturing information about nanomaterial characterization. However, the very flexibility of the frameworks and formats we discuss here, which are capable of including almost any information about a material, makes their use overwhelming. Furthermore, research standards need to not only address material characterization, but also information about the biological experiments performed. Finally, there needs to be (evolving) consensus about which characteristics of a material are important and reported. For instance, while nanoparticle size is undoubtedly an important determinant of cellular response, color, in itself, is almost certainly unimportant (though color is vital in other areas of material development).

Research and reporting standards mean little unless the community—both researchers and journals—agree to them. Journals are the gatekeepers of research standardization in a field. The adoption of checklist requirements and reporting practices, e.g., for biomedical research,\textsuperscript{113–115} for solar cells,\textsuperscript{116,117} and for a range of other materials and devices,\textsuperscript{118} is an encouraging step toward journal-expedited research standardization. This maturation parallels other areas at the forefront of biological research that have benefited from the establishment of minimum information standards. Examples include the MIQE guidelines for quantitative PCR,\textsuperscript{119} the MIAME standard for microarray data,\textsuperscript{120} and the MIRIAM standard for biochemical models,\textsuperscript{121} all of which have seen wide acceptance by their respective communities, and led to more productive research. The field continues to produce a dizzying number of new materials and
synthesis methodologies. We need to find a way to compare and determine which of these materials work, or do not work. Standardization, which is important for more mature materials (e.g., biomaterials in clinical use, especially when using a broad definition\textsuperscript{122}), is associated with obtaining and maintaining regulatory approval, which poses its own set of challenges.\textsuperscript{5,13,123–126}

**SUMMARY AND CONCLUSIONS**

The intersection of materials science and biomedicine has great potential, but to realize that potential, we need to re-examine our current approach. The themes of cumulative research, translational alignment, transparent science, illuminating dark data, and establishing research and reporting standards will accelerate progress. While the challenges discussed herein are considerable, we are fortunate that they are neither new, nor is our field alone in facing them. Through collaboration and importation of best practices from other fields, we can tackle these difficulties. The needs we outline reflect the growth of the bio-nano field from often exploratory, ad-hoc, and qualitative experimentation toward including more systematic, quantitative investigations.

The interdisciplinary nature of bio-nano science, combined with the parallel challenges faced by other fields, has led us to highlight convergence as an underlying theme and exciting solution. The union of multiple, disparate fields, has tremendous potential—much more than these fields working in isolation. There have been high-profile calls to increase convergent research\textsuperscript{11} and to re-imagine current fields along convergent lines.\textsuperscript{127} We strongly advocate for more convergent research, and are encouraged by our increasing difficulty in classifying modern projects as part of traditional disciplines. Another example is the increasing prevalence of “hybrid training”, for example across disciplines or by combining clinical and basic sciences.\textsuperscript{128,129} However, there are challenges to pursuing convergent research. Foremost and
most obvious is the difficulty of funding interdisciplinary research,\textsuperscript{130} which we need to move as a community to encourage. Additionally, there is inherent risk in expanding into new areas—misunderstandings can lead to slower output at best and retractions at worst. Similar to our discussion in cumulative research, our approach and advice is to accumulate experience in one area and expand into others through collaboration. Aligning research culture, incentives, and funding toward increasingly convergent science will require adjustments in both how we think about science and how we conduct research, but the potential benefits are enormous: both for increasing fundamental understanding of bio-nano interactions, and for translating material-based developments into real-world applications and improved clinical outcomes.

**AUTHOR INFORMATION**

**Corresponding Author**

*E-mail: fcaruso@unimelb.edu.au*

**ACKNOWLEDGMENTS**

This research was conducted and funded by the Australian Research Council (ARC) Centre of Excellence in Convergent Bio-Nano Science and Technology (Project CE140100036), and supported by the ARC under the Australian Laureate Fellowship scheme (FL120100030). We thank Alison E. Burke and Cassio Lynm for assistance with preparing the figure, and Edmund J. Crampin, Andrea O’Connor and Stephen J. Kent for valuable discussions.
REFERENCES
(30) Landry, J. J. M.; Pyl, P. T.; Rausch, T.; Zichner, T.; Tekkedil, M. M.; Stütz, A. M.; Jauch,


(38) van der Worp, H. B.; Sandercock, P. A. G. *BMJ* 2012, 345, e7837.


(64) Edwards, A. Nature 2016, 533, S70.
(89) Loder, E.; Tovey, D.; Godlee, F. *BMJ* 2014, 348, g2630.  
(90) Jack, A. *BMJ* 2014, 348, g2524.  
(91) Chalmers, I.; Glasziou, P. *Lancet* 2009, 374, 86.
(94) Pržulj, N.; Malod-Dognin, N. Science 2016, 353, 123.
ToC graphic