Journal of the American Heart Association

EDITORIAL

Artificial Intelligence: Evaluating AI Imaging for Cardiovascular Disease

David A. Bluemke, MD, PhD; Nadine Kawel-Boehm, MD

At present, at one of the author's institutions, artificial intelligence (AI) algorithms have already evaluated every chest, spine, and brain computed tomography (CT) well before they are evaluated by the radiologist. In our environment, the AI algorithms are lurking in the background, purposed to detect four clinically meaningful abnormalities: pulmonary embolus (PE), lung nodule, cervical spine fracture, and brain hemorrhage. Each of these diagnoses has well-defined clinical implications and treatments. The rapid triage of CT scans by the AI is expected to improve patient outcomes—the radiologist is alerted and reporting of urgent abnormalities for these patients can take precedence over the myriad of other patient abnormalities that can wait.

See Article by Canton et al.

For this "triage" function of AI, there are two main disadvantages. First, currently the AI is not always right. False positives are relatively common, often due to artifacts on less than perfect imaging examinations. Second, the AI is not well informed—eg, pulmonary emboli may be diagnosed previously; subsequent CT scans showing the same PE repeatedly over several days or weeks have no urgency. Yet the uninformed AI continues to alert the radiologist.

Despite these limitations, all radiologists understand the inevitable trend toward automated diagnosis by these AI assistants. Before long, the AI should be enabled to read the electronic health record and understand if a pulmonary embolus was previously reported. If already known, the PE becomes less urgent. The AI can enhance communication faster than medical staff: if a stroke is present, it is time to get the stroke team mobilized. Despite wonderful progress, a chest or brain CT has more abnormalities than PE or stroke. Each advanced imaging examination can have 30 or 40 diagnoses encompassing 50 to 1000 transaxial slices.

As work continues to focus on urgent diagnoses, there is another category of abnormality that has clinical meaning yet remains non-standard. We will refer to these as "opportunistic" imaging findings—radiology findings that primarily affect long-term patient health. One example: on an abdominal CT obtained for acute cholecystitis, the entire lumbar spine is also imaged. It would be a simple matter to have an AI to report lumbar spine bone density on *every* abdominal CT scan. The US Preventive Services Task Force (USPSTF) recommends dual-energy x-ray absorptiometry (DXA) for bone density screening.[1] But CT is much more accurate for bone density—abdominal CT is not a screening test for bone density simply because radiation dose is too high. But bone density—or liver fat or atherosclerotic extent—could be quantified for *all* CT scans once obtained for some other purpose, regardless of the original indication for CT. Hence the term "opportunistic"—the repurposing of imaging tests for reasons other than the clinical indication.

All physicians are already somewhat familiar with opportunistic imaging concepts—you may order the

Key Words: Editorials ■ arterial ■ artificial intelligence ■ atherosclerosis

The opinions expressed in this article are not necessarily those of the editors or of the American Heart Association.

Correspondence to: David A. Bluemke, MD, PhD, University of Wisconsin School of Medicine, 600 Highland Ave., Madison, WI 53703. E-mail: dbluemke@wisc.edu

For Disclosures, see page 3.

© 2021 The Authors. Published on behalf of the American Heart Association, Inc., by Wiley. This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

JAHA is available at: www.ahajournals.org/journal/jaha

Downloaded from http://ahajournals.org by on May 17, 2021

CT scan to check for the size of the aortic aneurysm. But along with that comes a host of radiologist statements about nearly every organ system. The current problem—those statements are *qualitative* and are often directed toward esoteric abnormalities. Perhaps there is a vascular variant, an accessory spleen that could simulate a lymph node. These are interesting but are not part of the "opportunistic" screening pathway. The most common targets of opportunistic imaging to date include (1) bone density, (2) coronary artery calcium/atherosclerotic disease, and (3) liver steatosis (a precursor to the more serious non-alcoholic steatohepatitis).

