Is There Value for Artificial Intelligence Applications in Molecular Imaging and Nuclear Medicine?

Gerold Porenta

Department of Nuclear Medicine, Ambulatorium Döbling, Vienna, Austria

Core competencies in molecular imaging and nuclear medicine include imaging with radioactive isotopes, pattern recognition, image interpretation, and report communication. In diagnostic imaging, nuclear medicine physicians communicate with referring physicians to establish indications for imaging studies, perform or supervise imaging procedures to obtain high-quality images, interpret images to compile medical reports, and communicate results to inform patients and their referring doctors. In many of these tasks, computers already are indispensable tools that enable or assist the work of physicians. How does artificial intelligence (AI) fit into this workflow (Fig. 1)?

There is no universal definition of AI (*1*), but in the context of practical applications, AI can be considered a scientific discipline that uses computers to perform tasks usually requiring human cognition. Research topics on AI include pattern recognition, natural language processing, machine learning, problem solving, and knowledge representation. Thus, AI techniques apply to many core competencies in metabolic imaging and nuclear medicine, and this has led to an enormous interest and even hype about AI in medical imaging (*2*). AI applications at present dominate the technical exhibitions of international imaging conferences such as the 2019 European Congress of Radiology (https://ecronline.myesr.org/ecr2019/) and are prominently featured in the product portfolios of most commercial companies of medical imaging technology.

AI applications have enabled a multibillion-dollar business model for the large Internet companies penetrating the consumer market in many ways, including Internet searches, social networks, or smartphone software. In consumer markets, value is measured by the willingness of a consumer to use or pay for a product. Therefore, the value of AI applications in consumer markets depends not on an intrinsic objective value but the perception of the customer. This value concept is well illustrated by a quote attributed to Charles Revson, the founder of a cosmetics company: "In the factory we make cosmetics, in the store we sell hope."

However, in medicine the business model and the value concept of a consumer market do not apply. Hope is essential when taking care of patients, but hope cannot serve as the foundation for medical decision making and patient stratification. Several decades ago, patients with extrasystoles after a myocardial infarction were commonly treated with antiarrhythmic drugs in the hope of preventing arrhythmic death. A large, randomized multicenter trial uncovered this hope as a deadly illusion and established that the antiarrhythmic therapy in fact increased arrhythmic mortality and caused premature death in thousands of patients. From this trial and others that were able to scientifically disprove suggested or assumed benefits based on hope and hype, rigorous and often laborious scientific methods have emerged that permit establishing, grading, or disproving the value of diagnostic and therapeutic procedures.

For patients, value is generated if quality of life is improved, morbidity is reduced, or preventable mortality is eliminated. It is difficult or impossible for both patients and physicians to directly assess the value of medical products, interventions, or technology. Thus, science is essential to firmly guide patient care. International guidelines compiled by professional societies assess the value of diagnostic and therapeutic methods based on scientific evidence and, for each method, clearly state the class of recommendation for a particular use with an associated level of evidence. This science-based model of grading value for patient care has been universally adopted in the medical community. Currently, there is little scientific evidence for the value of AI applications in medical imaging. Several AI applications have received Food and Drug Administration approval (*3*), but this does not imply that AI applications at present are relevant for medical practice or that their value has been firmly established.

Most AI applications are built from 3 essential components: complex computer algorithms, extensive computing resources, and large data sets.

Algorithms and computing facilities are easily available at little or no cost. Many powerful AI algorithms are published as open-source code, such as TensorFlow (https://www.tensorflow.org) and Core ML (https://developer.apple.com/machine-learning/). Extensive computing power is offered instantaneously on demand through the large Internet companies. Recently, even smartphones have been equipped with high-power computing hardware, including neural network chips that enable intensive AI applications, such as facial identification or speech recognition.

In contrast, large data sets are more difficult to collect and thus are the most critical component when building AI applications. Data are more important than hardware or software in determining the success of AI applications (*4*). Even highly complex AI algorithms cannot compensate for incomplete, inadequate, or low-quality data collection. Thus, the characteristics and validity of data sets need to be firmly established and made fully transparent when AI applications are investigated for a proposed clinical purpose.

