EBioMedicine 47 (2019) 20–21

Contents lists available at ScienceDirect

EBioMedicine

journal homepage: www.ebiomedicine.com

EBioMedicine
Published by THE LANCET

Commentary

Multicentric validation of radiomics findings: challenges and opportunities

Mathieu Hatt [a,*], François Lucia [a,b], Ulrike Schick [a,b], Dimitris Visvikis [a]

[a] LaTIM, INSERM, UMR 1101, Univ Brest, Brest, France
[b] Radiotherapy Department, CHRU Brest, Brest, France

Predicting patients' outcome after treatment is challenging. Usual clinical variables and standard exploitation of imaging fails to deliver actionable predictive models with sufficient accuracy in a number of cancer types and associated treatments, including cervical cancer, one of the most frequent malignant tumour in women with over 500,000 cases annually worldwide, associated with 270,000 deaths [1]. For women diagnosed with locally advanced cervical cancer (LACC), treatment in Western countries is usually pelvic external beam radiotherapy (EBRT) in association with cisplatin-based chemotherapy (CRT), followed by brachytherapy (BT). Alternatively, preoperative neoadjuvant chemotherapy (NACT) is increasingly being considered in Asian countries in order to reduce tumour volumes and allow for a surgical removal.

Within this context, there is a crucial need to identify biomarkers predictive of outcome, in order to personalize treatment. Radiomics has emerged as a potential solution to provide such biomarkers from available routine pre-treatment images, by considering that the content of images can be extracted through machine (deep) learning, beyond the capabilities of the human eye, even the expertly-trained one [2,3]. The resulting features can be mined, similarly to other -omics domains, in order to identify the relevant ones. One advantage of radiomics is that it exploits diagnostic images that are available already, so it does not require additional exams (imaging or biological).

Despite a number of encouraging results, radiomics faces challenges that have until now prevented its widespread use in clinical practice. These include false-positives [4], the lack of standards and resulting limited reproducibility [5], the lack of fully automated detection and delineation, and the lack of multi-centric validation [6]. This last point is addressed within the context of LACC by the study by Sun, et al. recently published in *EBioMedicine* [7]. They investigated the prediction of response to NACT in LACC using radiomic features extracted from pre-treatment magnetic resonance imaging (MRI) T1 and T2 sequences. They included 275 patients from 8 cohorts (5 were used as a training set, 3 as the testing set). IBSI (image biomarkers standardization initiative)-compliant radiomic features were extracted from the two MRIs, as well as from both the intratumoural and the peritumoural regions manually delineated by radiologists. In the training set, a support vector machine with recursive feature elimination was exploited to identify the most relevant features. Then a random forest was used to build models, which were then evaluated in the testing set. Importantly, the robustness of the findings was evaluated by training/testing the models on different training and testing sets. The models based on a single sequence and single tumour region reached AUCs ≥ 0.94, with 100% specificity but limited sensitivity (76% at best). Those based on combinations of sequences reached AUCs ≥ 0.98, with 100% specificity and sensitivity up to 84% for the best model. Moreover, these results were robust to different splitting of the data into training and testing sets. In a subset of 232 patients for which the clinical variables (age, FIGO stage, gross type) were available, the radiomic models significantly outperformed the model based on clinical features (AUC of 0.6). This is in line with the results of Lucia, et al. with the radiomic model reaching an accuracy of 0.90–0.97 versus 0.56–0.60 for the clinical model [8].

One challenging aspect of multicentric radiomic studies is the sensitivity of radiomic features with respect to the variability of scanner models, acquisition protocols and reconstruction algorithms and parameters. Recently, a radiomic model exploiting FDG positron emission tomography (PET) and MRI apparent coefficient diffusion (ADC) map to predict outcome in LACC was validated in a multicentric context (3 different centers), thanks to the use of the ComBat harmonization technique [8]. In that respect, the study of Sun, et al. showed that z-score normalization [9] of each slice of the MR images intensities as well as of the features of each patient could allow good performance, despite different 3.0-T and 1.5-T MRI scanner models from several different vendors across the 8 centres.

One important limitation of the study by Sun, et al. is that regions of interest were manually delineated. Although reproducibility was evaluated through blind analysis by two radiologists, a fully automated delineation will need to be developed so the proposed radiomic model can be exploited clinically. Artificial intelligence has already shown the potential of highly accurate image segmentation within the multimodality imaging context [10]. Another limitation is that clinical data were not available for all patients, compromising the comparison between the radiomic models and standard clinical variables carried out only on a subsample of the overall study population. Finally, the models might not be directly applicable to cohorts in the Western countries for which treatment management differ.

DOI of original article: https://doi.org/10.1016/j.ebiom.2019.07.049.
* Corresponding author.
E-mail address: hatt@univ-brest.fr (M. Hatt).

https://doi.org/10.1016/j.ebiom.2019.08.054

Nonetheless, the work by Sun, et al. contributes to the growing number of studies that show validating radiomic models in a multicentric context is feasible. Such results may help the radiomics approach to translate faster into the clinical routine practice.

Author disclosure

The authors declare no conflicts of interest.

References

[1] Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2018;68:394–424.

[2] Visvikis D, Cheze Le Rest C, Jaouen V, Hatt M. Artificial intelligence, machine (deep) learning and radio(geno)mics: definitions and nuclear medicine imaging applications. Eur J Nucl Med Mol Imaging 2019 Jul 6. https://doi.org/10.1007/s00259-019-04373-w in press, Epub ahead of print.

[3] Hatt M, Tixier F, Visvikis D, Cheze Le Rest C. Radiomics in PET/CT: more than meets the eye? J Nucl Med 2017;58:365–6.

[4] Welch ML, McIntosh C, Haibe-Kains B, Milosevic MF, Wee L, Dekker A, et al. Vulnerabilities of radiomic signature development: the need for safeguards. Radiother Oncol J Eur Soc Ther Radiol Oncol 2019;130:2–9.

[5] Vallières M, Zwanenburg A, Badic B, Cheze Le Rest C, Visvikis D, Hatt M. Responsible radiomics research for faster clinical translation. J Nucl Med Off Publ Soc Nucl Med 2018;59:189–93.

[6] Upadhaya T, Vallières M, Chatterjee A, Lucia F, Bonaffini PA, Masson I, et al. Comparison of radiomics models built through machine learning in a multicentric context with independent testing: identical data, similar algorithms, different methodologies. IEEE Trans Radiat Plasma Med Sci 2019;3:192–200.

[7] Sun C, Tian X, Liu Z, Li W, Li P, Chen J, et al. Radiomic analysis for pretreatment prediction of response to neoadjuvant chemotherapy in locally advanced cervical cancer: a multicentre study. EBioMedicine 2019;46:160–9.

[8] Lucia F, Visvikis D, Vallières M, Desseroit M-C, Miranda O, Robin P, et al. External validation of a combined PET and MRI radiomics model for prediction of recurrence in cervical cancer patients treated with chemoradiotherapy. Eur J Nucl Med Mol Imaging 2019;46:864–77.

[9] Chatterjee A, Vallières M, Dohan A, Levesque IR, Ueno Y, Saif S, et al. Creating robust predictive radiomic models for data from independent institutions using normalization. IEEE Trans Radiat Plasma Med Sci March 2019;3(2):210–5. https://doi.org/10.1109/TRPMS.2019.2893860.

[10] Guo Z, Li X, Huang H, Guo N, Li Q. Deep learning-based image segmentation on multimodal medical imaging. IEEE Trans Radiat Plasma Med Sci 2019;3:162–9.