



The cultivation of supply side data science in medical imaging: an opportunity to define the future of global health

Adam Kesner¹

Published online: 23 October 2021

© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2021

The field of nuclear medicine has long been defined with innovative technology seated at the cross roads of scientific and medical specialties. We have delivered inspiring on the promise of harnessing nuclear phenomena for the betterment of human health and established relevance broadly throughout modern medicine. Our joint missions of establishing best practices and supporting creative innovation remain as relevant as ever.

Many forefronts of the nuclear medicine field are advancing simultaneously—we are seeing exciting innovations in areas of pharmaceutical development [1], expansion into oncology [2] and nuclear medicine aided personalized medicine [3], dose enhancement strategies [4], and pharmacogenomics [5] and pretargeting techniques [6], to name a few. However, the innovations we produce in computing and instrumentation, a cornerstone of our field, are seemingly not living up to their potential—the technologies we build are stalling in their translation into clinical settings. EJNMMI recently published a two-article controversy series around the topic of embracing new positron emission tomography (PET) technologies. The series considers opportunities for affecting positive change in PET conjugated with the obstacles for translating innovation; provocative sentiments are presented at both ends of the discussion [7, 8]. The topic is timely as instrumentation has long been a pillar of our field and could continue to be, but this controversy highlights a quintessential antagonism and challenges us with a principal question pertaining to innovation: what is relevant? To many entering the field in my generation and later, the compass we have inherited appears antiquated and aligned with a synergy built to accommodate successes of the past.

We are witnessing many of the technologies, and domains we endeavor to master remain persistently academic at a time when remarkable technological innovations are transforming the world around us.

The two aforementioned papers are presented as opposing, but they congruently articulate a reality of our field: incremental technological innovations are having a limited impact. Perhaps through informed discussion and review of these technologies and their potential, we may be able to better appreciate them and thereby push them past obstacles and into clinical use. But the question of why we are not witnessing academic efforts fomenting change is bigger and more complex. As recognized in the articles, it is not the case that these technologies are not useful or capable of improving patient care. Rather, they are encumbered by hurdles, and ultimately are not having the transformational influence we have witnessed from innovations in years past. Nor are they producing peer leaders to demonstrate pathways demarcating success for the next generation. By metaphor, if we are having trouble finding the goal posts, as the title of one of the controversy articles suggests, it is worth the consideration that we may be playing on a wrong portion of the pitch.

From one perspective, the field of medical diagnostic imaging is maturing, supported with increasingly well-established models of technology, care delivery, and professional roles. By extension, we may be trying to innovate within a legacy infrastructure best suited for twentieth-century innovation, which has not evolved to match the rapidly shifting resources and changing environments in which the clinical applications of our field are practiced. Our efforts to seek transformation, while protecting our legacy models from disruption, may be antithetical. More specifically, access to resources, pathways for translation, and feedback systems for cultivating concepts of relevance are unduly constrained. This may hinder our ability to affect change and to provide meaningful leadership in the coming generations.

Traditional success with innovation entails matching innovation with innovation ecosystems [9], but the ecosystem itself is not immutable; it is an evolving configuration

This article is part of the Topical Collection on Technology

✉ Adam Kesner
kesnera@mskcc.org

¹ Department of Medical Physics, Memorial Sloan Kettering Cancer Center, 1250 First Avenue, Room S-1119E (Box 84), New York, NY 10065, USA

of existing resources that include not just technologies but also advances in knowledge, skills, concepts, organizations, and other social and institutional contexts [10]. We need to recognize the emergence of a significant new dimension of our field—a maturing landscape of data science. Reflection on its *de facto* versus ideal integration in the field, along with newly emerging ideas of open data, is now within the purview of academia and the greater community. That is the topic of this letter. Specifically, there exists today an unprecedented opportunity for establishing a concept/role for data that is unique from the hardware that creates it. The pivot could bolster all elements of research, discovery, translation, and inclusion and ultimately lay a foundation for meaningful growth in the coming decades [11].

