Importance of measurement repeatability of semi-quantitative imaging through PET-CT and PET-MR imaging in oncology

Adriano Massimiliano Priola, Sandro Massimo Priola

Department of Diagnostic Imaging, San Luigi Gonzaga University Hospital, Orbassano 10043, Torino, Italy
Correspondence to: Adriano Massimiliano Priola. Department of Diagnostic Imaging, San Luigi Gonzaga University Hospital, Regione Gonzole 10, Orbassano 10043, Torino, Italy. Email: adriano.priola@inwind.it.
Comment on: Fraum TJ, Fowler KJ, Crandall JP, *et al.* Measurement repeatability of 18F-FDG-PET/CT versus 18F-FDG-PET/MRI in solid tumors of the pelvis. J Nucl Med 2019;60:1080-6.

Positron emission tomography (PET)/computed tomography (CT), magnetic resonance (MR) imaging, and the hybrid PET/MR scanner of recent introduction have seen rapid expansion in both research and clinical application (1-3). At present, they are increasingly used in clinical practice, especially in the oncologic setting for tumour detection and diagnosis, for histological subtype correlation and stratification of disease, for staging, for predicting and monitoring treatment response during chemotherapy and at the end of treatments, for prognosis on disease-free and overall survival, and for follow-up of tumours, in different organs and compartments (2-7). These modern techniques have the advantage of providing semi-quantitative and quantitative data, in addition to morphological assessment obtained by CT or conventional MR imaging, which are obtained through the measurement of a lot of parameters, including the apparent diffusion coefficient (ADC) for diffusion-weighted (DW)/MR and the standardized uptake value (SUV) for PET/CT or PET/MR (8-10). The ADC obtained by using a mono-compartmental model is the most common quantitative method used in DW-MR imaging (2,11). It is inversely correlated to the tissue cellularity and the presence of barriers that reduce or impede the movement of the water molecules within the intracellular and extracellular compartments (2,3,11). Thus, malignant tumours usually present low ADC values because of the decrease in the extracellular and intracellular space due to the high cellularity and the high nucleus-cytoplasm ratio which have a limiting effect on water molecule diffusion (3,12). Furthermore, a significant increase in ADC values during treatment or at the end of chemotherapy demonstrates a positive response to treatments (2,13). In nuclear medicine, the SUV is the ratio of tissue radioactivity concentration at a specific time and administered dose at the time of injection divided by body weight. Beyond the body weight, the administered dose may also be corrected by the lean body mass (SUL), or body surface area. In addition, in fluoro-deoxy-glucose (FDG) studies, SUV can be corrected for plasma glucose level, because glucose transporters may be saturated by glucose (14). Generally, SUV is calculated on the highest image pixel in the tumour regions (SUVmax) because of its lower inter-observer variability compared to averaged SUV (SUVmean). In addition, other volume-based parameters such as metabolic tumour volume (MTV) and total lesion glycolysis (TLG) are also used since it has been proved that they may provide better prognostic indices than traditional SUV metrics (15). Nevertheless, SUV is yet used as the predominant metric for tumour quantification with ^{18}F-FDG PET. Glycolysis is increased in metabolically active tumours and in inflamed tissues. Hence, an increase in SUV values is expected in malignant tumours and decrease in SUV values reflects favourable response to treatments (1).

In studies that use quantitative or semi-quantitative imaging techniques with the aim to find accurate value thresholds to characterize tumours, the assessment of repeatability of measurements is fundamental for validating their use in clinical practice (2,3). Unfortunately, SUV and ADC are associated with multiple sources of variability that include type of scanner, parameters of acquisition, and method of measurement. Furthermore, the lack of standardization of the technical parameters and the

