



Quantitative imaging biomarkers in nuclear medicine: from SUV to image mining studies. Highlights from annals of nuclear medicine 2018

Martina Sollini^{1,2} · Francesco Bandera^{3,4} · Margarita Kirienko²

Received: 30 August 2019 / Accepted: 10 September 2019 / Published online: 5 November 2019
© Springer-Verlag GmbH Germany, part of Springer Nature 2019

Abstract

Objective Quantification in medical imaging is one of the main goals in research and clinical practice since it allows immediate understanding, objective communication, and comparison. Our aim was to summarize relevant investigations on quantification in nuclear medicine studies published in the volume 32 of *Annals of Nuclear Medicine*.

Methods In this article, we summarized the data of 14 selected papers from international research groups that were published between January and December 2018. This is a descriptive review with an inherently subjective selection of articles.

Results We discussed the role of parameters ranging from standardized uptake value to ratios, to flow within a region of interest, to volumetric parameters and to texture indices in different clinical scenarios in oncology, cardiology, and neurology.

Conclusions In all the medical disciplines in which nuclear medicine examinations play a role, quantification is essential both in research and in clinical practice. Standardization and high-quality protocols are crucial for the success and reliability of imaging biomarkers.

Keywords Standardized uptake value · Quantification · Radiomics · Artificial intelligence · Volumetric parameters

Introduction

In 2016, the *European Journal of Nuclear Medicine and Molecular Imaging* (EJNMMI) and *Annals of Nuclear*

Medicine (ANM) embarked on a joint initiative to publish highlights from studies of high interest for both European and Japanese readers and to strengthen the scientific cooperation between Europe and Japan. Since then, insightful papers have been published annually within this collaborative project [1–4]. These contributions have addressed impactful topics in preclinical studies, translational research, and clinical applications.

Currently, nuclear medicine research is focused mainly on new radiopharmaceuticals targeting specific biological processes, on the search for reliable imaging modalities for diagnosis, prognostication, and treatment response evaluation, and on innovative approaches to image and treat patients in a personalized manner. In particular, interest in image-derived biomarkers and advanced image analysis is growing, and this warrants a review designed to describe the key findings from a range of relevant investigations on quantification.

The use of any type of measurements and/or of related mathematics falls into the category of quantification [5]. Imaging is one of the domains that mostly require quantification [6]. In fact, image interpretation is somewhat operator dependent since it is influenced by personal experience, skills, and expertise, as well as by geographical

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

✉ Margarita Kirienko
margarita.kirienko@icloud.com

Martina Sollini
martina.sollini@cancercenter.humanitas.it

Francesco Bandera
francescobandera@gmail.com

¹ Nuclear Medicine, Humanitas Clinical and Research Center–IRCCS, Rozzano, Milan, Italy

² Department of Biomedical Sciences, Humanitas University, Pieve Emanuele, Milan, Italy

³ Department of Biomedical Sciences for Health, University of Milan, Milan, Italy

⁴ Department of Cardiology University, IRCCS Policlinico San Donato, San Donato Milanese, Milan, Italy

location [6]. Therefore, quantification is a necessity to increase consistency in image assessment and thereby guide medical decisions appropriately. Quantification in the digital age implies easier and automatic tools for image analysis, a myriad of biomarkers potentially derivable from any type of image (from pictures to medical images), and, consequently, a strong basis for personalized treatment decisions.

Biomarkers are defined as “measurable and quantifiable biological parameters (...) which serve as indices for health- and physiology-related assessments” [7]. The main advantage of the use of biomarkers lies in the fact that they are, or should be, easier and cheaper than direct measurement of the clinical endpoint [8]. Three crucial aspects—high reproducibility, a sizeable signal-to-noise ratio, and a swift change in response to events (e.g., treatment)—make a biomarker good [8]. In this regard, biomarkers derived from images have the potential to be among the most suitable.

In nuclear medicine, quantification—addressing the kinetics of a radiopharmaceutical—provides functional information for both diagnostic (e.g., split function of lungs or kidneys) and therapeutic (e.g., dosimetric estimates of radiation dose delivery) purposes [9]. Rather simple quantification methods are based on the assessment of tracer counts in one single pixel/voxel. Comparison of these counts with counts in a reference organ or on the contralateral side (ratio) is a common approach in several nuclear medicine investigations. Models based on fully quantitative methods, fitting dynamic data, are characterized by higher complexity. Assessment of tracer accumulation in either a region of interest (ROI) or the entire target volume allows the measurement of whole-lesion functional indices. Advanced image analysis (radiomics and artificial intelligence) has recently been introduced. Radiomics aims to extract numerous indices through mathematical methods with increasing levels of complexity to describe the tracer uptake heterogeneity as a function of the relationship between pixels or voxels within an ROI [10]. Artificial intelligence (AI) relies on the application of algorithms that identify relevant features in an image in a data-driven fashion. Figure 1 shows different types of quantitative approach in the imaging field.

The aim of the present review, within the joint EJNMMI/ANM initiative, is to provide an overview on quantification research in nuclear medicine focused on imaging biomarkers with an increasing degree of complexity: from SUV to radiomics. We subjectively selected the 14 most relevant investigations on imaging quantification from among all articles published in the *Annals of Nuclear Medicine* volume 32 (2018). We chose studies on the role of image-derived parameters in different clinical scenarios in oncology [11–20], cardiology [21, 22], and neurology [23, 24].

Quantitative imaging parameters

One pixel/voxel

The most common (semi)-quantitative parameter used in clinical practice for the assessment of positron emission tomography (PET) images is the standardized uptake value (SUV) and, particularly, the maximum SUV (SUV_{max}). SUV_{max} provides a measurement of the tracer uptake within the single pixel/voxel with the maximum decay counts among those belonging to a target region or volume, normalized to the injected activity [25]. SUV_{max} quantitation is affected by various patient- and image-related factors, including blood glucose level, concomitant medications, artifacts related to metal implants and devices, positive contrast, lesion size, scanner calibration, and acquisition time [25–30]. Abdel Gawad et al. proposed a 3D mathematical formulation of the SUV recovery coefficients. This formulation accounted for the most influential factors in SUV quantitation, including lesion size (volume or diameter) and lesion contrast (ratio of lesion concentration to background radiation level). Good performance of the proposed mathematical models has been reported in both phantom and clinical studies. In view of these preliminary results, the employed formulations are expected to allow more reproducible SUV measurements [11].