An anecdotal experience with imaging innovation

My own experience in imaging innovation stems from my graduate work with data-driven PET motion correction in the late 2000s. The technology, in brief, addressed the issue of PET image blur caused by patient motion during acquisition using fully software-based processing techniques, in contrast to hardware-based solutions; the technology is practical, standardizable, and scalable. The imaging field has more advanced data processing forefronts now, but this experience offers insight from one of our fields' first forays into data science translation, and it exemplifies the promise of data science: the concept of doing more with the data you have.

In my efforts to promote translation of this originally academic solution to clinical PET, I found that the obstacles I encountered had less to do with the substance of the new technology and more to do with a limited infrastructure for data science within our field. This is exemplified with some observations: first, for almost two decades, the subfield of PET motion correction has been supported by a synergy of vendors selling hardware gating systems, academics investigating and publishing with them, and hospitals purchasing them. Yet the principal feature of actual value has always been missing. The hardware gating technology is rarely used in diagnostic PET outside of research, and the limitation of patient motion persists. It thus appears that our field can sustain a technology that does not actually provide benefit. Second, clinical and translational research benefits from larger sample sizes/data sets. Despite our hospital environment filled with ample image data, and use of fully automated processing techniques capable of incorporating large numbers (thousands) of scans, we routinely performed research with very small populations, due to technical and legal constraints. Third, as my colleagues and I competed with ideas, I found that academic success—journal publications, awards,

grants—did not seem to correlate with relevance. In fact, my understanding of relevance differed significantly from many colleagues and vendors. I welcomed this diversity of thought as an opportunity to gauge my aptitude for the science, but I found academia to be bound by inertia and slow to process creative ideas on their merit. Fourth, when I could not find support in academia for what I saw as a solution ready to be ubiquitously deployed to solve a present problem, nor support working with the vendors towards building a clinical product (circa 2012), I set out to bring the product to market myself. I found that there is no pathway for third-party innovation around listmode data (raw data)-driven solutions because the necessary raw data formats were proprietary.

It is quite common that ideas and software from graduate students follow the path of publication then obscurity, an accepted inefficiency that squanders our greatest creative resource and/or filters progression of creative potential altogether. In the case of data-driven motion correction, recently PET vendors have brought their versions of software solutions to the market [7]. And indeed, it is worth noting that the industry is shifting significantly towards embracing data science and providing greater data access tools to customers.

The status quo for instrumentation innovation

The pathway of new ideas to clinical impact spans the community, across academia, industry, and clinical partners. Of significant consequence is the fact that the current landscape of imaging instrumentation is dominated by an ever decreasing number of large companies. The ecosystem has merit; these companies play a leading role in the field's cohesion and serve as gate keepers in promoting commercially viable technologies. Such an arrangement helps ensure competitive economic stability in our field. However, this paradigm also has significant inefficiencies in promoting alternative, creative, inventive, reinventive, often data-driven innovations that hinder the solutions from translating to the clinic. Solutions, based on reimagined data use, require raw data, but raw data remains under the auspices of equipment vendors who do not often have incentives to choose disruption over established technologies.

The dominating role of large companies in the translational pathways of innovation has many implications that shape our field, beginning with the very concept of relevance, which becomes defined/confined by the construct of what is possible with respect to the current industrial translation pathway paradigm. In innovation theory, there are several models in which technology is invented and disseminated [12]. Innovation through models of exclusivity, increased costs, and increased profitability—and presumably trickle-down diffusion—are currently dominant

in our system over alternative models, for example, those that expand access through reduced cost via unconventional technologies or creative redistribution of resources in our field. This predisposition should be examined because such opportunities are the promise of the emerging data science arena. Another implication of the status quo is that it leaves academia and clinicians to groom our peer expectations and standards with a vendor conforming view of what is relevant/possible. This disassociates us from opportunities born in the changing environment and hinders our ability to lead across generations. It also raises confusion of roles—cohesion and success become intertwined with supporting vendor products, possibly in lieu of due criticism or recognition of external feedback.

