Radiology: Cardiothoracic Imaging COMMENTARY

Radiomics in Cardiac MRI: Sisyphean Struggle or Close to the Summit of Olympus?

Tim Leiner, MD, PhD

Tim Leiner, MD, PhD, EBCR, FSCMR, FSCCT, FISMRM, is a tenured professor of radiology and holds the chair in cardiovascular imaging at Utrecht University Medical Center, Utrecht, the Netherlands. He works on development and clinical implementation of advanced cardiovascular imaging techniques, including machine learning algorithms for earlier recognition and improved diagnosis of disease of the heart and vessels.

> *"Everything that can be counted, does not necessarily count; everything that counts, cannot necessarily be counted."*
>
> – Albert Einstein (1879–1955)

The putative aim of radiomics—sometimes also referred to as *texture analysis*—is to detect and quantify features in images, most often not visible to the human eye, that represent the underlying biology and pathophysiology. Quantitative analysis of gray-scale pixel values, as well as their arrangement, may hold important clues about the pathophysiologic processes taking place in living tissues and organs, and their study is expected to become an important adjunct to qualitative visual image interpretation in the diagnosis and prognosis of various diseases. To date, most of the radiomics studies have used CT images in patients with various types of malignancies as input. For example, investigators have found that radiomic features can be used with considerable success to differentiate benign from malignant pulmonary nodules and local recurrence from radiation therapy–induced damage (1). In contrast, relatively little work has been done in cardiac imaging, although encouraging studies on radiomics of atherosclerotic plaques (2) and pericoronary fat (3) have highlighted the potential of radiomics to obtain important and actionable insights about cardiovascular diseases. Even less has been published on radiomics in cardiac MRI, and important questions remain regarding if, when, and in which patients radiomic analyses add clinical value. One of the most important issues in this regard is the reproducibility of radiomics. Reproducibility is especially important in MRI because of the unique MR image acquisition and formation process, which involves hundreds of modifiable parameters, many of which influence image appearance. Human observers are highly skilled at filtering out the resulting small variations in signal intensity that influence image contrast, texture, and delineation of anatomic details. However, these same variations can lead to detection of spurious radiomics features that have no underlying biologic correlate. To what extent this hampers application of radiomics in clinical cardiac MRI is an important knowledge gap.

The study by Jang et al (4), featured in this issue of *Radiology: Cardiothoracic Imaging*, is an important attempt to address this problem. To better understand the reproducibility of radiomics of cardiac MRI, several insightful experiments were carried out with a 3-T MRI unit of a single vendor using multiple commonly used sequences in cardiac MRI such as cine imaging, T1- and T2-weighted imaging, as well as quantitative T1 and T2 mapping. A phantom consisting of four different types of fruit, as well as 10 healthy participants and 51 patients referred for cardiac MRI for various clinical indications, were studied. The investigators also performed test-retest experiments for the phantom as well as the healthy participants and patients. Healthy participants were reimaged within a 2-week period. Patients were reimaged in the same session after all clinically required sequences were performed. To investigate interobserver variation, data were analyzed by three readers. For each imaging sequence and time point, 1023 radiomics features were calculated by using the PyRadiomics package. Eleven image filters and six feature families were applied. Results were reported according to the reporting guidelines recommended by the Image Biomarker Standardization Initiative (IBSI). IBSI is an international collaboration that works toward better understanding the sources of variation in radiomics with the goal of working toward a standardized and validated radiomics toolbox for different imaging modalities (5). The main findings of the study were that only a small minority of cardiac MRI radiomic features are reproducible—typically less than approximately 10%–15%, especially between different scanning sessions—and that the exact set of features that can be considered reproducible varies by sequence. These findings are not surprising. It has long been known that MR image appearance can vary as a function of many variables encountered in the MRI acquisition and image reconstruction process. A nonexhaustive list of important factors, independent of the underlying pathophysiology, that are known to affect MR image appearance and therefore likely to influence the reproducibility of radiomics are *(a)* the specific brand and type of scanner, *(b)* the field strength, *(c)* the type of pulse sequence used, *(d)*

From the Department of Radiology, Utrecht University Medical Center, Heidelberglaan 100, 3584CX Utrecht, the Netherlands. Received April 22, 2020; revision requested May 6; revision received May 6; accepted May 6. **Address correspondence to** the author (e-mail: *t.leiner@umcutrecht.nl*).

See also the article by Jang et al in this issue. Conflicts of interest are listed at the end of this article.

Radiology: Cardiothoracic Imaging 2020; 2(3):e200244 • https://doi.org/10.1148/ryct.2020200244 • Content codes: CA IN MR

the k-space sampling strategy (eg, Cartesian vs radial sampling), *(e)* the specific k-space undersampling strategy (eg, sensitivity encoding vs compressed sensing vs fingerprinting or multitasking), *(f)* choice of spatial resolution, *(g)* image reconstruction settings, including any filters and interpolation schemes used to alter image appearance, and *(h)* presence of artifacts (eg, due to arrhythmia or susceptibility).

Furthermore, the specific radiomics software package used may influence the results. Other factors that are specific for cardiac MRI include manual versus machine learning–based contouring of the ventricular myocardium and whether contrast agent was administered.