The increasing clinical application of PET/magnetic resonance (MR) has led to comparative studies aiming inter alia to assess the reproducibility of metrics between PET/CT and PET/MR systems. In this regard, Ringheim et al. [12] prospectively assessed the reproducibility of SUV_{max} in 30 prostate cancer patients with biochemical failure imaged by sequential ^{68}Ga -PSMA-11 PET/CT and PET/MR. They found good agreement, with a linear correlation between the SUV_{max} of same-day randomized ^{68}Ga -PSMA-11 PET/CT and PET/MR scans. However, the Bland-Altman analysis estimated a mean percentage difference of 18.7% (i.e., the PET/CT SUV_{max} was on average about 20% higher than the PET/MR SUV_{max}). Interestingly, based on their findings, they were able to propose a formula to compute PET/CT SUV_{max} starting from PET/MR SUV_{max} and vice versa: $Y(\text{PET/CT } SUV_{max}) = 0.75 + 1.00 \times (\text{PET/MR } SUV_{max})$. Overall, their results indicated that SUV_{max} from ^{68}Ga -PSMA-11 PET/CT and PET/MR should not be used interchangeably in either patient management or follow-up.

Ratios

The target-to-background ratio is a simple method to normalize tracer uptake within an ROI that can be used in planar scintigraphy, single-photon emission computed tomography (SPECT), and PET imaging. Particularly in the case of PET, the above-mentioned shortcomings of SUV have resulted in attempts to search for alternative, more valuable options. The

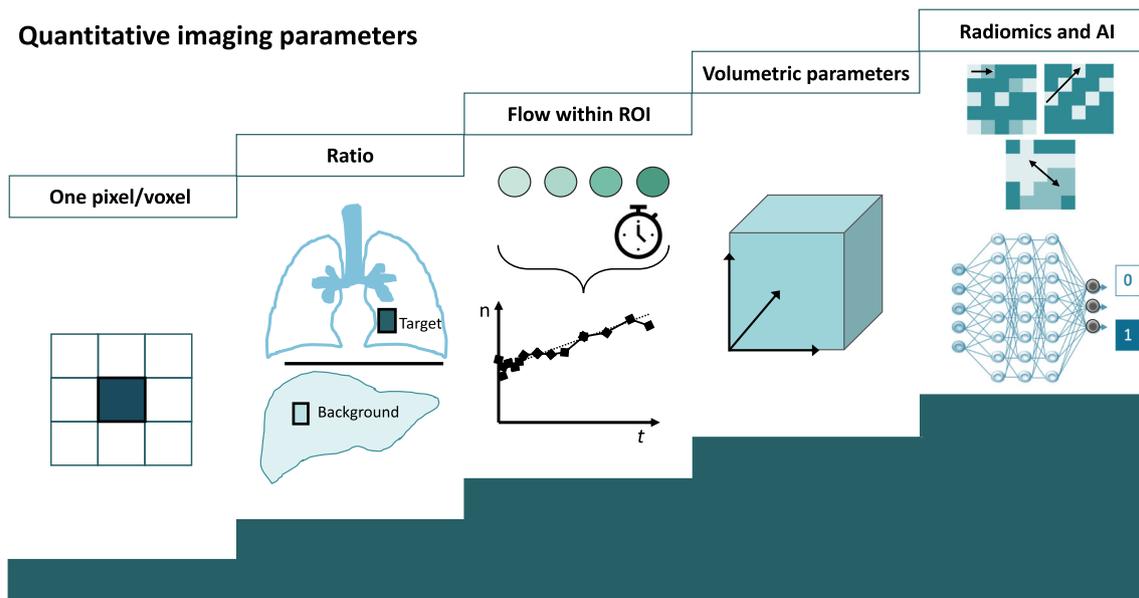


Fig. 1 Types of quantitative approaches in imaging

target-to-reference tissue ratio (TBR) is a simple (semi)-quantitative method to normalize the SUV_{max} . The liver, spleen, muscle, and blood have been proposed as reference organs. In the case of $[^{18}F]FDG$ studies, the tumor-to-blood standard uptake ratio (SUR) may be used as a surrogate for the metabolic rate. Moreover, a better correlation has been reported between the SUR and the metabolic rate of $[^{18}F]FDG$ compared with SUV, i.e., SUR is superior to SUV [30, 31]. The TBR is not dependent on tracer plasma clearance but, like SUV, it requires an adequate time for image acquisition. However, assuming the kinetics of $[^{18}F]FDG$ in the reference tissue to be irreversible and the shape of the arterial input function to be constant, both SUV and TBR may be accurately corrected for scan time variability [32].

The correlation between the target-to-liver ratio (TLR) and SUR has been investigated in order to assess whether the liver can be assumed to be a surrogate of arterial radiotracer uptake. However, the less than perfect results demonstrated the unsuitability of the liver as a surrogate of arterial tracer supply for SUV normalization via TLR computation [33]. Nevertheless, the prognostic role of SUR and TLR has been described in both solid tumors [33–36] and Hodgkin lymphoma [37]. Interestingly, Annunziata et al. [13] found that TLR exhibited a prognostic role also in follicular lymphoma. The authors compared the TLR with the five-point Deauville scale and to the International Harmonization Project criteria in 89 patients with follicular lymphoma treated with induction immuno-chemotherapy. The TLR and the five-point Deauville scale were found to be comparable in predicting progression-free survival at 5 years (PPV = 80% for both, NPV = 88% and 86%, respectively), with a better performance than the International Harmonization Project criteria. Accordingly, the TLR calculated using end-of-treatment

$[^{18}F]FDG$ PET/CT might be used as a prognostic index in follicular lymphoma, with the advantage over the five-point Deauville scale that it is independent from the subjective judgment of the observer [13].

As mentioned above, TBR may be computed with any radiopharmaceutical, including ^{68}Ga -PSMA, as shown by Komek et al. [20]. These authors provided evidence of the possibility of using SUV ratio cut-off values to identify metastases with poor or favorable outcomes in patients with advanced prostate cancer [20].