The current standard of framing raw data as an extension of hardware, under the auspices of equipment vendors, provides the elements of compartmentalization and efficiency to our field. However, let us consider that this understanding of data is a legacy from when data processing capacities were more limited and simplistic, and accordingly is due to be revisited. Importantly, leaving the entirety of hardware and software development priorities to the hardware vendors incentivizes the development of hardware solutions within established business models, further solidifying a hardware-centric identity of the field, with data science applications in medical imaging relegated to secondary, supportive, and vendor-specific peripheral products.

Presently, diagnostic scanners output images in DICOM format, with largely open accessibility. These images form the basis of clinical management and are used widely for research and third-party commercial products. The raw data that is used to create the images, which fully articulate the information captured in a patient scan, is restricted/proprietary. The academic community can be granted access to raw data with the help of vendors in select situations through a legal framework of research agreements. This system has benefits in that it provides support for academic research and development while also providing protection for vendors to invest in production and commercialization. It also ensures that non-regulatory approved products are not prematurely diffused into clinical use. However, the system precludes alternative innovation pathways. Researchers are somewhat obliged to stay in a vendor's good graces, arguably forfeiting intellectual independence, and intellectual property is forgone and/or made obsolete because third-party spinoff companies are not possible. Consequently, raw data-driven innovations only happen when they are selected by a vendor, developed to the vendor's specifications, and made available only in a vendor-specific format. In the example of data-driven motion correction for PET technology, patient benefiting solutions could have been brought to market a decade earlier, and in a vendor agnostic format, but were not because hardware systems were already selling. The problem

is reiterated when we examine a variety of software innovation, e.g., the active field of image reconstruction research—vendors have costs and minimal incentives to translate alternative reconstruction methods into their systems. Another implication of this industry led landscape is that vendors are increasingly moving towards offering vendor-specific suites of tools, compounding layers of vendor specificity into their products, and academic scientists are increasingly engaged in accepting a role of evaluating vendor products, rather than understanding/mastering/innovating new solutions. Looking forward, it is unclear what the future of this arrangement holds: either we need to accept that patients receive different care from different vendor supported systems or one vendor succeeds in dominating the imaging market, and we lose the engine of competitive development.

There is an important issue of image harmonization at many forefronts of translational technologies in PET, including radiomics, artificial intelligence (AI) segmentation, big data analysis, computer aided diagnosis, and dosimetry, which all stand to benefit from a reduction of variable biases [13, 14]. A commonality across all these subfields, as articulated in review articles, is the need to standardize and/or harmonize image data, thereby amplifying their capacity to inform. This is particularly true in large population research studies. The results of a recent study, working with a similar modality, single-photon emission computed tomography (SPECT), has shown that the largest source of variability across image collections from multiple vendors was attributed to differences in vendor image reconstruction [15]; it is reasonable to presume that the same is true for PET, a more inherently quantitative modality. Dedicated careers and the future of the field's subfields are dependent on the ability to achieve optimal harmonization. The biggest opportunity to normalize data across centers and sizably improve image harmonization is to establish reliable cross-vendor reconstruction platforms that support optional standardized advanced reconstruction. Notably, this observation is omitted from harmonization related literature. We might presume this is because this solution is commonly accepted as impossible and that the siloed formatting of vendor specific data is an insurmountable reality. However, it is technically possible.

The changing environment

We are in unprecedented times. Technology, technological capacities, and individual expertise are moving faster than the speed of careers. Data and communication traverse the globe at the speed of light, and market forces are global, as are communities.

The concept of “adapting to change” is well recognized in biological sciences, social sciences, and business.

Although the association was never made by Darwin himself, his seminal model of natural selection is often applied to social/industrial paradigms, because the principal concept is observed across fields: those with favorable adaptations in changing environments gain advantages and are best suited to survive and thrive. Opportunities to adapt and define our field with relevant innovations at our forefront include ensuring access to changing resources and, in contrast to the twentieth century, making space/mechanisms for the integration of change.

