



Global disease score (GDS) is the name of the game!

Poul F. Høilund-Carlsen^{1,2} · Lars Edenbrandt³ · Abass Alavi⁴

Received: 23 May 2019 / Accepted: 28 May 2019 / Published online: 10 June 2019
© The Author(s) 2019

For some decennia, CT and MRI have with beautiful images predominated medical in vivo imaging. However, because of their limited sensitivity and structural nature, they have not provided quite the key to improved understanding of disease or better solution to clinical problems that their impressive images promised that they could. Therefore, it is time for molecular imaging in the shape of PET/CT and PET/MRI to utilize the unique advantages of PET and hybrid imaging to accomplish these very important challenges. This requires easily accessible and reliable quantification procedures, which can assess the entire disease burden in the body and provide a simple score expressing the extent and activity of disease within minutes. In this editorial, we argue that the *global disease score (GDS)* is this measure, and must replace the maximal standardized value (SUVmax), which is an easily accessible, but often misleading, oversimplification.

PET was conceived in the late 1950s and came to birth in the 1970s. So did CT and MRI. They shared the same tomographic principle, but because of their higher spatial resolution, CT and MRI for some time put PET on the sideline, until a change began with the emergence of hybrid PET/CT at the turn of the millennium and was reinforced with the advent of PET/MRI approximately 10 years later. The wonderfully sharp images one can achieve with CT and MRI have helped

countless patients and thousands of doctors around the world to form an impression of what was wrong and what to rectify. However, the enthusiasm has gradually been damped by the realization that these modalities are hampered by non-negligible limitations, meaning that the tissue changes they depict are typically late-occurring events that seldom appear until cure is no longer an option or lesions no longer represent active disease. Together with their suboptimal sensitivity, these are the main reasons why CT and MRI have not delivered what they were supposed to in the understanding and management of a number of serious disorders including cancer and cerebral, cardiovascular, and musculoskeletal diseases. Thus, it is time for molecular imaging with PET to leave the sideline and step into play.

Besides being molecular and possessing a thousand times higher sensitivity than CT and MRI, PET is an inherently quantitative modality. Achieving optimal utilization of this potential is not straightforward and is something that has still not reached a stage where measurements at different institutions are directly comparable. For PET to become a clinical success in many more diseases, it is necessary to replace easily accessible measures like SUVmax, SUVpeak, and SULmax [1] with conceptually more correct indices of disease. By representing tracer uptake in only a few of the body's diseased cells, these values are not true indicators of any disease. They disregard the amount, extension, and activity of disease and the fact that many, if not most, diseases are heterogeneous. However, to replace them with other, more comprehensive and more accurate quantities is a challenge, as it requires great computing power and fast, user-friendly, and reproducible computer programs effectively providing proper segmentation and correction for background uptake and partial volume effect, of which the latter plays a major role especially in small lesions [2, 3]. We suggest instead a single number, the GDS, which expresses the amount and extent of disease and its activity. The concept was introduced already in 1993 for FDG-PET imaging of the brain [4], and can be used for the entire body, tumor tissue, or specific organs such as the heart [5–7]. Semi-automated programs enabling the calculation of GDS and other parameters have for some years been used to

This article is part of the Topical Collection on Advanced Image Analyses (Radiomics and Artificial Intelligence)

✉ Poul F. Høilund-Carlsen
pfhc@rsyd.dk

¹ Department of Nuclear Medicine, Odense University Hospital, 5000 Odense C, Denmark

² Research Unit of Clinical Physiology and Nuclear Medicine, Department of Clinical Research, University of Southern Denmark, Odense, Denmark

³ Region Västra Götaland, Sahlgrenska University Hospital, Department of Clinical Physiology, Gothenburg, Sweden

⁴ Department of Radiology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA

study dependence on age, gender, and body mass index, for instance, or potential correlation of PET measures with risk factors [8–13]. However, programs are still too time-consuming and somewhat operator-dependent, making them awkward in the daily routine. With the advent of artificial intelligence-based and deep learning procedures, this is about to rapidly change [14–16].