Additionally, TBR may be used for neuroreceptor quantification [38] and in patients with cardiovascular diseases [39]. Habert et al. [23] aimed to develop and validate a method for the quantification of amyloid burden on the basis of ^{18}F -florbetapir PET scans. They calculated in 53 cases (26 normal elderly controls, 11 patients with mild cognitive impairment, and 16 patients with Alzheimer's disease) the TBR in ROIs of the native space of the untransformed PET images, applying correction for the partial volume effect. They identified a positivity threshold and thereafter, using an in-house software to correlate it with other approaches, they validated the cut-off. The method was applied to an independent cohort of 318 cognitively normal subjects to assess the amyloid status, with positive results. This method may be useful in clinical practice, especially when the optimal cut-off cannot be obtained from healthy subjects [23].

Flow within ROIs

Dynamic image acquisition, analysis of multiple blood and/or urine samples, and extensive protocols for image and data post-processing are extremely accurate even if time consuming. Indeed, in PET, the net input constant (K_i) computed from

dynamic images applying Patlak analysis [40] better describes the [^{18}F]FDG metabolic rate than the surrogate endpoint of the metabolic rate of glucose consumption, SUR. Interestingly, Lebasnier et al. [21] evaluated the role of cardiac dynamic [^{18}F]FDG PET/CT in patients with suspicious cardiac sarcoidosis. They analyzed images of 28 patients, computing both global and segmental K_i . The heterogeneity of glucose metabolism was estimated by the normalized coefficient of variation and used to differentiate patients with and patients without cardiac sarcoidosis, with promising results (AUC = 0.96). Despite the retrospective study design, the paper benefitted from some positive aspects, including patient selection. In fact, it should be mentioned that the physiological myocardial uptake of [^{18}F]FDG requires specific protocols for patient preparation and image analysis [41]. Accordingly, two patients who did not meet the minimum criteria for adequate preparation were excluded from the analysis. Moreover, the combined evaluation of [^{18}F]FDG and myocardial perfusion images, which is the recommended radionuclide method for the evaluation of cardiac sarcoidosis [41], was worthy. Lastly, whole-body PET/CT images provided additional findings in more than half of the patients [21]. These results could have an important impact on patient management.

Volumetric parameters

The volumetric parameters used in PET/CT—metabolic tumor volume (MTV) and the total lesion glycolysis (TLG)—have been proposed for prognostication, radiation oncology planning, and treatment response assessment in a variety of tumors imaged by [^{18}F]FDG PET/CT. MTV and TLG represent the metabolically active burden and the level of glucose uptake within the total metabolically active volume, respectively. In the case of non-[^{18}F]FDG tracers, other definitions are used depending on the biological process traced by the radiopharmaceutical (e.g., functional volume, proliferative volume). Anwar et al. [15] retrospectively tested the prognostic role of baseline SUV_{max} , MTV, and TLG in 49 stage I NSCLC patients. All patients underwent staging [^{18}F]FDG PET/CT prior to any treatment, followed by complete tumor surgical resection (i.e., R0). Both 1- and 3-year disease-free survival were shorter in patients with higher MTV or TLG. MTV showed a higher prognostic value compared with TLG and SUV_{max} (AUC = 0.82, AUC = 0.79, and AUC = 0.72, respectively). Accordingly, baseline volumetric PET parameters may be useful in stage I NSCLC pre-treatment risk stratification, identifying patients who could benefit from wider resection and postoperative adjuvant therapy. These findings warrant further investigations since, if confirmed in independent cohorts, they might influence clinical practice.

Similarly, findings with a potential clinical impact were reported by Yoo et al. [17] in invasive ductal breast cancer. The authors analyzed baseline [^{18}F]FDG PET/CT images to

test whether volumetric parameters extracted within the primary tumor were predictors of axillary lymph node status (i.e., metastatic or not) in 135 patients with clinically negative axillary lymph nodes. All patients received breast surgery (breast-conserving surgery or mastectomy) with sentinel lymph node(s) biopsy and/or axillary lymph node dissection. The TLG was found to be an independent predictor of axillary lymph node status. Patients with a TLG greater than 5.74 had a risk of axillary lymph node metastases 17.36 times higher than patients with a TLG ≤ 5.74 [17].

Albano et al. [14] compared the prognostic value of visual assessment, SUVs, and volumetric parameters in 52 patients with primary brain lymphoma imaged by [^{18}F]FDG PET/CT. Patients with a low-to-moderate MTV and TLG ($< 9.8 \text{ cm}^3$ and < 94 , respectively) exhibited longer progression-free and overall survival. Neither qualitative visual assessment nor SUVs were associated with outcome. As addressed by the authors, the high physiological [^{18}F]FDG brain uptake may mask or interfere with brain lesion assessment, and [^{18}F]FDG may not be the most suitable tracer for PET imaging of brain tumors.

Recently, a study investigated the ability of 4-borono-2-[^{18}F]fluoro-phenylalanine ([^{18}F]FBPA) PET/CT to differentiate radiation-induced necrosis from recurrent brain tumor and to select patients with brain tumors for boron neutron capture therapy. Twelve patients (gliomas = 9, malignant meningioma = 1, hemangiopericytoma = 1, brain metastases from lung cancer = 1) previously treated with external beam radiotherapy were included. SUVs and volumetric parameters were calculated. Patients who experienced tumor recurrence ($n = 6$) showed significantly higher values than patients with radiation-induced necrosis. Interestingly, the total lesion FBPA uptake was the only parameter without any overlap between the two groups (121.01 ± 50.48 and 12.36 ± 9.70 , $p = 0.0029$) [16].

Radiomics and artificial intelligence

Texture is a general term to describe the appearance of a surface or a volume. Traditionally, in radiology, it refers to visual characteristics, with lesions being classified as homogeneous or heterogeneous [42]. More recently, the concept of texture in medical imaging has evolved such that it now refers to objective quantitation of lesion homogeneity/heterogeneity based on the extraction of parameters through mathematical methods. Texture analysis may be computed using different approaches (e.g., histogram, wavelet, or fractal based), software (in-house, free, or commercial), and imaging modalities (e.g., slides, planar X-ray, SPECT and SPECT/CT, PET/CT, CT, and MRI). Moreover, it may be combined within mixed approaches to AI algorithms. Otherwise, AI algorithms may be used as stand-alone tools. Radiomics, which relies on texture analysis, and AI-based algorithms—image mining—have

been approached as biomarkers with diagnostic prognostic or predictive role [43].