A notable shift in the environment of nuclear medicine has occurred, albeit slowly, with respect to the role of computing. The computational landscape has transformed from a supporting role to something much more significant. Nuclear medicine imaging is defined by all-digital modalities whose invention and capacity has always straddled the forefront of computing and data storage capacities. Decades of computational innovation described by Moore's Law of exponential processor growth (higher circuit densities at reduced cost) have created consequential implication [16]. Where the twentieth century witnessed a transition from the industrial age to the information age, we are now seeing a transition from the information age to a new era of exponential digital innovation, perhaps best epitomized by the maturing landscape of AI architecture that is driven with data [17]. The traditional role/value of data has been in storing information for reference. Now, in neural network paradigms, information within data is supplanting the task of analytical programming and supporting software generation, essentially integrated into the design of modern computing solutions. These same neural networks allow us to process massive amounts of data at micro detail levels. The limitations of our ability to create useful solutions in modern innovation frameworks quickly centers on data, and specifically its access, quality and abundance. Data has become a resource of interest, a new currency, and the most likely bottleneck of future evolution [18].

In the field of nuclear medicine, we collect a lot of data, more so than other medical specialties. Our data is distinctive with unique signal (from injected radioisotopes), logistically complex acquisitions (patients routed through hospitals), and collected with hyper-sensitive machines (PET/SPECT/probes). We know that it contains still untapped information (representations, correlations, interpretations, and insights), which can be extracted with new tools that expand our capacity to see it and mine it. The data is economically and socially relevant (patient health). A reorientation of understanding data as a new axis of our field should readily extend from these observations. Further, we do not need to know exactly how we will use it in the future to recognize and cultivate its value in preparation for the coming generations.

Over the last decade, we are seeing a growing concept of open science and open data and its relevance in the sciences [17, 19, 20]. While acknowledging the value of data, it should be recognized that the field would benefit from discussions around its stewardship. How do we retain its fidelity? How do we handle its vendor specificity? Importantly, who "owns" medical data? Patients, hospitals, vendors, health systems, or grant agencies? Where does privacy begin and end with respect to population data? What data access do we ideally want in the twenty-first century? [21] What are the obstacles in achieving it? What regulatory oversight is needed and/or appropriate? All these questions require forethought, and leadership.

Aside from changes in computing and data, an additional monumental shift has occurred with respect to globalization and global industrialization. The boundaries that once stratified our field across geography and populations have been undergoing monumental changes in the last few decades shifting the healthcare landscape and integrating billions of people through connectivity and technology. This expansion opens new markets and enhances our pool of professional talent, creative researchers, and entrepreneurs. Accordingly, it presents an opportunity to evolve our identity and care delivery models towards global, rather than local relevance, and to benefit from these new human and cross-cultural resources. Like the aforementioned opportunities of defining standards in data, normalizing a global, inclusive community for twenty-first century innovation could be an invaluable pathway for innovating within our field and improving care.

Also worth noting are the different skillsets we should anticipate that future generations will bring to our field. They will be, and are already, digital natives, computer literate virtually their whole lives. They are used to navigating, innovating, expressing themselves, failing, and succeeding via modern computing infrastructures. Students are trained today in AI with open architecture systems—open access training modules, programming environments, and neural networks. Their experience in, and expectations of, global communities, idea transmission, and reputation building have matured in systems built within industries that have already undergone tremendous disruption and transformation. Beyond traditional knowledge, they will need an era-accorded future-leaning infrastructure to be relevant.

Opportunities for evolution

Much of our collective success across academia, industry, and care providers has come from a positive collaboration, and it is this established productive synergy that continues to fuel a cohesive global field. We are partners; with history, with many shared values and with shared stakes in a

relevant future. We have the opportunity to build a vision for the future of our field.