PET is not a 100% accurate and reliable modality. Its main limitation is a spatial resolution of half a centimeter or more even with modern PET/CT scanners. This cannot be fully counterbalanced by the high sensitivity of PET. It means that a negative PET/CT scan cannot rule out ongoing cancer. However, while a negative FDG PET scan in a patient with a cancer above the size of the spatial resolution of PET is routinely interpreted as false-negative in terms of tumor detectability, it should in fact be interpreted as true-negative in terms of tumor biology, since tumors with low or no FDG uptake are histologically and clinically non-aggressive [17]. Moreover, the high sensitivity of PET and its molecular nature ensures disease detection much earlier than structural imaging, since molecular disease precedes functional disturbances, which in turn appear before structural tissue changes become detectable (Fig. 1). Thus, by its very nature, PET will become positive and demonstrate ongoing disease long before structural imaging will show late consequences of disease that may or may not represent active disease [18].

We are not aware of studies that have examined these challenges in depth. Two recent publications in the *Journal of Nuclear Medicine (JNM)* and *Radiology* based on the same material comprising 34 patients with newly diagnosed high-grade, resectable osteosarcoma illustrate the problem. They describe FDG PET/CT measures obtained at baseline and

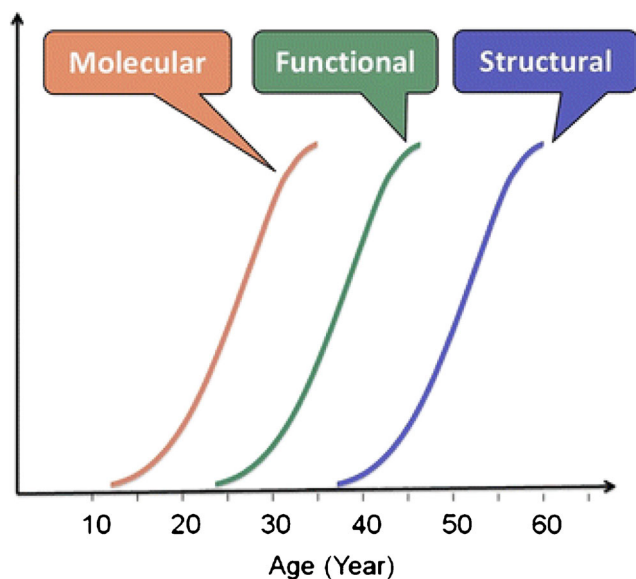


Fig. 1 Anticipated temporal relationship between molecular calcification, functional changes, and CT-visible arterial calcification. Reproduced with permission [13]

following 5-week and 10-week neoadjuvant chemotherapy as predictors of histologic response, event-free survival, and overall survival [19, 20]. The *JNM* report focuses on SUVmax, whereas the *Radiology* article concentrates on SUVpeak, metabolic tumor volume, and total lesion glycolysis, i.e., the product of SUVmean and metabolic tumor volume. In the *JNM* paper, the authors found that SUVmax at 5 weeks and 10 weeks and percentage change in SUVmax from baseline to 10 weeks were highly predictive of histologic response, whereas there was no association with event-free survival, which was not associated with either change or percentage change in SUVmax [19]. In the *Radiology* article, the authors found that SUVpeak, metabolic tumor volume, and total lesion glycolysis measured at baseline were predictors of event-free survival and overall survival, and that at 5 weeks and 10 weeks, all parameters including percent change were associated with histologic response and predictive of event-free survival [20]. By its definition, SUVmax cannot represent an entire tumor or the total burden of cancer in the body. Consequently, it is no surprise that parameters derived instead from the whole primary tumor and not a tiny fraction of it were predictive of event-free and overall survival and were more reliable endpoints of response to chemotherapy than histology, which was only borderline significantly predictive [20].

