Radiomics: Current Challenges in Clinical Validation

Faiq A. Shaikh, MD, University of Pittsburgh Medical Center; Omer Awan, MD; Christopher Deible, MD, PhD; Brian Kolowitz, MBA, DSc; Kenneth Hendrata, MBA

Background

Radiomics can be defined as the extraction and analysis of large amounts of advanced quantitative imaging features with high throughput from medical images obtained with various modalities [1]. Radiomics methods can be applied across various cancers to identify tumor phenotype characteristics in the images that correlate with their likelihood of survival, as well as their association with the underlying driving biology. Identifying this characteristic set of features called tumor signature, holds tremendous value in predicting cancer behavior and progression, which in turn has the potential to predict cancer’s response to various therapeutic options. In allowing us to have this capacity, radiomics holds the promise of driving the engine of precision medicine. However, there are numerous challenges in the validation methods needed to establish it as a clinically viable solution.

Evaluation

Traditional Picture Archiving and Communication System (PACS) platforms do not have the necessary capabilities to develop radiomics algorithms, nor do they have the ability to integrate the Radiomic outputs into the clinical workflows. Many actors (clinician researchers, data scientists, algorithm developers, etc.) must be able to collaborate with each other, and interact with multimodality datasets during the development of algorithms. Once developed, these outputs must seamlessly tie into the physician workflow which requires deep integration into exam worklists, imaging viewers, and the EMR. Today, PACS and the specialty Radiomic systems available lack the necessary integration points to advance the science and clinical practice. As PACS and EMR’s adopt newer standards including HL7, FHIR and DICOMweb, the ability to integrate with the systems becomes easier but a functionality gap remains for research, development, and clinical use.

Discussion

A disease may present similarly in different patient but have very different outcomes. Biopsy is limited in the information it provides as different parts of tumors have different molecular characteristics, which change over time. Personalized medicine aims to better predict individual patients’ outcomes, and select optimal treatments based on improved biomarker and radiogenomic data [2]. Radiomics reveals the tumor signature that defines its behavior and progression, which in turn directly determines the sensitivity profile to various therapeutic interventions [3].

Statistical challenges:

Given the novelty of these methods and the challenges associated with their application, there have been a limited number of studies performed performed to validate the performance of radiomics algorithms. These have been mostly retrospective studies, with a small sample size (50 - 150), yielding results with low statistical power and garnering low levels of confidence among physicians, who generally rely on major randomized studies to change practice behavior. Moreover, in order to apply neural network or machine learning algorithms, a large dataset is required for the regression models needed.
**Trial design:**
Challenges with using a bigger sample size for a radiomics study includes access to patient data (for a retrospective study) or patient selection (for a prospective study). The latter will be difficult to design given the issues with integration of radiomic application into the clinical workflow [4]. There are issues with recall bias when it comes to a retrospective study, but designing a randomized controlled prospective trial for quantitative image analysis is prohibitively difficult. There are issues surrounding patient selection (determining robust inclusion/exclusion criterion given that radiomics algorithms need large data sets to work on), challenges with blindfolding readers (unless a true fully automated radiomics application is running in the background, requiring no radiologist input), logistical challenges with data transfer (in a multicenter study), and perhaps most importantly, assessing outcomes.

**Outcomes Analysis, Survival Rates, & Clinical Follow-up**
Measuring outcomes for a major study in this domain can be difficult. While it is fairly straightforward to compare the machine’s performance with that of the radiologist, that may only be useful to assess the performance of a CAD software. Besides, if a machine/application is more sensitive at detecting lesions, it may be predisposed to over-calls, significantly reducing specificity and warranting unnecessary clinical workup, healthcare dollar waste and patient anxiety [5]. True outcomes analysis looking at survival rates and/or metrics of quality of life will require long follow-up periods. Ultimately, the most important question to ask is if a machine system helps in early detection of cancer, does it allow for early therapeutic intervention and whether that has been shown to impact patient survival rate. One potential benefit of a radiomics application could be better standardization of parameters used in treatment decisions across sites to make this larger data set gathering easier.

**Data transfer and deployment:**
Imaging data is sequestered for clinical consumption but radiomics applications require the whole raw dataset. Data transfer (for retrospective or prospective studies) is yet another challenge, requiring patient data transfer to a central repository, which involves HIPAA-compliance issues and requires heavy hardware and data bandwidth. Solutions based on onsite data collection, anonymization, sequestration and processing have been suggested by some companies working this space.

**The gold standard issue:**
Another important consideration in designing research studies for radiomics is that of choosing the gold standard. The limitations of pathology as the gold standard are well known - sampling errors, no longitudinal correlation, etc. In some cases, pathologic assessment is not warranted and therefore unavailable for comparison. Clinical follow-up in such cases becomes critical.

The new set of imaging biomarkers that radiomics offers need to be correlated with the serum or tissue biomarkers made available to us through digital pathology. A comparison needs to be made between these two kinds of biomarkers to assess their efficacy and impact in cancer prediction, assessment, stratification, therapy planning.

**Stratified analysis of Radiomics solutions:**
The results of radiomics are not binary - they are complex and stratified. What does slightly increased risk of malignant progression mean in terms of guiding management? What does a slightly better response prediction to a certain treatment mean? should it be administered? It is important to determine whether the statistical confidence is enough to stratify response to that information adequately. It can be argued that the results do not have to be binary to help - especially, with new information of the likelihood of better response based on imaging parameters, could we provide this to patients to help them make more appropriate ‘personalized’ decisions?
Conclusion

Overall, radiomics is a promising field that uses quantitative image analysis with the help of machine learning to glean useful insights into disease behavior, which then guides management in a highly precise and personalized manner. However, for it to be established as a reliable, reproducible and accurate clinical decision support system, robust statistical and outcomes analyses must be performed. This is an opportunity to scientifically address these challenges to find workable solutions that can allow us to tap into the true potential of Radiomics.

References


Keywords

radiomics; precision medicine; quantitative image analysis; challenges