

Radiomics in PET/CT: More Than Meets the Eye?

Mathieu Hatt¹, Florent Tixier², Dimitris Visvikis¹, and Catherine Cheze Le Rest²

¹LaTIM, INSERM, UMR 1101, University of Brest, IBSAM, Brest, France; and ²Academic Department of Nuclear Medicine, CHU Poitiers, Poitiers, France

Radiomics is defined as the high-throughput extraction of quantitative metrics from medical images (1). One of its main assumptions is that medical images are considered not merely pictures for visual assessment but rather minable quantitative data (2) that may not necessarily be captured by the human eye (3).

In this issue of *The Journal of Nuclear Medicine*, Orhac et al. present a study comparing visual assessment of uptake heterogeneity on PET images by experts and a subset of radiomics metrics, namely textural features (4). They exploited both clinical and simple simulated PET images, going further than previous studies performed using clinical data only (5–7). Such studies are useful because they provide additional understanding relative to the visual meaning of quantitative metrics that cannot easily be explained to nonspecialists. These studies have focused on the PET component and the ¹⁸F-FDG uptake heterogeneity. Similar analyses have been performed with CT (8) and MRI (9).

See page 387

One important finding is that textural features calculated after a relative quantization process (i.e., resampling the original image intensities into a variable number of bins of fixed width; e.g., 0.5 SUV (10)) correlate better with visual assessment than do those calculated after the usual quantization process (i.e., uniformly resampling the original intensities into a set number of bins; e.g., 64 or 128). These different observations can be related overall to different factors, such as the very different correlative relationships between texture parameters and either SUV_{max} or the number of voxels involved (tumor volume), which also have been previously reported (10–13). Other quantization processes (histogram equalization, Max–Lloyd clustering, and others) can lead to yet further differences in distribution and associated clinical value (14).

The consensus among experts was also substantially higher than in earlier studies, mostly because only 2 categories (heterogeneous vs. homogeneous) were considered, compared with 3 (5), 4 (7), or even 5 (8) in previous studies. In one study (5), the visual assessment

into 3 categories had limited prognostic value compared with textural features (5). Because there was no clinical endpoint (survival, outcome) in the study by Orhac et al., we cannot draw conclusions about the clinical value of the features that correlated well with visual assessment, although it is safe to assume that these features will be useful in clinical applications for which there is a correlation between patient outcome and the level of uptake heterogeneity visually assessed (or SUV_{max}, given the observed correlations).

The primary goal of radiomics is to build clinical models using machine-learning techniques (15) to predict patient outcome, thereby allowing better patient management. These multiparametric models, which are likely to be unintelligible even to experts because they combine a large number of high-order multimodality image features (13,16), should outperform visual analysis in terms of both accuracy and reproducibility. To associate a visual meaning to such models can be even more challenging because they can also incorporate information from other fields (demographics, histopathology, genomics). The human brain can take into account only a limited number of parameters in making a decision; therefore, these multiparametric models will not be easily apprehended by end users. These models will clearly demand a high level of precision and robustness in order to be accepted and relied on to formulate a clinical decision. Within this context, a rigorous process of model development (proper training) and validation (independent large cohorts) is needed, which is still far from being a standard, although some encouraging results have been published (17,18).

The current radiomics paradigm consists in adding quantitative information to the visual analysis by radiologists and nuclear medicine physicians, rather than replacing it entirely. For instance, it was recently shown that a set of semantic features obtained from visual assessment by radiologists could beneficially complement quantitative radiomics in determining epidermal growth factor receptor mutations in lung cancer (19). However, a recent trend in medical imaging is to exploit techniques from the field of deep learning (20), with examples in image segmentation (21) or radiomics-type studies (22). This will further complicate the issue of association with visual analysis. Indeed, on the one hand the standard radiomics workflow relies on the extraction of carefully designed features based on domain expertise (e.g., a specific calculation in the intensity histogram or in a predesigned texture matrix), some of which are clearly inspired by the human visual system. On the other hand, deep learning methods automatically discover features from data and the representations useful for the task at hand using a general-purpose learning procedure such as convolutional neural networks. These require substantial amounts of data not easily available in the field of medical imaging, particularly in PET/CT. Potential solutions include transfer learning, consisting in using convolutional neural networks trained for an unrelated task using large datasets, and adapting them to a

Received Oct. 4, 2016; revision accepted Oct. 11, 2016.
For correspondence or reprints contact: Mathieu Hatt, LaTIM, INSERM, UMR 1101, CHRU Morvan, 2 avenue Foch, 29609 Brest Cedex, France.
E-mail: hatt@univ-brest.fr
Published online Nov. 3, 2016.
COPYRIGHT © 2017 by the Society of Nuclear Medicine and Molecular Imaging.
DOI: 10.2967/jnumed.116.184655

different setting (23,24). If these tools were to advantageously replace the current workflow of radiomics, removing the need for tumor segmentation or the complex task of selecting relevant and reliable features (11,25), as well as improving the ability to handle standardization issues (26), the relationship with visual analysis by experts would not simply be more difficult but certainly unnecessary to establish.