Because SUVmax represents the uptake in just a single voxel and is influenced by noise, movement artefacts, and other sources of error, it is not a relevant measure of a disease and its development. Like many other cancers, osteosarcoma very heterogeneous [21]. As expressed by Lindsey et al. it is “extremely heterogeneous in both its origins and manifestations” and “continues rapidly modifying its genotype, thus making potential targeted molecular therapeutics increasingly impractical” [22]. For the same reason, more specific PET tracers than FDG for cancer imaging are a rare sight. In the *Radiology* report, metabolic tumor volume was as predictive as total lesion glycolysis, perhaps partly due to the chosen tumor type (high-risk and surgically removable), which could mean less heterogeneity than if all histological degrees were included. Thus, it is fair to assume that total tumor glycolysis, or rather the GDS, which is the weighted average uptake in all lesions, is a truer indicator of the amount and aggressiveness of the tumor burden than just the tumor volume, the SUVmax, or the diameter measures that are still retained in the Response Evaluation Criteria in Solid Tumors (RECIST) criteria [23]. Heterogeneity exists and changes spontaneously within the tumor volume, in particular when chemotherapy and other interventions kill or subdue part of a primary tumor. The Positron Emission Tomography (PET) Response Criteria in Solid Tumors (PERCIST) criteria was a step in the right direction by shifting from a geometric to an uptake

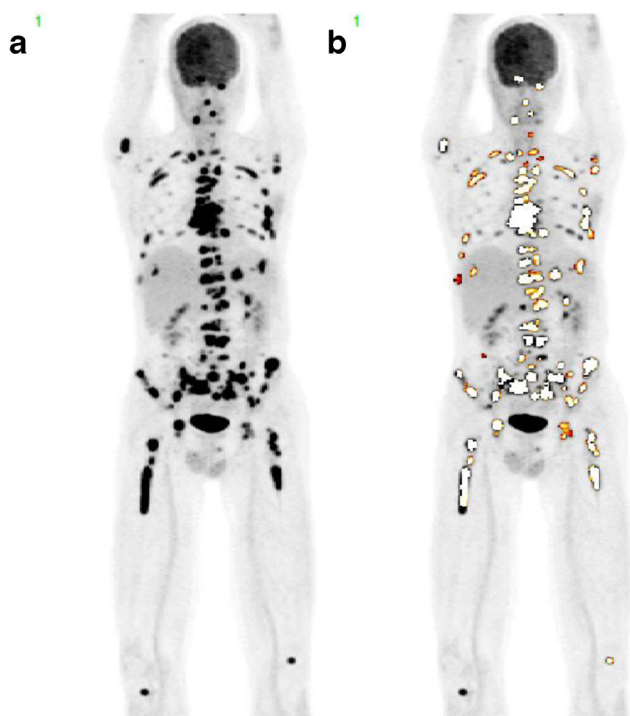


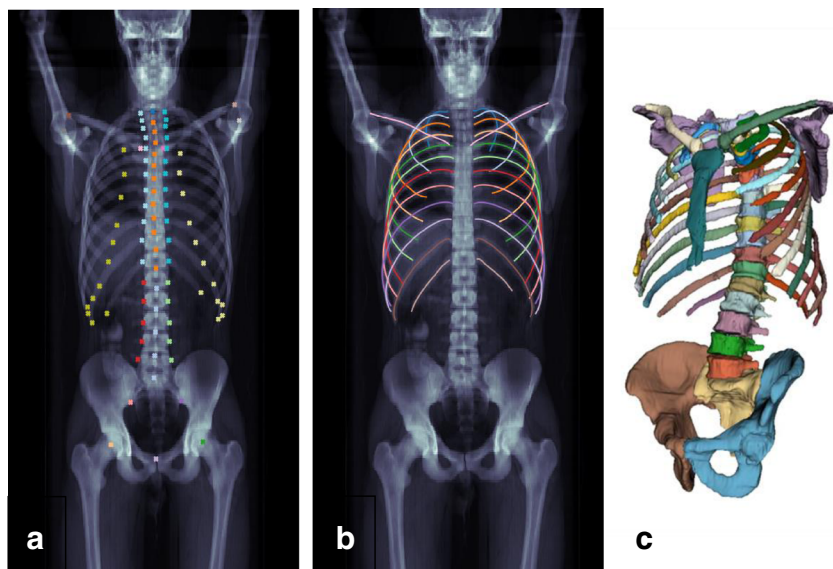
Fig. 2 **a** FDG-PET scan of a patient with multiple myeloma and osseous and extra-osseous lesions. **b** Lesions were analyzed using dedicated semi-quantitative software. Segmentation was done using a voxel-based iterative adaptive algorithm starting at 40% threshold of SUV_{max}. Note how difficult it is to determine which lesions should be included in the analysis and how difficult it might be for different observers to achieve the same result. Artificial intelligence-based analysis overcomes these challenges. Courtesy Dr. Brian Østergaard, Department of Hematology, Odense University Hospital, Odense, Denmark