acquisition protocols between different studies makes it impossible to compare results or even utilize proposed thresholds in everyday clinical practice (2,3). In DW-MR imaging, multiple factors affect quantitative ADC values, leading to unreliable measurements, such as field strength, repetition-time, echo-time, *b*-values selection, post-processing system and data analysis (2,3,16,17). As an example, the inclusion of low *b* values, which control the sensitivity to diffusion-based contrast in DW-MR, in mono-exponential ADC maps of DW-MR imaging, causes an ADC overestimation because of the inclusion of perfusion effects of tissues that are completely attenuated at *b* values greater than 100 sec/mm^2 (2). In the characterization of the anterior mediastinum, the use of perfusion-free ADC measurements (obtained by using *b* values of 150,500, and 800 sec/mm^2 for reconstructing ADC maps) significantly improved diagnostic accuracy of DW-MR imaging when compared to perfusion-sensitive measurements (obtained by including also the *b* value of 0 sec/mm^2 for reconstructing ADC maps) with perfusion effects which were greater for malignancies than benign conditions (18). In addition, the manual method of region-of-interest (ROI) positioning for measuring ADC may affect the ADC value and can lead to unreliable measurements (2). For instance, in breast cancer, Giannotti and colleagues recently observed that inter-reader variability is reduced by using larger ROIs to cover the entire lesion, as compared to a single small ROI (19). Beyond size, for the assessment of tumours of the anterior mediastinum, we demonstrated that repeatability of ADC measurements is also related to number of ROIs. Indeed, the whole tumour volume method of ADC measurement was found the most repeatable, suggesting that inclusion of the entire tumour volume leads to a decrease in inter-observer variation because this method, unlike other methods, does not require the reader to select the slice or the area with the lowest signal on the ADC map (20). Similar results were found for other tumours including pleural abnormalities, rectal cancer, endometrial cancer, and musculoskeletal lesions (4,17,21).

Thus, to best use SUV and ADC in clinical practice, an understanding of the repeatability of the technique is required. Test-retest studies involve repeated scanning of the same patient on the same scanner using the same protocol no more than a few days apart and provide basic information on the repeatability of the technique (22). The study of Fraum and colleagues faces this important issue for PET/CT and PET/MR imaging hybrid techniques (23). In particular, the authors investigated the test-retest repeatability of quantitative metrics obtained by using PET/MR imaging, including SUVmax, SULpeak, and ADCmedian, in patients with solid malignancies of the pelvis (23). Past studies demonstrated that metrics of PET/CT were highly repeatable, whereas measurement repeatability of PET/MR imaging metrics have been less investigated (22). The advantage of PET/MR imaging compared to PET/CT lies in its higher accuracy in local and systemic staging due to its superior contrast resolution compared to CT for morphological assessment (8,9,23). In their study, Fraum and colleagues prospectively enrolled 14 patients with newly diagnosed pelvic tumours (mostly cervical squamous cell carcinoma accounting for 79% of the whole cohort, with remaining cases of endometrial adenocarcinoma), except for one patient who had a pelvic recurrence of rectal adenocarcinoma (23). All these patients, who did not receive any oncologic treatment within 30 days prior to enrolment, underwent two imaging sessions of PET/CT and PET/MR separated each other from 1 to 7 days. On PET/CT imaging and on ADC map of DW imaging for PET/MR, the margin of the tumour was manually drawn in order to generate the whole tumour contour for obtaining several PET metrics and ADC values. For PET/MR, the ADC repeatability analysis was performed in 12 out of 14 patients enrolled. By comparing PET metrics obtained with PET/CT and PET/MR imaging, PET/MR presented lower metrics compared to PET/CT, which were significantly lower for 7 out of 12 metrics in the primary analysis and for 12 out of 72 metrics in the exploratory analysis, except for the metabolic tumour volume. The test-retest repeatability analysis demonstrated high repeatability of ADC measurements performed in the two defined sessions with identical values of ADCmedian (within-subject coefficient of variation, 2.4%; mean percent differences, 1.6%) and similar values of ADCmean, ADC20, and ADCtrough. For PET/CT, although the within-subject coefficient of variation was found in a narrow range (from –9.8% to 6.0%) for all PET metrics, the same coefficient was lower for ADC metrics compared to PET ones. Lastly, the mean percent differences in measurements between sessions were generally lower on PET/MR compared to PET/CT though without statistical significance.