Collaborative efforts are frequently required to compile the large data sets, which need to include many thousands of data entries. Recently, the Mozilla Common Voice project (https://voice.mozilla.org/en) has published an open-source multilanguage data set of voice recordings from 40,000 people in 18 languages to foster and enable research in natural language processing. If AI applications are to be applied successfully in medical imaging,

Received Mar. 18, 2019; revision accepted Apr. 3, 2019.
For correspondence or reprints contact: Gerold Porenta, Ambulatorium Döbling, Heiligenstädterstrasse 62-64, 1190 Vienna, Austria.
E-mail: gerold@porenta.com
Published online May 3, 2019.
COPYRIGHT © 2019 by the Society of Nuclear Medicine and Molecular Imaging.
DOI: 10.2967/jnumed.119.227702

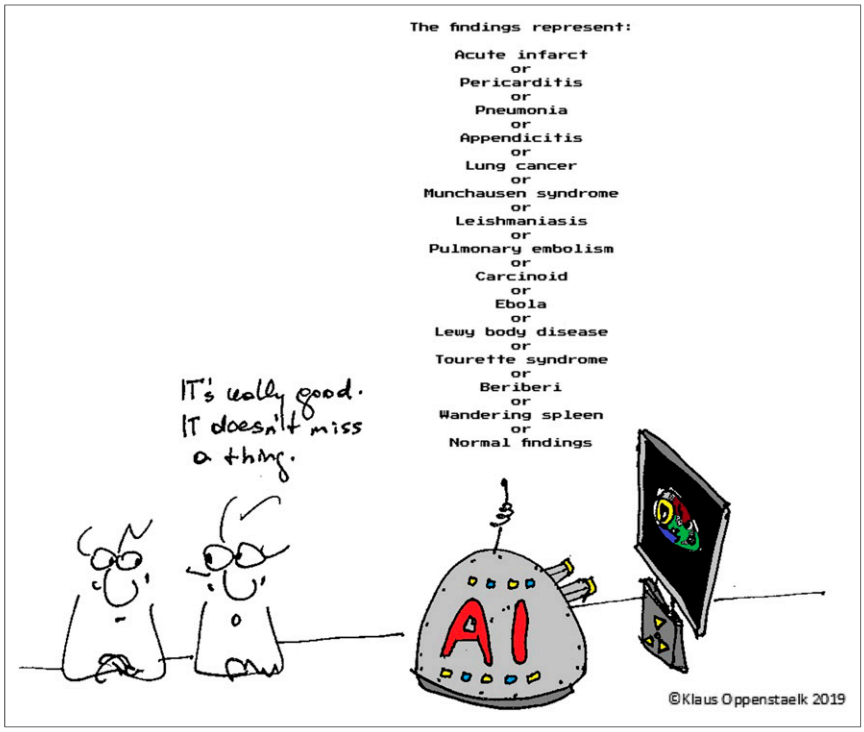

FIGURE 1. The value of AI in nuclear medicine? (Courtesy of Klaus Oppenstaelk.)

similar efforts are likely needed to collect validated data sets of adequate size to engage state-of-the-art AI technology for advancing image processing and interpretation. Such data sets are critical not only to develop AI applications but also to elucidate their performance and value in patient care.

In molecular imaging and nuclear medicine, large data sets are difficult to collect. The individual data in most health care systems belong to the patient and thus can be shared only when proper patient consent has been obtained. This is usually a laborious process and may be prohibitive for large-scale collections based on electronic patient records. New technology may provide better ways to share and collect patient data by putting the patient in control. Blockchain technology developed for cryptocurrencies has recently been suggested by, for instance, Nebula Genomics (https://www.nebula.org) as an interesting alternative to share genetic data between patients and research institutions and permit use for analysis in larger networks. In a blockchain-based workflow, patients retain full control over the access and use of their data. This is particularly important, because a labeled data set for machine learning needs to also include detailed phenotype and genotype data that likely preclude irreversible anonymization of the data records.

Although an assessment of AI applications needs to use established scientific tools, the intrinsic dynamic nature of AI applications poses novel challenges for the application of these methods. In particular, learning algorithms frequently evolve dynamically with growing data sets and cannot be confined to a static framework that can be subjected to standard randomized trials. The need for better tools to assess the value of AI applications has been recognized, and policies and concepts have been suggested on how these problems can be addressed (*5*). It will be interesting to observe how science can advance to accommodate the new needs for designing, developing, deploying, and validating dynamic learning AI applications.

Innovation and advances in technology can occur quickly, in a disruptive fashion, but can also happen in a more subtle and evolutionary way. AI is frequently considered a disruptive technology, but in nuclear medicine computer applications and in particular AI, technologies have been studied for more than 2 decades. Early research has demonstrated the feasibility of using neural networks to diagnose coronary artery disease from myocardial perfusion imaging (*6,7*). Although further progress since these early attempts has initially been slow, partly because of the lack of large data sets, AI technology has expanded exponentially since entering the consumer market. Machine learning has transformed into a mature and robust technology that appears to proceed to large-scale adoption in medical imaging. Novel AI applications come in many different forms with different scopes or targets, and their impact on patient care can range from negligible to potentially contributing to reduced morbidity and mortality.