measure [24]. Unfortunately, it suggests measurement of SUL_{peak} in the hottest lesion and potentially also in five selected lesions, which are the same from time to time. Since these will typically change and thus retain

their characteristics for only a limited period due to ongoing genetic and phenotypic changes, the risk of misinterpretation increases with the rate of these changes.

Only the entire burden of disease, its extent, and activity can truthfully characterize and monitor the disease and its response to treatment. That is exactly what is expressed by the GDS. To what extent this score should be supplemented with measurements of the rate of change in disease activity is a question to be highlighted in future research. The rationale behind double-time-point PET imaging using for instance 1- and 3-h acquisition is readily understandable, because this will reveal the degree of tumor aggressiveness [25]. In practice, however, this approach is cumbersome, and its ability to characterize the disease stage of individual patients may be somewhat limited [26, 27]. Another option is trend analysis based on repeat PET acquisitions, except that this requires three or more acquisitions, i.e., one at baseline and two or more during follow-up [28]. In addition, good reproducibility is absolutely necessary, since all involved measurement errors come into play twice when the measurement is repeated [29]. Manual methods for quantifying total disease volume and activity are too time-consuming and hampered by low reproducibility (Fig. 2). The use of modern artificial intelligence-based methods for single and repeat GDS measurements makes sense, because these can be obtained in a completely automated and operator-independent way within minutes or seconds (Fig. 3), which means that they can be applied and their performance validated much more rigorously. Importantly, the AI-based approach allows for the first time for comparison of GDS results obtained at different institutions with different vendor-specific scanners, reconstruction algorithms, and processing procedures, and that is exactly what is required in the clinical context and for research purposes. Objectivity and

Fig. 3 **a** Maximum-intensity projection of a CT scan with annotated landmarks. **b** Detected center lines for ribs and clavicles. **c** Surface reconstruction of the resulting segmentation enabling calculation of the skeletal volume in relation to which the sum of the metabolically active volume of bone lesions is expressed. Reproduced with permission [15]



reproducibility are critical according to the FDA biomarker qualification review team, which states that the clinical association of a quantitative biomarker is significant only if the marker is measured consistently under pre-defined conditions [30], which again is a prerequisite for performing analytical and clinical validation studies such as those conducted recently for the Bone Scan Index [31].

In conclusion, the time has come to abandon illogical and misleading PET imaging measures representing but a fraction of vastly heterogeneous diseases in the body, in favor of ultra-fast, operator-independent, highly reproducible AI-based measurements of the GDS, which is a much truer expression of the body's disease load that is constantly undergoing spontaneous or therapy-induced changes.

Compliance with ethical standards

Conflict of interest None.

