Integrating molecular nuclear imaging in clinical research to improve anticancer therapy

de Vries, Elisabeth G E; Kist de Ruijter, Laura; Lub-de Hooge, Marjolijn N; Dierckx, Rudi A; Elias, Sjoerd G; Oosting, Sjoukje F

Published in:
Nature Reviews Clinical Oncology

DOI:
10.1038/s41571-018-0123-y

Publication date:
2019

OPINION

Integrating molecular nuclear imaging in clinical research to improve anticancer therapy

Elisabeth G. E. de Vries, Laura Kist de Ruijter, Marjolijn N. Lub-de Hooge, Rudi A. Dierckx, Sjoerd G. Elias and Sjoukje F. Oosting

Abstract | Effective patient selection before or early during treatment is important to increasing the therapeutic benefits of anticancer treatments. This selection process is often predicated on biomarkers, predominantly biospecimen biomarkers derived from blood or tumour tissue; however, such biomarkers provide limited information about the true extent of disease or about the characteristics of different, potentially heterogeneous tumours present in an individual patient. Molecular imaging can also produce quantitative outputs; such imaging biomarkers can help to fill these knowledge gaps by providing complementary information on tumour characteristics, including heterogeneity and the microenvironment, as well as on pharmacokinetic parameters, drug–target engagement and responses to treatment. This integrative approach could therefore streamline biomarker and drug development, although a range of issues need to be overcome in order to enable a broader use of molecular imaging in clinical trials. In this Perspective article, we outline the multistage process of developing novel molecular imaging biomarkers. We discuss the challenges that have restricted the use of molecular imaging in clinical oncology research to date and outline future opportunities in this area.

Anticancer drugs are the most abundant agents in the pharmaceutical pipeline and more clinical drug trials are conducted in oncology than in any other field of medicine[1,2]. Nevertheless, the clinical benefit of new anticancer therapies is often modest, especially in unselected patient cohorts[3–6]. Better patient selection before or at an early stage of therapy is essential to enhancing the benefits of treatment[7]. Most validated predictive biomarkers used for patient selection, such as HER2 expression in breast cancer and *BRAF*$^{V600E/K}$ mutation status in melanoma[8,9], require the analysis of tumour tissue samples. In addition, liquid biopsy approaches for serial, minimally invasive biomarker measurements are currently under development. The FDA has already approved a liquid biopsy test to analyse driver gene mutation in circulating cell-free tumour DNA from patients with non-small-cell lung cancer (NSCLC). Tumour biospecimens, however, provide limited information about the entire tumour burden. For example, tumour tissue samples might not encompass all tumour lesions with potential genotypic and/or phenotypic heterogeneity, whereas liquid biopsies might overcome issues surrounding tumour heterogeneity but lack information on the specific characteristics of individual tumour lesions. Molecular nuclear imaging provides potential solutions to these problems by enabling minimally invasive in vivo visualization of cellular functions and characteristics, and tracking of molecular processes. For this purpose, labelled probes or 'tracers' are administered and visualized. In clinical cancer research, numerous tumour characteristics can be analysed; radionuclide probes are currently available to visualize all established hallmarks of cancer, apart from the hallmark enabling replicative immortality[10,11]. In addition, molecular imaging can generate imaging biomarkers predicated on spatially delineated information derived from measurements made on an image, which can be used to guide patient selection for anticancer therapies[12]. Furthermore, labelling of drugs for imaging can provide insights into their whole-body distribution and pharmacokinetics, which can support drug development.

Several molecular imaging modalities are available, including PET, MRI, single-photon emission computed tomography (SPECT), and optical imaging. SPECT and PET, which are forms of nuclear imaging assessments, are most widely used in the clinic. The advantages of PET over SPECT include better spatial and temporal resolution and an intrinsically superior potential for quantification. PET imaging with 2-deoxy-2-[^{18}F]fluoro-D-glucose (FDG-PET), which enables the visualization of glucose consumption of tissues, is the molecular imaging modality for which the most knowledge is available, as underscored by the large numbers of publications and clinical trials using this technology. FDG-PET has important roles in the diagnosis, staging, and follow-up assessment of patients with cancer in daily clinical practice, and for that reason has been included in several clinical guidelines[13,14].

Approximately 2,210 trials of oncological PET imaging are currently listed in the ClinicalTrials.gov database. Around 1,025 of these studies are ongoing, at least 490 (48%) of which are using a PET tracer other than FDG. These studies are often small and are predominantly designed to investigate the feasibility of imaging with particular tracers. Few robust clinical trials powered to demonstrate clinical utility of molecular imaging with novel tracers have been conducted[15]. The example of hypoxia imaging with ^{18}F-fluoromisonidazole (^{18}F-FMISO)-PET illustrates the long development periods that typically precede the launch of large-cohort studies incorporating molecular imaging as an integral biomarker assessment (TABLE 1). The protracted nature of this process is due in part to the necessary involvement of

Table 1 | **Summary of all trials with reported outcomes of ^{18}F-FMISO-PET analyses in patients with HNSCC**