Synergy around a product or technology is good, but synergy around an infrastructure for meaningful innovation is better. Every transformative innovation starts off as an idea with questionable relevance and requires an ecosystem that propels promising opportunities [10]. A healthy eco-system for data-driven innovation should have: encouragement for new ideas, access to resources (in our case raw data), access to translational pathways, and critical debate [14]. Accordingly, we should now consider that raw data is a resource for innovation in the twenty-first century. We should proactively endeavor to guide the field such that data is accessible for creative innovators of the coming generations. Congruently, and within the same effort, we can make space for alternative commercial pathways for third-party startup companies so that data-driven disruption can occur if/when/where appropriate.

The importance of enabling free market competition in the innovation space is fundamental to our ability to foster, and ultimately integrate, competitive transformational innovation. Venture startups can be creative, bold, driven with vision, and can accept risks in ways well established entities cannot. Their risks can be positively disruptive—to all our benefit—or fail without bringing the wider field with them. Free market competition is a feature of a healthy field that can help us ensure that large vendors enjoy large market shares because they have the most competitive products, not because we have suppressed competition.

Whether looking at academic pursuits or commercial creativity, a common denominator for opportunity is access to data. There is great precedent for this in the medical imaging field, e.g., the DICOM image standard revolution that took place in the medical imaging field in the 1990s [22]. Although establishing imaging data to a standard format faced skepticism at the time, it opened a new age of innovation and research capacity in our field. We have ubiquitous use of DICOM data in research and numerous companies excel in providing products wholly based in the DICOM image space. Looking further across our field, we see pharmaceutical development driven by startups—exemplifying the free market model of risks for few, rewards for many. Adjacent image analysis fields, like pathology, are already being disrupted by innovative startups dependent on access to high fidelity data [23]. As we move forward into an AI-supported nuclear medicine field, it would be a mistake to limit progress to 1990s data standards (i.e., DICOM). Whatever fortitude or foresight it required to establish the DICOM standards in the 1990s, it was clearly the right idea for the time. It is time now to examine the current obstacles and potential opportunities to muster that foresight again.

The final and possibly most important opportunity worthy of mention is the prospect of using scalable data science solutions to expand the medical imaging field. There are currently large disparities in access to medical imaging across the globe [24]. Increasing access to quality healthcare is a principal value to prioritize in the evolution of our field. A recent review presented opportunities to improve cancer survival in lower income countries more than a factor of 10, in part through expansion of imaging access [25]. Accordingly, enabling lower cost hardware and investing in scalable data science solutions while improving the infrastructure to bolster them can expand our influence more efficiently and may be a key element to a more inclusive era of medicine [24, 26].

Ultimately, through all the complexities of translating innovation, synergies, biases, healthcare inequalities, definitions of success and leadership, and imaginative ideas we have not thought of yet, the solution to bolster the next generation may be quite simple in concept: democratization of medical imaging data across the imaging community - making data accessible at its highest fidelity [11, 27]. A large step in that direction could be as simple as evolving expectations for hardware vendors to publicly describe their raw data, or as complex as creating standard data formats across modalities or innovating new data sharing networks and spaces. As a medical imaging community, we should discuss the who, what, where, why, when, and how so that everyone who wants to be a part of the change can contribute.

A data oriented moonshot

Currently, scientific momentum and motivation in nuclear medicine and PET instrumentation largely coalesces around (often expensive) hardware-centric innovations and associated data processing. The twenty-first century data science landscape may propel new alternative models for the evolution of our field. The imaging data-centric moonshot for global health may look like this:

In 20 years, we will (a) expand access to imaging by a factor of 10, (b) witness > 100 startup companies utilizing raw medical image data come to fruition, and (c) deliver care to > 1 million nuclear medicine patients using AI networks trained with > 1 million data sets (d) throughout > 100 countries.

This moonshot aims to foster a stewardship of data that incorporates a modern vision of its use. Importantly, this proposal allows cost reductions, viable through expanded access, through cultivation of an innovation landscape

that encourages scalable technology development and application.

Closing remarks

The biggest hurdle to evolution towards medical imaging data democratization across the field may be intellectual inertia. There is a famous quotation attributed the Willie Sutton, an American bank robber in the early twentieth century. When asked why he robbed banks, he replied concisely “I rob banks because that’s where the money is.”