The impact of AI technologies on patient outcome needs to be considered when gauging the intensity needed to assess the value of AI applications. AI algorithms that provide transparent information to physicians that can be entirely corrected or edited by manual interaction, such as identification of the cardiac axes in image reorientation or detection of myocardial borders, do not have a direct impact on patient care. These AI-based software tools assist in image processing but remain under the full control of human experts. Therefore, a focused technical evaluation rather than an extensive clinical validation may be adequate for these ancillary AI tools.

In contrast, AI applications that mimic human intelligence without transparency or an optional correction by human experts, and therefore have the potential to affect patient care, need a thorough scientific validation similar to that for drugs and medical devices. For example, AI applications that use deep-learning strategies to recommend for or against coronary revascularization by analysis of myocardial perfusion or metabolic images without physician interaction potentially exert a significant and critical impact on patient care and thus need extensive and rigorous validation. Such AI applications will likely be accessible globally and may affect the lives of many more patients than a single doctor could. Given the potential to harm thousands of patients if AI applications contain errors, scrutiny is required when their value in clinical use is established. One could envision that such AI applications—especially when based on machine learning—would have to undergo a repetitive recertification process such as human experts do.

Advances in imaging technology and personalized medicine have dramatically increased the amount of data that need to be reviewed and put into individual context when taking care of patients. As image resolution in all imaging modalities has continuously increased, as multimodality and multidimensional imaging has been adopted in routine clinical work, and as treatment cycles have accelerated, physicians have been confronted with a rapidly

growing data load that precludes comprehensive interpretation by humans alone. Imaging data, patient phenotype, and increasingly also genotype, in addition to a growing evidence base published in guidelines and scientific literature, all need to be considered and put to proper use when managing individual patients with individual clinical challenges.

Because this data tsunami far exceeds the capacity of even the brightest human minds, data analysis support is urgently needed on multiple levels in the clinical workflow. AI applications likely could significantly alleviate many human shortcomings. They have shown potential in minimizing radiation exposure, improving image quality, accelerating image analysis, supporting image interpretation, generating structured reports, and performing risk stratification. Most importantly, all these achievements by knowledge-centered AI applications are available 24–7, with a precisely defined quality in the setting of personalized medicine tailored to individual patients and not as population-based estimates for which individual features of patients are largely eliminated by statistical methodology. However, whereas the potential value of AI applications appears evident and even cogent, scientific validation is still the missing link that at present impedes large-scale clinical adoption of AI imaging applications.

In summary, currently there is little scientific evidence that AI applications in molecular imaging and nuclear medicine can assist in improving the quality of life for patients, reduce morbidity, or eliminate premature mortality. AI technology is a main driver for innovations in computing technology, has been instrumental in data science, and will certainly soon exert a profound impact on medical care. However, the value of AI applications in medical care can be confirmed only when professional guidelines provide recommendations for their use in specific clinical settings and patient populations.

A final personalized facet: I am ready to concede value to an AI imaging application if it participates in the imaging challenge at the annual meeting of the Society of Nuclear Medicine and Molecular Imaging as a regular contestant and is able to score better than at least 1 group of human experts. From my past experience with AI technology, I do not expect this to happen soon, but I would hope to witness it before I retire in 10 years.

DISCLOSURE

No potential conflict of interest relevant to this article was reported.

ACKNOWLEDGMENT

The editorial assistance of Ellen Casey is very much appreciated.

REFERENCES

1. Russel S, Norvig P. *Artificial Intelligence: A Modern Approach.* 3rd ed. Essex, England: Pearson Education; 2016.
2. Ranschaert ER, Morozov S, Algra PR, eds. *Artificial Intelligence in Medical Imaging: Opportunities, Applications and Risks.* New York, NY: Springer; 2019.
3. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. *Nat Med.* 2019;25:44–56.
4. Halevy A, Norvig P, Pereira F. The unreasonable effectiveness of data. *IEEE Intelligent Systems.* 2009;24(2):8–12.
5. Evidence standards framework for digital health technologies. National Institute for Health and Care Excellence website. www.nice.org.uk/digital-evidence-standards. Published March 2019. Accessed July 18, 2019.
6. Fujita H, Katafuchi T, Uehara T, Nishimura T. Application of artificial neural network to computer-aided diagnosis of coronary artery disease in myocardial SPECT bull's-eye images. *J Nucl Med.* 1992;33:272–276.
7. Porenta G, Dorffner G, Kundrat S, Petta P, Duit-Schedlmayer J, Sochor H. Automated interpretation of planar thallium-201-dipyridamole stress-redistribution scintigrams using artificial neural networks. *J Nucl Med.* 1994;35:2041–2047.