Iwabuchi et al. [24] compared the fractal dimension (FD), a quantitative feature of tracer distribution, with the conventional quantitative specific binding ratio (SBR), which has been widely used to measure the total striatal [^{123}I]ioflupane uptake. The authors firstly performed a phantom study which confirmed the relationship between FD and SBR, and then proved that FD could detect dot and diffuse deterioration patterns. Thereafter, they moved to the clinical setting, retrospectively analyzing [^{123}I]ioflupane SPECT images of 150 patients, including 110 with parkinsonian syndrome. Here, FD performed better than SBR (AUC 0.90 and 0.94, respectively) in distinguishing between patients with and without parkinsonian syndrome. As expected, due to their complementarity, SBR was significantly empowered when combined with FD (AUC 0.90 vs 0.96, $p < 0.001$) [24].

Molina-García et al. [18] found that textural features derived from baseline [^{18}F]FDG PET/CT were able to predict pathological complete response to neoadjuvant chemotherapy and outcome (in terms of progression-free and overall survival) in 68 patients with locally advanced breast cancer. Conversely, in 82 patients with aggressive B cell lymphoma, textural features derived from baseline [^{18}F]FDG PET/CT failed to predict first-line therapy response, though some were correlated with disease-free and overall survival [19].

So far, AI and texture analysis applications in medical imaging have received the same level of attention from the scientific community. However, it should be stressed that data validation is an essential requirement for image mining studies. In fact, the significance, goodness and strength of the results rely on generalizability, which is proved only by independent data validation [43]. The study published by Nakajima et al. [22] is representative for the value of

validation. They tested the diagnostic performance of version 1.1 of an artificial neural network (ANN) applied to myocardial perfusion imaging in order to detect perfusion defects and ischemia and subsequently compared the results with those obtained using the previous version (v1.0). The ANN v1.0 was trained on data from 1051 Swedish patients, while the ANN v1.1 was retrained on data from 1001 Japanese patients. The cohort used in this study to validate the v1.1 was the same as was used for the validation of the v1.0 (106 Japanese patients). Results showed that the slight improvement in the AUC obtained by the v1.1 compared with the v1.0 in detecting stress defects was statistically meaningless (0.95 and 0.93, respectively; $p = 0.27$), while the v1.1 performed better than the v1.0 in detecting induced ischemia (0.98 and 0.88, respectively; $p = 0.0093$) [22].

Discussion

Our overview confirms that quantification permeates every type of nuclear imaging approach, providing helpful functional information regardless of the clinical setting (e.g., diagnosis, staging, prognostication) and the disease (e.g., neurological, cardiological, oncological).

Quantification makes the observations objective and is the prerequisite for the standardization of any type of assessment. It allows easier exchange of information and encourages medical debate at both the research and the clinical practice level [5, 6]. The identification of image-based biomarkers derived from nuclear medicine and molecular imaging techniques is one of the main topics of investigation. Several initiatives, aiming to change patient care by making imaging a more quantitative and reliable science, have been introduced, such as the Quantitative Imaging Biomarkers Alliance® (QIBA)

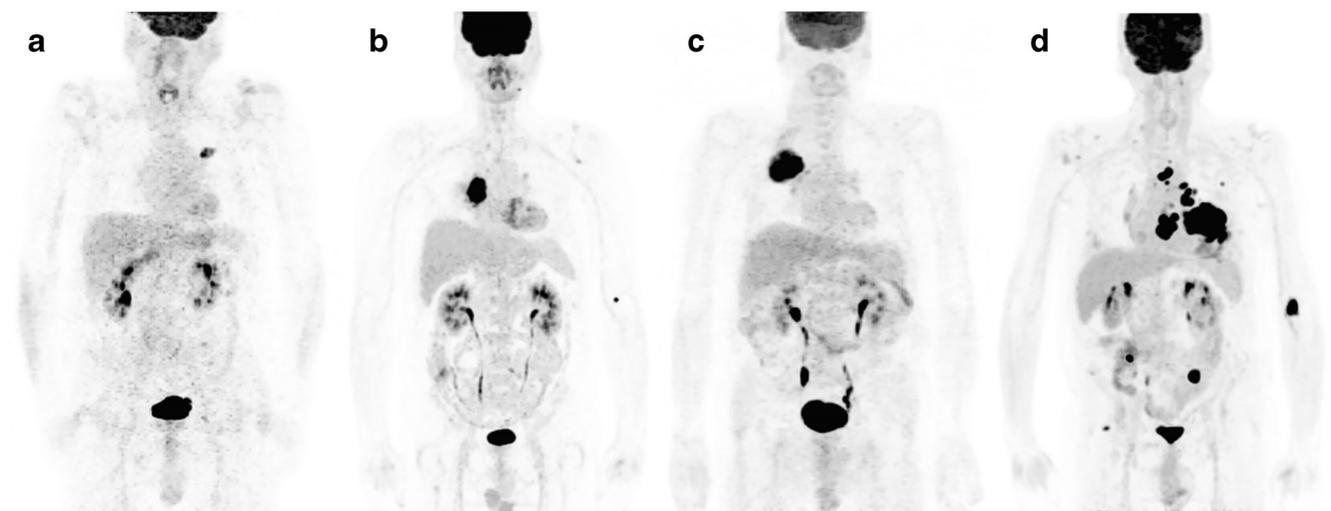


Fig. 2 [^{18}F]FDG PET maximum intensity projection coronal images of patients with non-small cell lung cancer at stages I (a), II (b), III (c), and IV (d) according to the 8th edition of the American Joint Committee on Cancer Staging Manual

