Artificial Intelligence in Medical Imaging: An Overview of a Decade of Experience

Orit Wimpfheimer MD and Yotam Kimmel BSc

Nanox Vision, Petah Tikva, Israel

ABSTRACT Medical imaging data has been at the frontier of artificial intelligence innovation in medicine with many clinical applications. There have been many challenges, including patient data protection, algorithm performance, radiology workflow, user interface, and IT integration, which have been addressed and mitigated over the last decade. The AI products in imaging now fall into three main categories: triage artificial intelligence (AI), productivity AI, and augmented AI, each providing a different utility for radiologists, clinicians, and patients. Adoption of AI products into the healthcare system has been slow, but it is growing. It is typically dictated by return on investment, which can be demonstrated in each use case. It is expected to lead to wider adoption of AI products in imaging into the clinical workflow in the future.

IMAJ 2024; 26: 128–131

KEY WORDS: artificial intelligence (AI), artificial intelligence (AI) radiology workflow, coronary artery calcium, osteoporosis, population health

Artificial intelligence has become the new buzzword to signify innovation in the digital world with its myriad of use cases in the world of manufacturing, marketing, security, banking, education, and healthcare. Within healthcare, diagnostic imaging was seen as the low hanging fruit to enable artificial intelligence algorithms to thrive, given the immense amount of imaging data available for training. In fact, many companies have invested considerable amounts of funding to make artificial intelligence (AI) in imaging a reality and to gain global recognition and success.

With almost a decade of experience in the world of artificial intelligence in imaging, these companies encountered and overcame many challenges as they tried to penetrate healthcare organizations. Some challenges, however, are still unresolved. The challenges involved many domains including algorithm performance, patient data protection, radiology workflow, user interface (UI), clinical use cases, IT challenges, and return on investment (ROI).

The first few years of AI in imaging were dominated by challenges in the performance of the algorithms. The algorithmic challenge was not only to have a high enough sensitivity for a given algorithm's clinical indication but to maintain high specificity. While sensitivity to find acute conditions was critical, it typically came at the expense of specificity, particularly for the low prevalence findings such a pneumothorax or acute intracranial hemorrhage. For a radiologist, highlighting several false positive cases to detect a positive case was an unwelcome interference in the daily workflow of the already busy radiologist. In institutions where radiologists were not immediately available to read all the studies, these AI results were sometimes available to interns and residents in the intensive care unit or in the hospital wards. These practitioners were sometimes misled by the false positive findings, which occasionally led to medical confusion. As time progressed, this issue was minimized, although not completely resolved, by using larger data sets for training and also by providing confidence levels of the algorithm results, reminding the user that the algorithm is not a binary assessment of pathology but, like the trained radiologist, there are confidence levels of finding that should be considered.

The issue of the generalizability of an algorithm became a hot topic for a while. Would an algorithm have the same performance on varied, unseen data, as compared to the training data sets? While this problem was a key area of concern in the early stages of AI imaging, creating skepticism in the early phases of deployment, it was a relatively easily surmountable problem. As there were a limited number of medical imaging device manufacturers, imaging protocols were relatively similar across the field and patient ethnicity had little impact on the images themselves. If the training set was robust, the algorithm should be generalizable across institutions. Over time, clinician skepticism around this issue has been dramatically curtailed.

The second concurrent challenge was the radiologist workflow. As busy physicians manage the workload of cases, any additional clicks that are required limit the workflow of the radiologist and must be justified [1]. Many companies installed a widget on the radiologist workstations with a UI designed to attract the radiologist to look at the positive cases. Often the widget went unaddressed, given the extra clicks it required. Thus, AI algorithms were installed but were often underused. Improvements such as re-organizing the PACS worklist to prioritize the positive cases detected by the algorithm did improve the utility of the AI for acute findings. Another option was to create a secondary capture series within the study such that the algorithm results became fully incorporated into the case and saved on PACS. While this procedure was fully integrated into the radiology workflow, it created its own challenges as the AI results, which were not 100% accurate, became part of the permanent medical record of the patient. In addition to making the AI imaging results available, making the text results available either by PDF or within the radiology dictation system were additional attempts to make the results accessible in a workflow amenable to the radiologist. Yet, despite these various attempts, only a small percentage of institutions paid for commercially available imaging AI products.