DISCLOSURE

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

REFERENCES

1. Lambin P, Rios-Velazquez E, Leijenaar R, et al. Radiomics: extracting more information from medical images using advanced feature analysis. *Eur J Cancer*. 2012;48:441–446.
2. Gillies RJ, Kinahan PE, Hricak H. Radiomics: images are more than pictures, they are data. *Radiology*. 2016;278:563–577.
3. Aerts H. Radiomics: there is more than meets the eye in medical imaging [abstract]. In: *Proc SPIE*. 2016;9785:978500.
4. Orhac F, Nioche C, Soussan M, Buvat I. Understanding changes in tumor texture indices in PET: a comparison between visual assessment and index values in simulated and patient data. *J Nucl Med*. 2017;58:387–392.
5. Tixier F, Hatt M, Valla C, et al. Visual versus quantitative assessment of intratumor ¹⁸F-FDG PET uptake heterogeneity: prognostic value in non-small cell lung cancer. *J Nucl Med*. 2014;55:1235–1241.
6. Kim D-H, Jung J-H, Son SH, et al. Quantification of intratumoral metabolic macroheterogeneity on ¹⁸F-FDG PET/CT and its prognostic significance in pathologic N0 squamous cell lung carcinoma. *Clin Nucl Med*. 2016;41:e70–e75.
7. Watabe T, Tatsumi M, Watabe H, et al. Intratumoral heterogeneity of F-18 FDG uptake differentiates between gastrointestinal stromal tumors and abdominal malignant lymphomas on PET/CT. *Ann Nucl Med*. 2012;26:222–227.
8. Hodgdon T, McInnes MDF, Schieda N, Flood TA, Lamb L, Thornhill RE. Can quantitative CT texture analysis be used to differentiate fat-poor renal angiomyolipoma from renal cell carcinoma on unenhanced CT images? *Radiology*. 2015;276:787–796.
9. Meinel LA, Stolpen AH, Berbaum KS, Fajardo LL, Reinhardt JM. Breast MRI lesion classification: improved performance of human readers with a backpropagation neural network computer-aided diagnosis (CAD) system. *J Magn Reson Imaging*. 2007;25:89–95.
10. Leijenaar RTH, Nalbantov G, Carvalho S, et al. The effect of SUV discretization in quantitative FDG-PET radiomics: the need for standardized methodology in tumor texture analysis. *Sci Rep*. 2015;5:11075.
11. Hatt M, Tixier F, Pierce L, Kinahan P, Cheze Le Rest C, Visvikis D. Characterization of PET images using texture analysis: the past, the present... any future? *Eur J Nucl Med Mol Imaging*. June 6, 2016 [Epub ahead of print].
12. Orhac F, Soussan M, Chouahnia K, Martinod E, Buvat I. ¹⁸F-FDG PET-derived textural indices reflect tissue-specific uptake pattern in non-small cell lung cancer. *PLoS One*. 2015;10:e0145063.
13. Desserot M-C, Visvikis D, Tixier F, et al. Development of a nomogram combining clinical staging with ¹⁸F-FDG PET/CT image features in non-small cell lung cancer stage I-III. *Eur J Nucl Med Mol Imaging*. 2016;43:1477–1485.
14. Desserot M-C, Tixier F, Cheze Le Rest C, et al. Comparison of three quantization methods for the calculation of textural features in PET/CT images: impact on prognostic models in non-small cell lung cancer. Presented at: IEEE NSS-MIC; Strasbourg, France; October 29–November 5, 2016.
15. Parmar C, Grossmann P, Bussink J, Lambin P, Aerts HJWL. Machine learning methods for quantitative radiomic biomarkers. *Sci Rep*. 2015;5:13087.
16. Upadhaya T, Morvan Y, Stindel E, Le Reste P-J, Hatt M. A framework for multimodal imaging-based prognostic model building: preliminary study on multimodal MRI in glioblastoma multiforme. *IRBM*. 2015;36:345–350.
17. Aerts HJWL, Velazquez ER, Leijenaar RTH, et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. *Nat Commun*. 2014;5:4006.
18. Desserot M-C, Tixier F, Guillemin R, et al. Nomogram for NSCLC exploiting clinical staging, tumor volume and PET/CT heterogeneity features: development using support vector machines in a retrospective cohort and first validation results in prospectively recruited patients [abstract]. *J Nucl Med*. 2016;57 (suppl 2):437.
19. Velazquez ER, Liu Y, Parmar C, Narayan V, Gillies R, Aerts H. MO-DE-207B-08: radiomic CT features complement semantic annotations to predict EGFR mutations in lung adenocarcinomas [abstract]. *Med Phys*. 2016;43:3706.
20. LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature*. 2015;521:436–444.
21. Havaei M, Davy A, Warde-Farley D, et al. Brain tumor segmentation with deep neural networks. *Med Image Anal*. 2016;35:18–31.
22. Ypsilantis P-P, Siddique M, Sohn H-M, et al. Predicting response to neoadjuvant chemotherapy with PET imaging using convolutional neural networks. *PLoS One*. 2015;10:e0137036.
23. Antropova N, Huynh B, Giger M. SU-D-207B-06: Predicting breast cancer malignancy on DCE-MRI data using pre-trained convolutional neural networks [abstract]. *Med Phys*. 2016;43:3349–3350.
24. Huynh B, Drukker K, Giger M. MO-DE-207B-06: Computer-aided diagnosis of breast ultrasound images using transfer learning from deep convolutional neural networks [abstract]. *Med Phys*. 2016;43:3705.
25. Lian C, Ruan S, Denœux T, Jardin F, Vera P. Selecting radiomic features from FDG-PET images for cancer treatment outcome prediction. *Med Image Anal*. 2016;32:257–268.
26. Lasnon C, Majdoub M, Lavigne B, et al. ¹⁸F-FDG PET/CT heterogeneity quantification through textural features in the era of harmonisation programs: a focus on lung cancer. *Eur J Nucl Med Mol Imaging*. 2016;43:2324–2335.