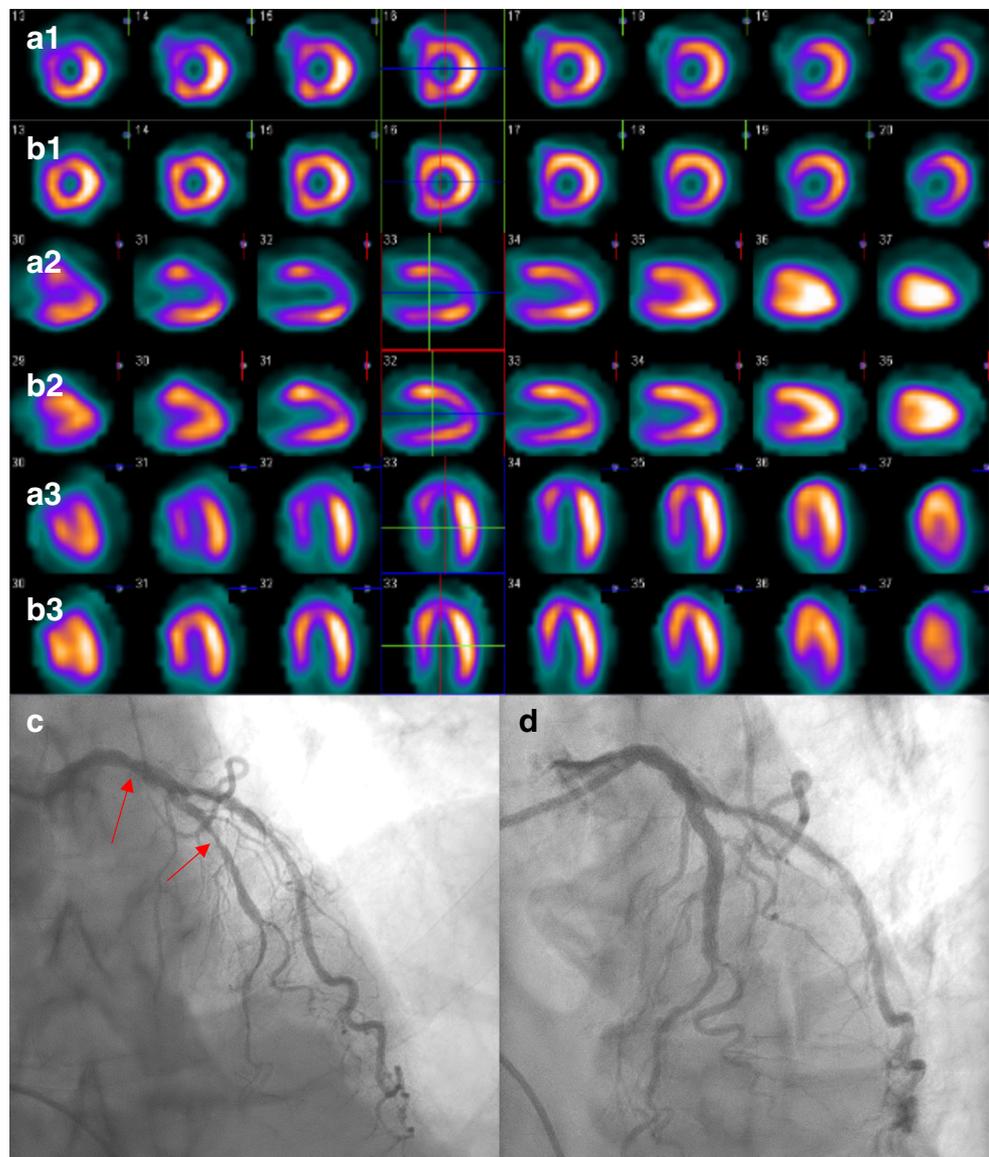
[44] and the Image Biomarker Standardisation Initiative (IBSI) [45]. Adherence to the procedural and clinical guidelines is an essential prerequisite for biomarker development and clinical implementation.

From a clinical point of view, in daily practice, quantification is crucial since (a) it is an easy way to assess the severity of a disease, helping the clinician to rapidly frame the general situation (e.g., cancer disease staging, cf. Fig. 2); (b) it facilitates communication, especially among colleagues, medical staff, and healthcare personnel, thereby avoiding possible misunderstanding and expediting patient management in both emergency and non-urgent situations; (c) it guides treatment decisions (e.g. through quantification of reversible myocardial ischemia, cf. Fig. 3), allows subsequent monitoring of treatment response and efficacy, and potentially results in reductions in drug dosage and side effects; (d) it makes patients

aware of their conditions, which may directly influence the quality of life.

Nonetheless, there are some meaningful caveats that must be considered when using a quantitative parameter. First, knowledge of general context (disease, comorbidities, symptoms, collateral imaging findings) is an essential requirement. Accordingly, a comprehensive assessment of each event or situation should be carefully performed to avoid misdiagnosis and/or disease over-/underestimation (e.g., evaluation of response to anticancer therapy, cf. Fig. 4). Moreover, physicians should be aware that indices (e.g., glomerular filtration rate or SUV) may derive from different techniques or approaches and, therefore, may differ significantly, consequently proving incomparable. Furthermore, methods for quantification may be hampered by some intrinsic inaccuracies that should be weighed when interpreting numbers.

Fig. 3 ^{99m}Tc -sestamibi gated SPECT/CT myocardial perfusion imaging (a.1–b.3) and coronary angiograms (c and d) in an 81-year-old patient with hypertension, dyslipidemia, and previous non-Q anterior myocardial infarct who presented with typical symptoms of angina. Stress images (a.1, a.2, and a.3, showing short, long vertical, and long horizontal axis views, respectively) and rest images (b.1, b.2, and b.3, showing short, long vertical, and long horizontal axis views, respectively) demonstrate reversible defects in all segments of anteroseptal and anterior regions. The percentage of perfusion defect was 23% (reversible 21% + irreversible 2%), consistent with inducible ischemia of large extent. A 10% threshold is recognized as an indication for percutaneous revascularization. Coronary angiogram (c) shows significant stenosis of the distal main branch and left anterior descending artery (red arrows), responsible for the extensive inducible ischemia. Panel d shows the final result after percutaneous transluminal coronary angioplasty plus two drug-eluting stents of the main branch and left anterior descending artery