IMAGING DATA IS A PRIME OPPORTUNITY TO USE ARTIFICIAL INTELLIGENCE TO IMPROVE RADIOLOGIST PRODUCTIVITY AND SENSITIVITY TO FACILITATE IMPROVED PATIENT MANAGEMENT.

Over time, it became clear that while the AI product performance, workflow, and UI were extremely important, the clinical use cases were critical in determining the viability of a commercial AI product. The clinical use cases were grouped into one of three general categories. The first use case was the triage model. The algorithms identified positive acute urgent results to ensure that those cases were read more quickly by the radiologist, and that the radiologist did not miss even subtle positive cases. At first glance, this protocol seemed the most natural fit for AI. In addition, regulating such a triage product with the U.S. Food and Drug Administration (FDA) was a relatively simple process. As such, many companies followed this protocol, although most of those companies struggled to produce significant revenue from findings, such as intracranial hemorrhage and pneumothorax.

The first real success in the world of acute pathology is stroke imaging, where every minute counts as patients with large vessel occlusion (LVO) are directed to the neurovascular suite for intervention [2]. In fact, the first AI product to receive new technology add on payment (NTAP) was for detection of LVO on head computed tomography (CT) angiograms. What further made the use of AI in stroke imaging unique, was that the results of the algorithm was available not only to the radiologist on call but simultaneously to the stroke team on call via an application on the phone. This usage was a game changer because it enabled a measurable decrease in time to interventional procedure for large vessel occlusion that led to improved clinical outcomes and decreased length of hospital stays. As such, many hospitals in the United States currently deploy AI solution for stroke imaging.

Another well-received AI product was the detection of incidental pulmonary embolism, as the product showed that many small peripheral pulmonary emboli were often missed by radiologists evaluating non-acute oncology patients whose routine CT scans could take several days to analyze. In select institutions, pulmonary care teams also had an application to alert them about patients with large pulmonary emboli who might need interventional thrombectomy [3].

The second use case in which many companies created commercial AI products was radiology productivity. The time of a radiologist, especially with the critical shortage of trained radiologists, is expensive. There is also increasing burnout of radiologists who often feel the stress of never-ending daily workloads. AI products that can identify small findings, measure them accurately, and compare them to prior exams are most useful for a radiologist's productivity. These types of products also act as a safety net to radiologists and administrators who fear missing findings that could have potential medical and legal ramifications. These companies often focus on pulmonary nodules and breast imaging as fields where this use case is particularly suited. Yet, while some of these companies are successful, most radiology centers do not use them. Typically, the reason cited is financial. Most radiologists or radiology departments do not want to purchase these tools as an added expense, particularly if the system seems to work sufficiently without that monetary investment [4]. In some cases, radiologists lure fellow radiologists to their facilities with these types of added workflow enhancements, as the competition for radiologists in the current market is fierce.

In addition to providing a second reader safety net for radiologists, there is another emerging use case of using imaging AI as a companion to the non-radiologists who analyze X-rays such as in the emergency department (ED), acute care facilities, or inpatient settings [5]. Often, the clinical workflow does not include the radiologist reporting X-ray images immediately due to the volume of X-rays obtained. As such, chest X-rays and skeletal X-rays get evaluated by non-radiology clinicians such as ED doctors, general physicians, and other subspecialists seeing the patient. Having AI tools to detect fractures and to provide a variety of common chest X-ray findings may be quite useful in those settings.

The most recent of the use cases has been termed *AI augmentation*. This term refers to AI tools that empower a radiologist to provide more detailed reports without adding significant additional time and effort by the radiologist, typically focusing on chronic conditions that are unaddressed by a radiologist who is focused primarily on the acute clinical indication. Findings such as osteoporosis and chronic vertebral body compression fractures often go unreported and, as these findings are typically clinically silent, they are severely undertreated. As the number of CT scans has grown exponentially in the last 20 years, many patients in the 40-to-70-year age range get a CT scan for some clinical indication [6]. Abdominal pain, renal colic, cough, and trauma are some of the many acute clinical indications that often result in a CT scan. While osteoporosis can often lead to a devastating hip fracture, bone mineral density and chronic vertebral body compression fractures go unaddressed in most CT reports. The cohort with the highest likelihood of hip fracture often go unreported. Yet, osteoporosis treatment and management can significantly reduce the number of elderly who present with hip fractures. Using AI powered tools to highlight such chronic findings has the potential to empower radiologists to decrease the burden of such chronic diseases by directing patients to appropriate clinical management. By getting patients the preventative treatment early in their disease process, we can decrease the impact of those chronic diseases which in the case of osteoporosis would mean fewer hip fractures [7].