New roles around data leadership are emerging because that is where the opportunity is. Our diligence now can impact the shape and reach of medicine for coming generations. Literature discussing the unique opportunities in embracing open science and democratization of data is already plentiful. Many industries are moving towards open access, valuation of data, expanding opportunities for inclusion, and transcending bottlenecks across uses of data. The field of nuclear medicine is well suited to provide leadership for change: our images are rich with meaningful data, and we already have extensive experience in cohesively integrating a variety of specialists. As new landscapes of opportunity for data-driven innovation open, we can lead by becoming stewards of imaging data, opening access to its highest fidelity embodiment, and promoting principles of accessibility and free market competition in the data use space.

To close with a metaphor, the wind of change is blowing, it would be a shame not to take the opportunity to reposition our sails.

Funding AK is supported by the National Institute of Biomedical Imaging and Bioengineering (U01EB028234).

Data availability Not applicable.

Code availability Not applicable.

References

- Sharma P, Mukherjee A. Newer positron emission tomography radiopharmaceuticals for radiotherapy planning: an overview. *Ann Transl Med.* 2016;4:53. <https://doi.org/10.3978/j.issn.2305-5839.2016.01.26>.
- Mittra E, Bodei L. Landscape analysis of phase 2 and 3 clinical trials for targeted radionuclide therapy. *J Nucl Med.* 2021;62:1031–2. <https://doi.org/10.2967/jnumed.120.258103>.
- Herrmann K, Schwaiger M, Lewis JS, Solomon SB, McNeil BJ, Baumann M, et al. Radiotheranostics: a roadmap for future development. *Lancet Oncol.* 2020;21:e146–56. [https://doi.org/10.1016/S1470-2045\(19\)30821-6](https://doi.org/10.1016/S1470-2045(19)30821-6).
- Dacek MM, Veach DR, Cheal SM, Carter LM, McDevitt MR, Punzalan B, et al. Engineered cells as a test platform for radiopharmaceuticals in pretargeted imaging and radioimmunotherapy applications. *Bioconjug Chem.* 2021;32:649–54. <https://doi.org/10.1021/acs.bioconjugchem.0c00595>.
- Scott AM, Bodei L. Pharmacogenomics in radionuclide therapy: impact on response to theranostics. *J Nucl Med.* 2021;62:884–5. <https://doi.org/10.2967/jnumed.120.254995>.
- Bodei L, Raj N, Do RKG, Krebs S, Reidy D, Kidd M, et al. Blood based genomic assessment of the clinical efficacy and toxicity of peptide receptor radionuclide therapy (PRRT). *J Nucl Med.* 2021;62:78.
- Aide N, Lasnon C, Kesner A, Levin CS, Buvat I, Iagaru A, et al. New PET technologies - embracing progress and pushing the limits. *Eur J Nucl Med Mol Imaging.* 2021;48:2711–26. <https://doi.org/10.1007/s00259-021-05390-4>.
- Rogasch JMM, Boellaard R, Pike L, Borchmann P, Johnson P, Wolf J, et al. Moving the goalposts while scoring—the dilemma posed by new PET technologies. *Eur J Nucl Med Mol Imaging.* 2021;48:2696–710. <https://doi.org/10.1007/s00259-021-05403-2>.
- Adner R. Match your innovation strategy to your innovation ecosystem. *Harv Bus Rev.* 2006;84(9):89–107:48.
- Chae B. A General framework for studying the evolution of the digital innovation ecosystem: the case of big data. *Int J Inf Manage.* 2019;45:83–94. <https://doi.org/10.1016/j.ijinfomgt.2018.10.023>.
- Payne PRO, Shah NH, Tenenbaum JD, Mangravite L. Democratizing health data for translational research. *Pac Symp Biocomput.* 2018;23:240–6.
- Joly P-B. Beyond the competitiveness framework? Models of innovation revisited. *Journal of Innovation Economics.* 2017;22:79. <https://doi.org/10.3917/jie.pr1.0005>.
- Meikle SR, Sossi V, Roncali E, Cherry SR, Banati R, Mankoff D, et al. Quantitative PET in the 2020s: a roadmap. *Phys Med Biol.* 2021;66:06RM1. <https://doi.org/10.1088/1361-6560/abd4f7>.
- Pogue BW, Zhang R, Gladstone DJ. A roadmap for research in medical physics via academic medical centers: The DIVERT Model. *Med Phys.* 2021;48:3151–9. <https://doi.org/10.1002/mp.14849>.
- Tran-Gia J, Denis-Bacelar AM, Ferreira KM, Robinson AP, Calvert N, Fenwick AJ, et al. A multicentre and multi-national evaluation of the accuracy of quantitative Lu-177 SPECT/CT imaging performed within the MRTDosimetry project. *EJNMMI Phys.* 2021;8:55. <https://doi.org/10.1186/s40658-021-00397-0>.
- Schaller RR. Moore’s law: past, present and future. *IEEE Spectr.* 1997;34:52–9. <https://doi.org/10.1109/6.591665>.
- Kesner A, Laforest R, Otazo R, Jennifer K, Pan T. Medical imaging data in the digital innovation age. *Med Phys.* 2018;45:e40–52. <https://doi.org/10.1002/mp.12794>.
- Moulik S. Data as the New Currency—How open source toolkits have made labeled data the core value in the AI marketplace. *Acad Radiol.* 2020;27:140–2. <https://doi.org/10.1016/j.acra.2019.09.016>.
- Carris NW, Cheon B, Wolfson J. Open data, trials and new ethics of using others’ work. *J Med Ethics.* 2020. <https://doi.org/10.1136/medethics-2019-105898>.
- Vicente-Saez R, Martinez-Fuentes C. Open Science now: a systematic literature review for an integrated definition. *J Bus Res.* 2018;88:428–36. <https://doi.org/10.1016/j.jbusres.2017.12.043>.
- Larson DB, Magnus DC, Lungren MP, Shah NH, Langlotz CP. Ethics of using and sharing clinical imaging data for artificial intelligence: a proposed framework. *Radiology.* 2020;295:675–82. <https://doi.org/10.1148/radiol.2020192536>.