Ethical approval This article does not contain any studies with human participants or animals performed by any of the authors.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Boellaard R, Delgado-Bolton R, Oyen WJ, Giammarile F, Tatsch K, Eschner W, et al. FDG PET/CT: EANM procedure guidelines for tumour imaging: version 2.0. *Eur J Nucl Med Mol Imaging*. 2015;42(2):328–54.
- Alavi A, Werner TJ, Høilund-Carlsen PF, Zaidi H. Correction for partial volume effect is a must, not a luxury, to fully exploit the potential of quantitative PET imaging in clinical oncology. *Mol Imaging Biol*. 2018;20:1–3.
- Chawluk JB, Alavi A, Dann R, Hurtig HI, Bais S, Kushner MJ, et al. Positron emission tomography in aging and dementia: effect of cerebral atrophy. *J Nucl Med*. 1987;28(4):431–7.
- Alavi A, Newberg AB, Souder E, Berlin JA. Quantitative analysis of PET and MRI data in normal aging and Alzheimer's disease: atrophy weighted total brain metabolism and absolute whole brain metabolism as reliable discriminators. *J Nucl Med*. 1993;34(10):1681–7.
- Basu S, Zaidi H, Salavati A, Hess S, Carlsen PF, Alavi A. FDG PET/CT methodology for evaluation of treatment response in lymphoma: from “graded visual analysis” and “semiquantitative SUVmax” to global disease burden assessment. *Eur J Nucl Med Mol Imaging*. 2014;41:2158–60.
- Alavi A, Werner TJ, Høilund-Carlsen PF. What can be and what cannot be accomplished with PET: rectifying ongoing misconceptions. *Clin Nucl Med*. 2017;42:602–5.
- Alavi A, Werner TJ, Høilund-Carlsen PF. What can be and what cannot be accomplished with PET to detect and characterize atherosclerotic plaques. *J Nucl Cardiol*. 2018;25:2012–5.
- Basu S, Zaidi H, Houseni M, Bural G, Udupa J, Acton P, et al. Novel quantitative techniques for assessing regional and global function and structure based on modern imaging modalities: implications for normal variation, aging and diseased states. *Semin Nucl Med*. 2007;37:223–39.
- Basu S, Saboury B, Torigian DA, Alavi A. Current evidence base of FDG-PET/CT imaging in the clinical management of malignant pleural mesothelioma: emerging significance of image segmentation and global disease assessment. *Mol Imaging Biol*. 2011;13:801–11.
- Beheshti M, Saboury B, Mehta NN, Torigian DA, Werner T, Mohler E, et al. Detection and global quantification of cardiovascular molecular calcification by fluoro-18-fluoride positron emission tomography/computed tomography—a novel concept. *Hell J Nucl Med*. 2011;14:114–20.
- Blomberg BA, de Jong PA, Thomassen A, Lam MG, Vach W, Olsen MH, et al. Thoracic aorta calcification but not inflammation is associated with increased cardiovascular disease risk: results of the CAMONA study. *Eur J Nucl Med Mol Imaging*. 2017;44:249–58.
- McKenney-Drake ML, Moghbel MC, Paydary K, Alloosh M, Houshmand S, Moe S, et al. 18F-NaF and 18F-FDG as molecular probes in the evaluation of atherosclerosis. *Eur J Nucl Med Mol Imaging*. 2018;45:2190–200.
- Moghbel M, Al-Zaghal A, Werner TJ, Constantinescu CM, Høilund-Carlsen PF, Alavi A. The role of PET in evaluating atherosclerosis: a critical review. *Semin Nucl Med*. 2018;48:488–97.
- Lindgren Belal S, Sadik M, Kaboteh R, Hasani N, Enqvist O, Svärm L, et al. 3D skeletal uptake of 18F sodium fluoride in PET/CT images is associated with overall survival in patients with prostate cancer. *EJNMMI Res*. 2017;7(1):15. <https://doi.org/10.1186/s13550-017-0264-5>.
- Lindgren Belal S, Sadik M, Kaboteh R, Enqvist O, Ulén J, Poulsen MH, et al. Deep learning for segmentation of 49 segmented bones in CT scan: first step in automated PET/CT-based 3D quantification of skeletal metastases. *Eur J Radiol*. 2019;113:89–95.
- Mortensen MA, Borrelli P, Poulsen MH, Gerke O, Enqvist O, Ulén J, et al. Artificial intelligence-based versus manual assessment of prostate cancer in the prostate gland: a method comparison study [revision submitted].
- Kwee TC, Basu S, Saboury S, Ambrosini V, Torigian DA, Alavi A. A new dimension of FDG-PET interpretation: assessment of tumor biology. *Eur J Nucl Med Mol Imaging*. 2011;38:1158–70.
- Høilund-Carlsen PF, Hess S, Werner TJ, Alavi A. Cancer metastasizes to the bone marrow and not to the bone: time for a paradigm shift! *Eur J Nucl Med Mol Imaging*. 2018;45:893–7.
- Davis JC, Daw NC, Navid F, Billups CA, Wu J, Bahrami A, et al. 18F-FDG uptake during early adjuvant chemotherapy predicts histologic response in pediatric and young adult patients with osteosarcoma. *J Nucl Med*. 2018;59:25–30.
- Im HJ, Zhang Y, Wu H, Wu J, Daw NC, Navid F, et al. Prognostic value of metabolic and volumetric parameters of FDG PET in pediatric osteosarcoma: a hypothesis-generating study. *Radiology*. 2018;287:303–12.
- Schiavone K, Garnier D, Heymann MF, Heymann D. The Heterogeneity of Osteosarcoma: The role played by cancer stem cells. *Adv Exp Med Biol*. 2019;1139:187–200.
- Lindsey BA, Markel JE, Kleinerman ES. Osteosarcoma overview. *Rheumatol Ther*. 2017;4:25–43.
- Schwartz LH, Litière S, de Vries E, Ford R, Gwyther S, Mandrekar S, et al. RECIST 1.1 – update and clarification: from the RECIST committee. *Eur J Cancer*. 2016;62:132–7.