Another example is cardiovascular disease, which is the number one cause of death worldwide. It has been well established in the cardiology literature that coronary artery calcifications are a better risk predictor for a future cardiovascular event than any of the risk calculators used in clinical practice [8]. A standard coronary artery calcium score was established using electrocardiogram-gated non-contrast chest CT. The technique was designed to evaluate the coronary artery calcium using Agatston units, although it is only used and reimbursed in certain clinical circumstances. With improved image acquisition in the newer CT scans, the effect of cardiac motion has decreased. In addition, AI tools have been developed that can identify, segment, and measure the coronary artery calcium, excluding the densities created by cardiac motion on standard non-contrast CT. Using this tool, all patients can be opportunistically screened for cardiovascular disease on a CT scan they obtained for another clinical indication. Few radiologists mention these calcifications. Some state that there are coronary calcifications without committing to severity while even fewer analyze the amount, which has high inter-reader variability. Even if there is a mention of coronary artery calcium in the radiology report, often patients do not get the preventative cardiology and clinically recommended treatment because cardiologists treat patients based on cardiac risk categories that are highly correlated with calcium score categories, something not inherent in the radiology lexicon and therefore not present in the radiology report. Augmentative AI can change that paradigm so that these patients are identified by the appropriate calcium category to provide the referring physicians with the information they need to direct patients to the appropriate guidelines based preventative healthcare management [9].

At the Society of Cardiovascular Computed Tomography conference in 2023, the cardiology department at Belinson Hospital, Petah Tikvah, Israel, part of Clalit Health Services, presented a poster showing results from their study, which highlighted the use of a commercially available AI product to measure coronary artery calcium on non-contrast CT scans. Of the 326 patients who met the inclusion criteria for the study, ages 30-75 with access to EMR information, the authors found that 101 patients (31%) had severe coronary artery calcium (score of > 400) but had no prior history of cardiovascular disease. Those patients were directed to a dedicated preventative care cardiology clinic. In addition, they found that 88 patients (27%) had moderate amounts of coronary artery calcium (27%) and directed those patients to their primary care physicians for further evaluation and treatment.

In the sphere of IT integration, the first hurdle was patient data protection and HIPAA compliance. However, that concern was allayed using anonymization tools enabling vendors to access the medical images and related information images for the AI algorithms but without the patient identifying data. Whether using on premise solutions for the more conservative institutions or cloud-based integration for the more innovative institutions, this constraint should no longer be a limiting factor.

The newest hurdle for broader imaging AI adoption is the IT departments at institutions. As there are now over 500 FDA cleared AI applications, many institutions have become inundated with IT complexities as they try to work with each vendor's platform to implement various AI solutions, one vendor at a time. This difficulty created a rate-limiting bottleneck for innovation even in the most progressive of institutions. As such, over the last 2 years, the market has gone from individual vendor installations toward a marketplace paradigm where many Imaging AI vendors integrate into one or more of several available marketplaces. After an institution chooses a marketplace and successfully integrates that platform, all the AI products they offer are available without an additional IT effort, which significantly enables integration of a broad spectrum of AI solutions efficiently [10].

The AI in imaging market has matured over the last 10 years, with considerable advancements in algorithm development, generalizability, UI, workflow, health data protection, use cases, and more recently marketplace integration. Yet, the ultimate hurdle of demonstrating a ROI remains as healthcare institutions are chronically underfunded and must legitimize the commercial investment of these new AI tools. Creating a viable ROI is a challenge faced by many imaging AI companies, although some use cases offer an easier path to success. Within the triage use case, stroke imaging is the best return on investment as using AI to decrease time to the interventional suite can reduce the length of hospital stay and improve patient recovery. In Europe, where a second reader is required for screening mammograms, using AI as an alternate second reader has a clear ROI. While radiologists enjoy the benefits of productivity tools subjectively, a measurable ROI of such productivity tools has been harder to prove.