The topic of opportunistic atherosclerotic disease evaluation is the subject of an important article by Yuan et al in this issue of the *Journal of the American Heart Association* (*JAHA*).[2] The authors report their experience using a deep learning algorithm for atherosclerosis evaluation based on knee magnetic resonance imaging (MRI) obtained for cartilage evaluation, as part of the Osteoarthritis Initiative (OAI). The OAI (URL: https://www.clinicaltrials.gov. Unique identifier: NCT00080171) is a highly respected multi-center cohort study designed to evaluate biomarkers for osteoarthritis as potential surrogate endpoints for arthritis disease onset and progression in ≈4800 study participants.[3] The rationale for studying arthritis is compelling: the number of Americans with symptomatic osteoarthritis was 27 million in 2005. At that time, half a million knee replacements were performed for osteoarthritis. As the US population ages, this need will continue—up to 3.5 million knee replacements are expected by 2030.[1] The OAI investigators aim to find reproducible imaging markers that will help assess the effectiveness of drugs that prevent or postpone the onset of arthritis.

Yuan et al used knee MRI examinations from the OAI in an unexpected manner—to design and test an AI to study the development of atherosclerosis of the popliteal artery. The characterization of peripheral artery disease (PAD) is a compelling topic equal to that of osteoarthritis. An estimated 27 million individuals in Europe and North America are affected by PAD, with more than 400 000 annual patient admissions attributed to PAD. Yuan et al evaluated more than half a million magnetic resonance MRI slices from 4688 study participants in the OAI. The extraordinary extent of data analysis was made possible by a deep learning algorithm that identified more than 9100 popliteal arteries (bilateral arteries were evaluable in most subjects). After finding the popliteal artery on the knee MRI, the algorithm preceded to determine three parameters: the popliteal artery lumen diameter, the wall thickness, and the total artery diameter.

An automated, AI-driven analysis of biologically important arteries on this order of this magnitude has never been previously reported. The important implications of this work are wide-ranging and multifactorial. We begin with the contribution of Canton et al to advantage of knowledge regarding vascular biology and atherosclerosis. Ever since the groundbreaking work of Glagov et al on compensatory arterial enlargement in early atherosclerosis,[4] the concept of positive and negative remodeling has been central to detection, treatment, and prevention of CAD. Glagov and co-authors studied coronary arteries in 136 hearts at autopsy, elegantly showing “coronary arteries enlarge in relationship to plaque area and that functionally important lumen stenosis may be delayed until the lesion occupies 40% of the internal elastic lamina area.” Dr Glagov and co-authors presciently noted that normal lumen areas could be present despite the presence of a large plaque.

On the other hand—autopsy studies like that of Glagov et al have that one, singularly important bias: inclusion only of individuals who have died. Pathologic specimens contract and artifacts occur during histology preparation. What about in vivo proof? To date, noninvasive imaging studies have repeatedly verified the so-called Glagov hypothesis in the coronary arteries. But is the Glagov hypothesis universal? Do all small to moderate size arteries, such as the renal, carotid, behave the same way as the coronary arteries? Noninvasive imaging has said “yes.”

The study by Yuan et al gives us yet an additional vote for yes, the best evidence to date for popliteal artery remodeling. Remarkably enlarged popliteal arteries were identified by the AI algorithm in response to atherosclerosis. Over a 2-fold range of observed arterial wall thicknesses, Canton et al found the corresponding vessel size was greater by a factor of 4. Yet this apparent vascular enlargement had a limit: beyond a wall thickness of 0.9 mm, the size of the arterial lumen was smaller, while the overall vessel diameter was essentially constant. Like the observations of Glagov et al, the observations were cross-sectional rather than longitudinal change. Yet the observations of Canton et al remained consistent after adjustment for body size, gender, and cardiovascular risk factors. Age was the single strongest factor that correlated with mean arterial wall thickness, followed by gender. Factors associated with maximal (rather than mean) thickness in each study participant were not evaluated at this time.

As suggested above, there are broad implications of the author’s results far beyond the OAI data set. Knee MRI examinations are extremely common; the authors report more than 800 000 such evaluations were done in 2017. The initial development of AI is the most challenging. Researchers must train, validate, and test the algorithm by hand. However, subsequent AI software modification, tuned to different MRI scanners and

Downloaded from http://ahajournals.org by on May 17, 2021

acquisition details, are less intensive. The AI by Canton et al, called FRAPPE, could be the basis for wider scale algorithms for opportunistic PAD screening based on knee MRI. Early treatment, prior to clinically evident PAD, would be likely to reduce later morbidity.