24. O JH, Lodge MA, Wahl RL. Practical PERCIST: a simplified guide to PET response criteria in solid tumors 1.0. *Radiology*. 2016;280:576–84.
25. Basu S, Kung J, Houseni M, Zhuang H, Tidmarsh GF, Alavi A. Temporal profile of fluorodeoxyglucose uptake in malignant lesions and normal organs over extended time periods in patients with lung carcinoma: implications for its utilization in assessing malignant lesions. *Q J Nucl Med Mol Imaging*. 2009;53(1):9–16.
26. Christlieb SB, Strandholdt CN, Olsen BB, Mylam KJ, Larsen TS, Nielbse AL, et al. Dual time point FDG PET/CT and FDG uptake and related enzymes in lymphadenopathies: preliminary results. *Eur J Nucl Med Mol Imaging*. 2016;43(10):1824–36.
27. Mortensen MA, Vilstrup MH, Gerke O, Høilund-Carlsen PF, Lund L. A prospective study on dual time 18F-FDG PET/CT in high-risk prostate cancer patients. *BMC Res Notes*. 2018;11(1):871. <https://doi.org/10.1186/s13104-018-3985-2>.
28. Gerke O, Ehlers K, Motschall E, Høilund-Carlsen PF, Vach W. PET/CT-based response evaluation in cancer – a systematic review of design issues. *Mol Imaging Biol*. 2019. <https://doi.org/10.1007/s11307-019-01351-4>.
29. Høilund-Carlsen PF, Lauritzen SL, Marving J, Rasmussen S, Hesse B, Folke K, et al. The reliability of measuring left ventricular ejection fraction by radionuclide cardiography: evaluation by the method of variance components. *Br Heart J*. 1988;59:653–62.
30. Biomarkers Definitions Working Group. Biomarkers and surrogate endpoints: preferred definitions and conceptual framework. *Clin Pharmacol Ther*. 2001;69:89–95.
31. Armstrong AJ, Anand A, Edenbrandt L, Bondesson E, Bjartell A, Widmark A, et al. Phase 3 assessment of the automated bone scan index as a prognostic imaging biomarker of overall survival in men with metastatic castration-resistant prostate cancer: a secondary analysis of a randomized clinical trial. *JAMA Oncol*. 2018;4:944–51.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.