22. Bidgood WD Jr, Horii SC. Introduction to the ACR-NEMA DICOM standard. *Radiographics*. 1992;12:345–55. <https://doi.org/10.1148/radiographics.12.2.1561424>.
23. Campanella G, Hanna MG, Geneslaw L, Miraflor A, Werneck Krauss Silva V, Busam KJ, et al. Clinical-grade computational pathology using weakly supervised deep learning on whole slide images. *Nat Med*. 2019;25:1301–9. <https://doi.org/10.1038/s41591-019-0508-1>.
24. Hricak H, Abdel-Wahab M, Atun R, Lette MM, Paez D, Brink JA, et al. Medical imaging and nuclear medicine: a Lancet Oncology Commission. *Lancet Oncol*. 2021;22:e136–72. [https://doi.org/10.1016/S1470-2045\(20\)30751-8](https://doi.org/10.1016/S1470-2045(20)30751-8).
25. Ward ZJ, Scott AM, Hricak H, Abdel-Wahab M, Paez D, Lette MM, et al. Estimating the impact of treatment and imaging modalities on 5-year net survival of 11 cancers in 200 countries: a simulation-based analysis. *Lancet Oncol*. 2020;21:1077–88. [https://doi.org/10.1016/S1470-2045\(20\)30317-X](https://doi.org/10.1016/S1470-2045(20)30317-X).
26. Frija G, Blažić I, Frush DP, Hierath M, Kawooya M, Donoso-Bach L, et al. How to improve access to medical imaging in low- and middle-income countries? *EClinicalMedicine*. 2021;38. <https://doi.org/10.1016/j.eclinm.2021.101034>.
27. Cornelissen J. The democratization of data science. *Harv Bus Rev*. 2018.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.