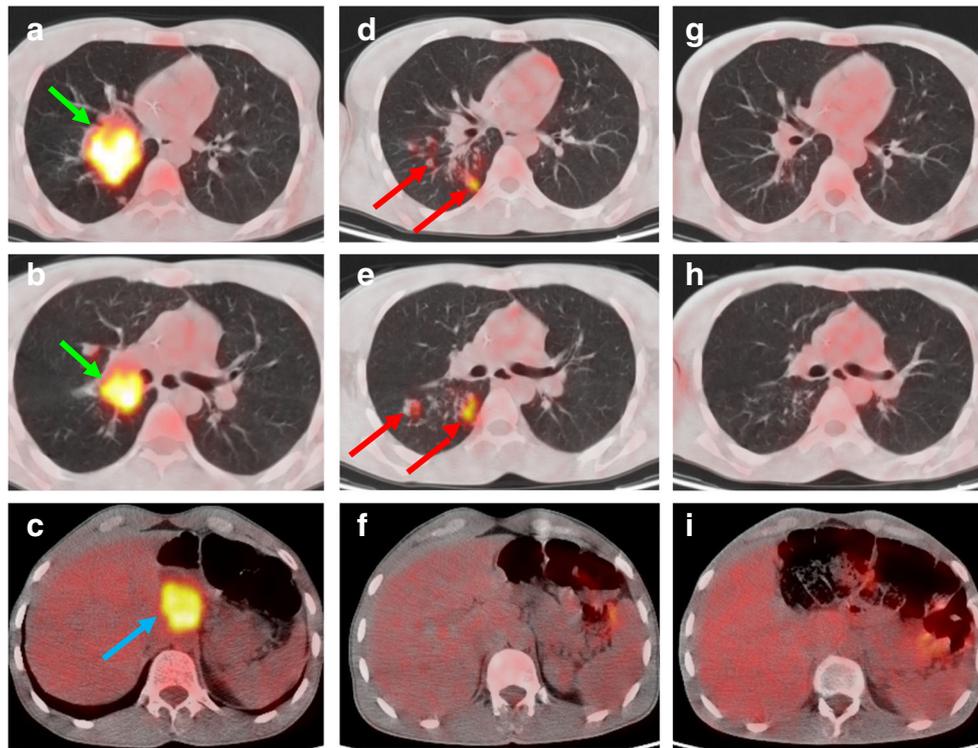


Fig. 4 [^{18}F]FDG PET/CT images of a 26-year-old male with refractory Hodgkin lymphoma treated with immune checkpoint inhibitors. Before the start of nivolumab treatment, PET/CT images (a–c) show intense [^{18}F]FDG uptake in the right hilar pulmonary region (Deauville score 5, $\text{SUV}_{\text{max}} = 15.81$; green arrows) and in a celiac lymph node (Deauville score 5, $\text{SUV}_{\text{max}} = 12.44$; blue arrow). Images obtained after 3 months (d–f) show the appearance of FDG-avid lung opacities (Deauville score 5, $\text{SUV}_{\text{max}} = 9.68$; red arrows) within the right lung. Both lesions previously

shown at the right lung hilum and in the celiac region had disappeared. The “tree in bud” appearance of the right lung lesions was interpreted as inflammatory rather than malignant. Additionally, the patient had a performance status of 0, without clinical deterioration. Accordingly, a diagnosis of lung infection was suggested. A CT-guided biopsy of one of the right lung lesions confirmed chronic inflammatory disease. [^{18}F]FDG PET/CT examination 3 months later was negative (g–i), and the patient was classified as being in complete response

Risk factors (e.g., heavy smoker), signs, symptoms (e.g., severe pain, light fever, high blood pressure, abnormal heart rate), and disease severity (e.g., end-stage heart failure, multiple metastatic tumor lesions, mild cognitive impairment) may also be quantified using a scale. Each level/stage within a scale is defined according to specific criteria, which generally refer to quantitative parameters or range. However, risk factors, signs and symptoms, and diseases may be scaled differently according to different definitions or guidelines; therefore, even if the categorizations or adjectives used to characterize a disease are the same, their meaning may change (e.g., oligometastatic disease may be defined as a tumor with up to three or up to five metastases [46]). Moreover, different scales may be available to quantify the same condition (e.g., five-point vs ten-point scale). Therefore, whatever quantification method is used, the scale to which it refers needs to be mentioned. Collectively, quantification has great value and is easily applicable, keeping in mind that it should be appropriately used and that quantitative parameters constitute just one piece of the entire puzzle.

All the above-described quantification approaches have some advantages and limitations. In fact, the simple methods

may fall short in terms of accuracy. SUV_{max} is commonly used in both oncological and non-oncological conditions with different purposes (e.g., diagnosis and prognostication), thanks to its easy computation that is independent from the clinical condition (health or disease), and with different tracers ([^{18}F]FDG or non-[^{18}F]FDG), acquisition protocols (e.g., whole-body, total-body, dual-point, dynamic), and scanners (PET/CT or PET/MRI). However, to guarantee reproducibility and comparability of measurements across patients’ scans and between centers, SUV should be used within specific standardization programs (e.g., the EARL initiative [47]).

Kinetic modeling provides the most accurate estimates of a condition, but it requires time-consuming protocols for image acquisition and analysis. Further studies on the development of algorithms for easier computation of kinetic biomarkers are of paramount importance.

Image mining (i.e., radiomics and artificial intelligence [43]) workflow requires extensive protocols for image pre-processing and data analysis as well as large datasets. Additionally, the pathophysiological relationship between the image mining biomarkers and the relevant clinical endpoints needs to be determined [8]. Finally, even if artificial

intelligence powers digital medicine [48], data reproducibility in biomarker discovery appears one of the main concerns [49]. Nonetheless, the available results encourage further investigations. In parallel with investigations of technical and clinical matters, issues concerning the “black box”, legal responsibility, and acceptance should be addressed. Since the adoption of innovations depends not only on the scientific evidence (“logos”, one of the components of the rhetorical triangle) but also on the integrity (“ethos”) and the values, necessities, and emotions (“pathos”) of the recipients, all these aspects need to be addressed before a large-scale implementation [49, 50].

In conclusion, in all the medical disciplines in which nuclear medicine examinations play a role, including especially oncology, cardiology, and neurology, quantification is essential. Both in research and in clinical practice, imaging findings expressed in terms of measures allow immediate understanding, objective communication, and comparison. Standardization and high-quality protocols in research and clinical routine are essential for the success and reliability of imaging biomarkers.

Acknowledgements MK PhD scholarship was funded by the AIRC grant IG-2016-18585. We thank all colleagues from the Nuclear Medicine Department of Humanitas Clinical and Research Center for their collaboration and Prof. Carlo Stella for cooperation with the Hematology Department.

Author contributions MS and MK conceptualized the study, MK performed data selection, MS and FB drafted the paper, FB commented on the paper, and MK reviewed the paper; all the authors approved the manuscript.