The strengths of this study were the amount and detail of work aimed to investigate reproducibility in phantoms, healthy participants, and patients. Furthermore, the investigators are to be commended for making available both the data as well as the software code to other investigators. The obvious limitation of the study was that just one MRI system operating at a magnetic field strength of 3.0 T from a single vendor was studied. Also, because of the small sample size, a comparison of the value of radiomics between different disease entities was not possible. More work is therefore needed to understand the reproducibility of cardiac MRI radiomics at different field strengths, as well as with different types of MRI machines from different vendors and in different disease entities.

Does this mean that radiomics in cardiac MRI is a futile pursuit? Probably not. There is preliminary evidence that clinically relevant manifestations of disease can be detected with cardiac radiomics. Baessler et al demonstrated that radiomics can identify the presence and transmurality of myocardial scar from non–contrast-enhanced cine images with high accuracy (6), and that it is capable of detecting myocardial tissue alterations from native T1 maps in the setting of hypertrophic cardiomyopathy with excellent accuracy (7). These findings underscore that even if just a small percentage of radiomic features are reproducible, the technique can have a big clinical impact. On the other hand, efforts to standardize MRI acquisition, reconstruction, postprocessing, and analysis will be needed to bring radiomics to routine clinical cardiac MRI. Another crucial part of good radiomics practices is appropriate use of statistics to avoid overfitting and false-positive findings. Techniques such as cluster analysis and principal component analysis including advanced versions such as *t*-stochastic neighbor embedding (*t*-SNE) (8) are especially important to visualize high-dimensional data and to identify redundant features. In this regard, it is also important to underscore that current radiomics software packages contain large libraries of pre-engineered features. It will be interesting to see how deep learning with its capability of identifying structure in data without prespecifying the features to look for can be exploited to find additional clinically relevant patterns in radiomics data.

So, where are we now with radiomics in cardiac MRI and what is needed to bring it to clinical practice? The key to successful deployment is to identify canonical features of disease irrespective of the settings used above. In addition to following the IBSI recommendations, there are a number of additional ways to achieve this aim. First, it is important to make large data sets with known ground truth publicly available. This enables different groups of researchers to benchmark their results against one another. Conversely, it is of paramount importance that promising radiomic markers are validated in independent data sets that accurately represent the variety in image appearance one is likely to encounter across different machines, hospitals, and clinical settings. Finally, investigators should be encouraged to make the source code of their algorithms openly available. This not only promotes transparency about how an algorithm works, but it also allows other investigators to independently study how a particular algorithm performs in different settings.

In summary, Jang et al have performed an important study in the reproducibility of radiomics in cardiac MRI. This field is in its infancy, and their findings underscore the need for further standardization of cardiac MR image acquisition, postprocessing, and image analysis as well as the need for large publicly available data sets. There is a long climb ahead, but the view from the summit will be magnificent.

Disclosures of Conflicts of Interest: T.L. Activities related to the present article: disclosed no relevant relationships. Activities not related to the present article: institution receives grant from Dutch Technology Foundation (P15-26, 12726), Netherlands Heart Foundation (147141), with participation of Pie Medical Imaging and Philips Healthcare; institution receives payment for lectures from Bayer Healthcare and Philips Healthcare; coinventor, U.S. patent no. 10,395,366 (royalty agreement with Pie Medical; no royalties have been paid so far); royalties from Pie Medical; cofounder and shareholder in Quantib-U (no money paid to author or institution). Other relationships: disclosed no relevant relationships.

References

1. Rogers W, Thulasi Seetha S, Refaee TAG, et al. Radiomics: from qualitative to quantitative imaging. Br J Radiol 2020;93(1108):20190948.
2. Kolossváry M, Park J, Bang JI, et al. Identification of invasive and radionuclide imaging markers of coronary plaque vulnerability using radiomic analysis of coronary computed tomography angiography. Eur Heart J Cardiovasc Imaging 2019;20(11):1250–1258.
3. Oikonomou EK, Williams MC, Kotanidis CP, et al. A novel machine learning-derived radiotranscriptomic signature of perivascular fat improves cardiac risk prediction using coronary CT angiography. Eur Heart J 2019;40(43):3529–3543.
4. Jang J, Ngo LH, Mancio J, et al. Reproducibility of Segmentation-based Myocardial Radiomic Features with Cardiac MRI. Radiol Cardiothorac Imaging 2020;2(3):e190216.
5. Image Biomarker Standardisation Initiative Web site. https://ibsi.readthedocs.io/en/latest/index.html. Accessed April 20, 2020.
6. Baessler B, Mannil M, Oebel S, Maintz D, Alkadhi H, Manka R. Subacute and Chronic Left Ventricular Myocardial Scar: Accuracy of Texture Analysis on Nonenhanced Cine MR Images. Radiology 2018;286(1):103–112.
7. Baessler B, Mannil M, Maintz D, Alkadhi H, Manka R. Texture analysis and machine learning of non-contrast T1-weighted MR images in patients with hypertrophic cardiomyopathy-Preliminary results. Eur J Radiol 2018;102:61–67.
8. van der Maaten L, Hinton G. Visualizing data using t-SNE. J Mach Learn Res 2008;9:2579–2605. http://www.jmlr.org/papers/v9/vandermaaten08a.html.