There has been much interest lately in population health screening of chronic findings such as osteoporosis, coronary artery calcium, and aortic aneurysms as previously identifying these undetected patients brought in new revenue in fee-for-service institutions in the short term while also improving long-term patient outcome potentially decreasing the financial burden on patients as their conditions deteriorated. In the United States, reimbursable current procedural terminology (CPT®) codes is another strategy being pursued by several vendors to try to impact the ROI of using such products. There are currently several CPT III codes for using AI tools for population health screening on CT scans; however, CPT III codes are designed for innovative products and do not have financial reimbursement. Those CPT III codes can serve as a first step toward CPT I codes, which do come with financial reimbursement if it is demonstrated that there is widespread use of the code [11].

CONCLUSIONS

The last decade has seen the development of an entire industry of artificial intelligence in imaging, which has overcome many hurdles on its path to maturity. Unlike equipment for medical imaging where the burden of entry is high, companies operating in the field of AI are young and nimble and able to meet each new challenge with prompt solutions. The hope and belief by many in the industry is that facilities will benefit from deeper and broader infiltration of commercial AI products across all medical departments where imaging plays a key role in patient healthcare.

Conflicts of interest

Orit Wimpfheimer MD is a U.S. board certified diagnostic radiologist and was Clinical Director and then Chief Medical Officer at Zebra Medical Imaging from 2019 until its acquisition by Nanox Vision, after which she became Chief Medical Officer at Nanox.

Correspondence

Dr. O. Wimpfheimer
Nanox Vision, Petah Tikva 4951025, Israel
Email: oritwimpfheimer@gmail.com

References

1. Bailey CR, Bailey AM, McKenney AS, Weiss CR. Understanding and Appreciating Burnout in Radiologists. *Radiographics* 2022; 42 (5): E137-9.
2. Matsoukas S, Morey J, Lock G, et al. AI software detection of large vessel occlusion stroke on CT angiography: a real-world prospective diagnostic test accuracy study. *J Neurointerv Surg* 2023; 15 (1): 52-6.
3. Ebrahimian S, Digumarthy SR, Homayounieh F, Bizzo BC, Dreyer KJ, Kalra MK. Predictive values of AI-based triage model in suboptimal CT pulmonary angiography. *Clin Imaging* 2022; 86: 25-30.
4. Huisman M, Ranschaert E, Parker W, et al. An international survey on AI in radiology in 1,041 radiologists and radiology residents part 1: fear of replacement, knowledge, and attitude. *Eur Radiol* 2021; 31 (9): 7058-66.
5. Woznitza N, Piper K, Burke S, Bothamley G. Chest X-ray interpretation by radiographers is not inferior to radiologists: a multireader, multicase comparison using JAFROC (Jack-knife Alternative Free-response Receiver Operating Characteristics) analysis. *Acad Radiol* 2018; 25 (12): 1556-63.
6. Smith-Bindman R, Miglioretti DL, Johnson E, et al. Use of diagnostic imaging studies and associated radiation exposure for patients enrolled in large integrated health care systems, 1996-2010. *JAMA* 2012; 307 (22): 2400-9.
7. Dagan N, Elnekave E, Barda N, et al. Automated opportunistic osteoporotic fracture risk assessment using computed tomography scans to aid in FRAX underutilization. *Nat Med* 2020; 26 (1): 77-82.
8. Budoff MJ, Young R, Burke G, et al. Ten-year association of coronary artery calcium with atherosclerotic cardiovascular disease (ASCVD) events: the multi-ethnic study of atherosclerosis (MESA). *Eur Heart J* 2018; 39 (25): 2401-8.
9. Sandhu AT, Rodriguez F, Ngo S, et al. Incidental coronary artery calcium: opportunistic screening of previous Nongated Chest Computed Tomography Scans to Improve Statin Rates (NOTIFY-1 Project). *Circulation* 2023; 147 (9): 703-14.
10. Leiner T, Bennink E, Mol CP, Kuijf HJ, Veldhuis WB. Bringing AI to the clinic: blueprint for a vendor-neutral AI deployment infrastructure. *Insights Imaging* 2021; 12 (1): 11.
11. Pickhardt PJ, Correale L, Hassan C. AI-based opportunistic CT screening of incidental cardiovascular disease, osteoporosis, and sarcopenia: cost-effectiveness analysis. *Abdom Radiol (NY)* 2023; 48 (3): 1181-98.