The major strength of the present study lies in the investigation of DW imaging repeatability on simultaneous PET/MR, a topic that was not previously analyzed because two previous studies on PET/MR evaluated only repeatability of PET metrics (24,25). Overall, beyond PET metrics, the study of Fraum and colleagues showed

that ADC metrics are also quantitatively robust for their inclusion in clinical trials (23).

However, authors are aware of the limits of their results which mainly lie in the small study cohort that have precluded to obtain statistical significance differences between defined groups. In addition to other limitations listed by authors, a potential major limitation of this study lies in the lack of information about acquisition of data for extracting quantitative values of DW imaging for PET/MR. As an example, the authors declared that "ADC maps were generated by the VA40 console software" and in the supplemental table 1 a three-point b-value acquisition is reported with values of 50, 500, and 1,000 sec/mm^2. However, if ADC maps were reconstructed by including the low b value of 50 sec/mm^2, perfusion sensitive values were obtained with ADC overestimation due to perfusion effects within the tumour (2,3,18). Furthermore, it is likely that, in the study of Fraum and colleagues, quantitative measurements were manually extracted from the lesion of interest from the same reader for all cases, whereas measurements should be performed at least by two readers in order to assess inter-reader variability of quantitative or semi-quantitative data.

In conclusion, the study of Fraum and colleagues add valuable and useful information in the field of hybrid quantitative imaging about repeatability of PET metrics and DW metrics, underlying the robustness of such metrics for their utilization in clinical trials with quantitative end-point and for their incorporation into treatment response assessment algorithms.

Acknowledgments

Footnote

Provenance and Peer Review: This was commissioned and reviewed by the Section Editor Hengwei Liu, MD, PhD (Department of Obstetrics and Gynecology, Union Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China).

Conflicts of Interest: Both authors have completed the ICMJE uniform disclosure form (available at http://dx.doi.org/10.21037/tcr.2019.05.15). The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

References

1. Singnurkar A, Poon R, Metser U. Comparison of 18F-FDG-PET/CT and 18F-FDG-PET/MR imaging in oncology: a systematic review. Ann Nucl Med 2017;31:366-78.
2. Padhani AR, Liu G, Koh DM, et al. Diffusion-weighted magnetic resonance imaging as a cancer biomarker: consensus and recommendations. Neoplasia 2009;11:102-25.
3. Priola AM, Gned D, Veltri A, et al. Chemical shift and diffusion-weighted magnetic resonance imaging of the anterior mediastinum in oncology: current clinical applications in qualitative and quantitative assessment. Crit Rev Oncol Hematol 2016;98:335-57.
4. Nguyen TL, Soyer P, Fornès P, et al. Diffusion-weighted MR imaging of the rectum: clinical applications. Crit Rev Oncol Hematol 2014;92:279-95.
5. Priola AM, Priola SM. Primary mediastinal Hodgkin lymphoma and rebound thymic hyperplasia: differentiation with chemical-shift magnetic resonance imaging after treatment. Int J Hematol 2009;90:8-10.
6. Taghipour M, Marcus C, Sheikhbahaei S, et al. Clinical indications and impact on management: fourth and subsequent posttherapy follow-up 18F-FDG PET/CT scans in oncology patients. J Nucl Med 2017;58:737-43.
7. Priola AM, Priola SM, Gned D, et al. Nonsuppressing normal thymus on chemical-shift MR imaging and anterior mediastinal lymphoma: differentiation with diffusion-weighted MR imaging by using the apparent diffusion coefficient. Eur Radiol 2018;28:1427-37.
8. Priola AM, Priola SM, Gned D, et al. Comparison of CT and chemical-shift MRI for differentiating thymoma from non-thymomatous conditions in myasthenia gravis: value

of qualitative and quantitative assessment. Clin Radiol 2016;71:e157-69.

9. Priola AM, Priola SM. Imaging of thymus in myasthenia gravis: from thymic hyperplasia to thymic tumor. Clin Radiol 2014;69:e230-45.