But the implications of the work of Canton et al go further: a busy academic radiology practice in a moderate size hospital *frequently* performs CT angiography of the extremities. These examinations entail about 2500 CT transaxial slices. Current reporting is crude and mostly subjective—qualitative adjectives of PAD disease severity, with words like mild, moderate, or severe. Medical imaging examinations simply have *too much information* for words in a clinical radiology report. On the other hand, an AI has little concern for the number of imaging slices: means, maximums, arterial segments can be accurately and rapidly labeled. Accurate measurement leads to accurate assessment of therapeutic response—a step that is currently subjective and incomplete in the work-up of most patients for PAD.

Atherosclerosis does not stop below the knee, or above the coronary arteries. Our radiology departments perform 40 or more different types of imaging examinations for arterial segments throughout the body. In the majority of hospitals, none of these imaging tests are routinely quantified beyond rough approximately (eg, quartiles of stenosis). Yet Canton's work and the underlying foundation of Glagov's work are clear that early atherosclerotic disease is associated with normal or even slightly enlarged luminal diameters—precisely those vessels that are currently classified as "normal" on today's CT, MRI, or ultrasound studies. Clearly many of these are far from normal—atherosclerosis in the arterial wall is the short-term harbinger of severe disease that will ultimately develop. Still, we concentrate today's efforts on treatment of severe arterial disease using stents and dilatations. Medical therapies are already available that reverse atherosclerotic changes. Yet diagnostic imaging methods for early atherosclerosis are mostly unused. The work of Canton et al helps us envision a future where early atherosclerotic disease is readily identified by an AI, ready and waiting for that next imaging examination. Once PAD is identified, risk factors can then be assessed, therapy prescribed, and assessments performed to ensure adequate therapeutic response.

The authors are to be congratulated for their work in the OAI study. We hope that the next 10 years sees continued progress in this area, along with advanced therapeutic trials that convincingly demonstrate benefits of early atherosclerotic treatment.

ARTICLE INFORMATION

Affiliations

Departments of Radiology and Medical Physics, University of Wisconsin, Madison, WI (D.A.B.); Department of Radiology, Kantonsspital Graubuenden, Chur, Switzerland (N.K.); and Institute for Diagnostic, Interventional and Pediatric Radiology (DIPR), Inselspital, University Hospital of Bern, Switzerland, (N.K.).

Disclosures

None.

REFERENCES

1. Curry SJ, Krist AH, Owens DK, Barry MJ, Caughey AB, Davidson KW, Doubeni CA, Epling JW, Kemper AR, Kubik M, et al. Screening for osteoporosis to prevent fractures US Preventive Services Task Force recommendation statement. *JAMA*. 2018;319:2521–2531. DOI: 10.1001/jama.2018.7498.
2. Canton G, Hippe DS, Chen L, Waterton JC, Liu W, Watase H, Balu N, Sun J, Hatsukami TS, Yuan C. Atherosclerotic burden and remodeling patterns of the popliteal artery as detected in the magnetic resonance imaging Osteoarthritis Initiative dataset. *J Am Heart Assoc*. 2021;10:e018408. DOI: 10.1161/JAHA.120.018408.
3. Schneider E, NessAiver M, White D, Purdy D, Martin L, Fanella L, Davis D, Vignone M, Wu G, Gullapalli R. The Osteoarthritis Initiative (OAI) magnetic resonance imaging quality assurance methods and results. *Osteoarthritis Cartilage*. 2008;16:994–1004. DOI: 10.1016/j.joca.2008.02.010.
4. Glagov S, Weisenberg E, Zarins CK, Stankunavicius R, Kolettis GJ. Compensatory enlargement of human atherosclerotic coronary arteries. *N Engl J Med*. 1987;316:1371–1375. DOI: 10.1056/NEJM198705283162204.

Downloaded from http://ahajournals.org by on May 17, 2021