Compliance with ethical standards

Ethical statement In view of the nature of the present article (i.e., review), ethical approval was considered unnecessary. Figures are based on anonymized images, taken from existing research database, published with patient consent.

Conflict of interest The authors declare that they have no conflict of interest.

References

- Farolfi A, Ghedini P, Fanti S. Highlights from 2017: impactful topics published in the annals of nuclear medicine. *Eur J Nucl Med Mol Imaging*. 2019;46:217–23.
- Jadvar H. Highlights of articles published in annals of nuclear medicine 2016. *Eur J Nucl Med Mol Imaging*. 2017;44:1928–33.
- Inubushi M, Tatsumi M, Yamamoto Y, Kato K, Tsujikawa T, Nishii R. European research trends in nuclear medicine. *Ann Nucl Med*. 2018;32:579–82.
- Inubushi M, Kaneta T, Ishimori T, Imabayashi E, Okizaki A, Oku N. Topics of nuclear medicine research in Europe. *Ann Nucl Med*. 2017;31:571–4.
- Shryock RH. The history of quantification in medical science. Isis. The University of Chicago Press. The History of Science Society. 1961;52:215–37.
- Weisz G. Body counts: medical quantification in historical and sociological perspective. In: Weisz G, Jorland G, Opinel A, editors. *Body Counts Med Quantif Hist Sociol Perspect Hist Sociol sur la Quantif médicale*. Montreal, Kingston, London, Ithaca: McGill-Queen’s University Press; 2005. p. 377–93.
- Biomarkers [Internet]. Available from: <https://www.ncbi.nlm.nih.gov/mesh/68015415>. Accessed 26 Aug 2019.
- Aronson JK, Ferner RE. Biomarkers – a general review. *Curr Protoc Pharmacol*. Hoboken, NJ, USA: John Wiley & Sons, Inc.; 2017;9.23.1–9.23.17.
- Sheikh A. Evolution of quantification in clinical nuclear medicine: a brief overview of salient uses and upcoming trends. *J Nucl Med Radiat Ther*. 2018;09:375.
- Sollini M, Cozzi L, Antunovic L, Chiti A, Kirienko M. PET radiomics in NSCLC: state of the art and a proposal for harmonization of methodology. *Sci Rep*. 2017;7:358.
- Abdel Gawad H, Khalil MM, W. Shafaa M, Al Ramlawy S. Development of anatomically and lesion contrast-guided partial volume correction: new 3D formalisms and validation in phantom and clinical studies. *Ann Nucl Med*. 2019;33:481–94.
- Ringheim A, Campos Neto G de C, Martins KM, Vitor T, da Cunha ML, Baroni RH. Reproducibility of standardized uptake values of same-day randomized 68Ga-PSMA-11 PET/CT and PET/MR scans in recurrent prostate cancer patients. *Ann Nucl Med*. 2018;32:523–31.
- Annunziata S, Cuccaro A, Tisi MC, Hohaus S, Rufini V. FDG-PET/CT at the end of immuno-chemotherapy in follicular lymphoma: the prognostic role of the ratio between target lesion and liver SUVmax (rPET). *Ann Nucl Med*. 2018;32:372–7.
- Albano D, Bertoli M, Battistotti M, Rodella C, Statuto M, Giubbini R, et al. Prognostic role of pretreatment 18F-FDG PET/CT in primary brain lymphoma. *Ann Nucl Med*. 2018;32:532–41.
- Anwar H, Vogl TJ, Abougabal MA, Grünwald F, Kleine P, Elrefaie S, et al. The value of different 18 F-FDG PET/CT baseline parameters in risk stratification of stage I surgical NSCLC patients. *Ann Nucl Med*. 2018;32:687–94.
- Beshr R, Isohashi K, Watabe T, Naka S, Horitsugi G, Romanov V, et al. Preliminary feasibility study on differential diagnosis between radiation-induced cerebral necrosis and recurrent brain tumor by means of [18 F]fluoro-borono-phenylalanine PET/CT. *Ann Nucl Med*. 2018;32:702–8.
- Yoo J, Kim BS, Yoon HJ. Predictive value of primary tumor parameters using 18 F-FDG PET/CT for occult lymph node metastasis in breast cancer with clinically negative axillary lymph node. *Ann Nucl Med*. 2018;32:642–8.
- Molina-García D, García-Vicente AM, Pérez-Beteta J, Amo-Salas M, Martínez-González A, Tello-Galán MJ, et al. Intratumoral heterogeneity in 18F-FDG PET/CT by textural analysis in breast cancer as a predictive and prognostic surrogate. *Ann Nucl Med*. 2018;32:379–88.
- Parvez A, Tau N, Hussey D, Maganti M, Metser U. 18F-FDG PET/CT metabolic tumor parameters and radiomics features in aggressive non-Hodgkin’s lymphoma as predictors of treatment outcome and survival. *Ann Nucl Med*. 2018;32:410–6.
- Komek H, Can C, Yilmaz U, Altindag S. Prognostic value of 68 Ga PSMA I&T PET/CT SUV parameters on survival outcome in advanced prostatic cancer. *Ann Nucl Med*. 2018;32:542–52.
- Lebasnier A, Legallois D, Bienvenu B, Bergot E, Desmots C, Zalcman G, et al. Diagnostic value of quantitative assessment of cardiac 18F-fluoro-2-deoxyglucose uptake in suspected cardiac sarcoidosis. *Ann Nucl Med*. 2018;32:319–27.
- Nakajima K, Okuda K, Watanabe S, Matsuo S, Kinuya S, Toth K, et al. Artificial neural network retrained to detect myocardial ischemia using a Japanese multicenter database. *Ann Nucl Med*. 2018;32:303–10.