10. Priola AM, Galetto G, Priola SM. Diagnostic and functional imaging of thymic and mediastinal involvement in lymphoproliferative disorders. Clin Imaging 2014;38:771-84.

11. Priola AM, Priola SM, Giraudo MT, et al. Chemical-shift and diffusion-weighted magnetic resonance imaging of thymus in myasthenia gravis: usefulness of quantitative assessment. Invest Radiol 2015;50:228-38.

12. Priola AM, Priola SM, Giraudo MT, et al. Diffusion-weighted magnetic resonance imaging of thymoma: ability of the apparent diffusion coefficient in predicting the World Health Organization (WHO) classification and the Masaoka-Koga staging system and its prognostic significance on disease-free survival. Eur Radiol 2016;26:2126-38.

13. Jalaguier-Coudray A, Villard-Mahjoub R, Delouche A, et al. Value of dynamic contrast-enhanced and diffusion-weighted MR imaging in the detection of pathologic complete response in cervical cancer after neoadjuvant therapy: a retrospective observational study. Radiology 2017;284:432-42.

14. Chauvie S, Bergesio F. The strategies to homogenize PET/CT metrics: the case of onco-haematologic clinical trials. Biomedicine 2016;15:e26-46.

15. Vallius T, Hynninen J, Kemppainen J, et al. 18F-FDG-PET/CT based total metabolic tumor volume change during neoadjuvant chemotherapy predicts outcome in advanced epithelial ovarian cancer. Eur J Nucl Med Mol Imaging 2018;45:1224-32.

16. Schmidt H, Gatidis S, Schwenzen NF, et al. Impact of measurement parameters on apparent diffusion coefficient quantification in diffusion-weighted magnetic resonance imaging. Invest Radiol 2015;50:46-56.

17. Inoue C, Fujii S, Kaneda S, et al. Apparent diffusion coefficient (ADC) measurement in endometrial carcinoma: effect of region of interest methods on ADC values. J Magn Reson Imaging 2014;40:157-61.

18. Priola AM, Priola SM, Gned D, et al. Diffusion-weighted quantitative MRI to diagnose benign conditions from malignancies of the anterior mediastinum: improvement of diagnostic accuracy by comparing perfusion-free to perfusion-sensitive measurements of the apparent diffusion coefficient. J Magn Reson Imaging 2016;44:758-69.

19. Giannotti E, Waugh S, Priba L, et al. Assessment and quantification of sources of variability in breast apparent diffusion coefficient (ADC) measurements ad diffusion weighted imaging. Eur J Radiol 2015;84:1729-36.

20. Priola AM, Priola SM, Parlatano D, et al. Apparent diffusion coefficient measurements in diffusion-weighted magnetic resonance imaging of the anterior mediastinum: inter-observer reproducibility of five different methods of region-of-interest positioning. Eur Radiol 2017;27:1386-94.

21. Priola AM, Priola SM, Gned D, et al. Diffusion-weighted quantitative MRI of pleural abnormalities: intra- and interobserver variability in the apparent diffusion coefficient measurements. J Magn Reson Imaging 2017;46:769-82.

22. Lodge MA. Repeatability of SUV in oncologic 18F-FDG PET. J Nucl Med 2017;58:523-32.

23. Fraum TJ, Fowler KJ, Crandall JP, et al. Measurement repeatability of 18F-FDG-PET/CT versus 18F-FDG-PET/MRI in solid tumors of the pelvis. J Nucl Med 2019;60:1080-6.

24. Rasmussen JH, Fischer BM, Aznar MC, et al. Reproducibility of 18F-FDG PET uptake measurements in head and neck squamous cell carcinoma on both PET/CT and PET/MR. Br J Radiol 2015;88:20140655.

25. Groshar D, Bernstine H, Goldberg N, et al. Reproducibility and repeatability of same-day two sequential FDG PET/MR and PET/CT. Cancer Imaging 2017;17:11.