23. Habert MO, Bertin H, Labit M, Diallo M, Marie S, Martineau K, et al. Evaluation of amyloid status in a cohort of elderly individuals with memory complaints: validation of the method of quantification and determination of positivity thresholds. *Ann Nucl Med*. 2018;32:75–86.
24. Iwabuchi Y, Nakahara T, Kameyama M, Yamada Y, Hashimoto M, Ogata Y, et al. Quantitative evaluation of the tracer distribution in dopamine transporter SPECT for objective interpretation. *Ann Nucl Med*. 2018;32:363–71.
25. Lucignani G, Paganelli G, Bombardieri E. The use of standardized uptake values for assessing FDG uptake with PET in oncology: a clinical perspective. *Nucl Med Commun*. 2004;25:651–6.
26. Sollini M, Berchiolli R, Delgado Bolton RC, Rossi A, Kirienko M, Boni R, et al. The “3M” approach to cardiovascular infections: multimodality, multitracers, and multidisciplinary. *Semin Nucl Med*. 2018;48:199–224.
27. Hoffman EJ, Huang SC, Phelps ME. Quantitation in positron emission computed tomography: 1. Effect of object size. *J Comput Assist Tomogr*. 1979;3:299–308.
28. Huang SC, Hoffman EJ, Phelps ME, Kuh DE. Quantitation in positron emission computed tomography: 2. Effects of inaccurate attenuation correction. *J Comput Assist Tomogr*. 1979;3:804–14.
29. Hickeson M, Yun M, Matthies A, Zhuang H, Adam LE, Lacorte L, et al. Use of a corrected standardized uptake value based on the lesion size on CT permits accurate characterization of lung nodules on FDG-PET. *Eur J Nucl Med*. 2002;29:1639–47.
30. van den Hoff J, Oehme L, Schramm G, Maus J, Lougovski A, Petr J, et al. The PET-derived tumor-to-blood standard uptake ratio (SUR) is superior to tumor SUV as a surrogate parameter of the metabolic rate of FDG. *EJNMMI Res*. 2013;3:77.
31. Hofheinz F, van den Hoff J, Steffen IG, Lougovski A, Ego K, Amthauer H, et al. Comparative evaluation of SUV, tumor-to-blood standard uptake ratio (SUR), and dual time point measurements for assessment of the metabolic uptake rate in FDG PET. *EJNMMI Res*. 2016;6:53.
32. van den Hoff J, Lougovski A, Schramm G, Maus J, Oehme L, Petr J, et al. Correction of scan time dependence of standard uptake values in oncological PET. *EJNMMI Res*. 2014;4:18.
33. Hofheinz F, Bütof R, Apostolova I, Zöphel K, Steffen IG, Amthauer H, et al. An investigation of the relation between tumor-to-liver ratio (TLR) and tumor-to-blood standard uptake ratio (SUR) in oncological FDG PET. *EJNMMI Res*. 2016;6:19.
34. Kunikowska J, Matyskiel R, Toutouchi S, Grabowska-Derlatka L, Koperski Ł, Królicki L. What parameters from 18F-FDG PET/CT are useful in evaluation of adrenal lesions? *Eur J Nucl Med Mol Imaging*. 2014;41:2273–80.
35. Bahce I, Vos CG, Dickhoff C, Hartemink KJ, Dahele M, Smit EF, et al. Metabolic activity measured by FDG PET predicts pathological response in locally advanced superior sulcus NSCLC. *Lung Cancer*. 2014;85:205–12.
36. Tournoy KG, Maddens S, Gosselin R, Van Maele G, van Meerbeeck JP, Kelles A. Integrated FDG-PET/CT does not make invasive staging of the intrathoracic lymph nodes in non-small cell lung cancer redundant: a prospective study. *Thorax*. 2007;62:696–701.
37. Annunziata S, Cuccaro A, Calcagni ML, Hohaus S, Giordano A, Rufini V. Interim FDG-PET/CT in Hodgkin lymphoma: the prognostic role of the ratio between target lesion and liver SUVmax (rPET). *Ann Nucl Med*. 2016;30:588–92.
38. Ichise M, Meyer JH, Yonekura Y. An introduction to PET and SPECT neuroreceptor quantification models. *J Nucl Med*. 2001;42:755–63.
39. Osborn EA, Jaffer FA. The advancing clinical impact of molecular imaging in CVD. *JACC Cardiovasc Imaging*. 2013;6:1327–41.
40. Patlak CS, Blasberg RG, Fenstermacher JD. Graphical evaluation of blood-to-brain transfer constants from multiple-time uptake data. *J Cereb Blood Flow Metab*. 1983;3:1–7.
41. Slart RHJA, Glaudemans AWJM, Lancellotti P, Hyafil F, Blankstein R, Schwartz RG, et al. A joint procedural position statement on imaging in cardiac sarcoidosis: from the Cardiovascular and Inflammation & Infection Committees of the European Association of Nuclear Medicine, the European Association of Cardiovascular Imaging, and the American Society of Nuclear Cardiology. *Eur Heart J Cardiovasc Imaging*. 2017;18:1073–89.
42. Sollini M, Cozzi L, Chiti A, Kirienko M. Texture analysis and machine learning to characterize suspected thyroid nodules and differentiated thyroid cancer: where do we stand? *Eur J Radiol*. 2018;99:1–8.
43. Sollini M, Antunovic L, Chiti A, Kirienko M. Towards clinical application of image mining: a systematic review on artificial intelligence and radiomics. *Eur J Nucl Med Mol Imaging*. 2019. <https://doi.org/10.1007/s00259-019-04372-x>.
44. Sullivan DC, Obuchowski NA, Kessler LG, Raunig DL, Gatsonis C, Huang EP, et al. Metrology standards for quantitative imaging biomarkers 1. *Radiology*. 2015;277:813–25.
45. Zwanenburg A, Leger S, Vallières M, Löck S. Initiative for the IBS. Image biomarker standardisation initiative. 2016.
46. Reyes DK, Pienta KJ. The biology and treatment of oligometastatic cancer. *Oncotarget*. 2015;6:8491–524.
47. Kaalep A, Sera T, Oyen W, Krause BJ, Chiti A, Liu Y, et al. EANM/EARL FDG-PET/CT accreditation – summary results from the first 200 accredited imaging systems. *Eur J Nucl Med Mol Imaging*. 2018;45:412–22.
48. Fogel AL, Kvedar JC. Artificial intelligence powers digital medicine. *NPJ Digit Med*. 2018;1:5.
49. Stuppel A, Singerman D, Celi LA. The reproducibility crisis in the age of digital medicine. *NPJ Digit Med*. 2019;2:2.
50. Van de Ven AH, Schomaker MS. Commentary: the rhetoric of evidence-based medicine. *Health Care Manag Rev*. 2002;27:89–91.

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