Study	*n* (total cohort)	Treatment	Timing of PET scan	Study design	Biomarker usage	Hypoxia indicators (reference tissue)	^{18}F-FMISO-related end point or outcomes
Thorwarth et al. (2005)[124]	15	Standard RT or CRT	Baseline	Prospective cohort	Integrated	TBR 1.4 (blood) and SUV_{max}	Hypoxia on 4-hour ^{18}F-FMISO-PET correlates with worse PFS after RT
Rajendran et al. (2006)[125]	73	Standard CRT or RT, or surgery followed by CRT or RT	Baseline	Prospective cohort	Integrated	TBR 1.2 (blood) in μCi/ml	Low (<median) baseline T/B_{max} and HV associated with longer OS
Rischin et al. (2006)[126]	45	CRT ± tirapazamine	• Baseline • Week 4–5	Prospective cohort (substudy of RCT)	Integrated	Visually qualitative (greater than background in adjacent or mirrored soft tissue)	Mildly-to-moderately higher ^{18}F-FMISO uptake than background at baseline correlates with high risk (HR 7.1) of LRF
Eschmann et al. (2007)[127]	14	Standard RT or CRT	• Baseline • After 30 Gy of RT	Prospective cohort	Integrated	TBR (muscle), no threshold defined; SUV_{mean}, and washout type	Decrease in ^{18}F-FMISO uptake during RT indicates radio-induced oxygenation
Dirix et al. (2009)[128]	15	Standard CRT	• Baseline • Week 4	Prospective cohort	Integrated	TBR 1.2 (blood)	High T/B_{max} (>median of 1.17) at baseline correlates with longer DFS
Nehmeh et al. (2008)[129]	20 (28)[a]	NA	Baseline (×2)	Feasibility study	Integrated	TBR 1.2 (blood) and SUV	PET scans performed 3 days apart show variability in tumour ^{18}F-FMISO uptake
Lee et al. (2008)[130]	10 (28)[a]	Standard RT (^{18}F-FMISO-guided dose escalation modelled but not actually performed)	Baseline	Feasibility study	Integral	TBR 1.3 (blood)	Dose-escalation based on ^{18}F-FMISO-PET–CT-guided dose painting does not compromise normal tissue sparing
Lee et al. (2009)[131]	20 (28)[a]	Standard CRT	• Baseline (×2) • Week 4	Prospective cohort	Integrated	Qualitatively greater than background (reference tissue NR)	Presence or absence of hypoxia not correlated with patient outcomes (LRF, RPFS, DMFS, or OS)
Kikuchi et al. (2011)[132]	17	NAC + surgery, RT or CRT	Baseline	Prospective cohort	Integrated	TBR median 1.3 (muscle) and SUV_{mean}	High ^{18}F-FMISO uptake (>median SUV_{max} of 2.3) correlates with shorter DSS
Yamane et al. (2011)[133]	13	NAC	• Baseline • 2–4 weeks after NAC	Prospective cohort	Integrated	TBR (muscle), no threshold reported; SUV_{mean} + 2 standard deviations, and SUV_{max}	Baseline ^{18}F-FMISO-PET data are not predictive of NAC outcome
Okamoto et al. (2013)[134]	11	NA	Baseline (×2)	Feasibility study	Integrated	TBR 1.5 (blood) and TBR 1.25 (muscle)	^{18}F-FMISO uptake between two separate PET scans was highly reproducible
Sato et al. (2013)[135]	23	Standard surgery	Baseline	Cross-sectional	Integrated	SUV_{max} and HIF1α expression	^{18}F-FMISO uptake was higher in tumours expressing HIF1α
Henriques de Figueiredo et al. (2013)[136]	15	Standard RT	Baseline	Feasibility study	Integrated	TBR 1.4 (NR)	Delineated HV differs among fixed threshold (≥1.4), adaptive threshold, and FLAB methods

Table 1 (cont.) | **Summary of all trials with reported outcomes of ^{18}F-FMISO-PET analyses in patients with HNSCC**

Study	*n* (total cohort)	Treatment	Timing of PET scan	Study design	Biomarker usage	Hypoxia indicators (reference tissue)	^{18}F-FMISO-related end point or outcomes
Bittner et al. (2013)[137]	16	Standard CRT	• Baseline • 2 weeks into CRT	Prospective cohort and feasibility	Integrated	TBR 1.5 (normal tissue)	In persistent hypoxia, hypoxic subvolumes had relatively stable geographical conformations
Wiedenmann et al. (2015)[138]	16	Standard CRT	• Baseline • 2 weeks into CRT • 5 weeks into CRT	Prospective cohort	Integrated	TBR 1.4 (blood)	Reoxygenation starts early during CRT and correlates with a higher probability of local control
Okamoto et al. (2016)[139]	20	Standard RT	• Baseline • During RT (after 30 Gy) • After RT (after 70 Gy)	Prospective cohort	Integrated	TBR 1.25 (muscle)	Intensity and volume of tumour hypoxia rapidly decreases in the early phase of RT
Lee et al. (2016)[73]	33 (216)[a,b]	CRT (standard versus ^{18}F-FMISO-guided dose reduction)	• Baseline • 1 week into CRT	Prospective feasibility	Integral	TBR 1.2 (muscle)	De-escalation of RT to normoxic lymph nodes resulted in high (100%) 2-year local and regional PFS and OS
Grkovski et al. (2017)[140]	123 (216)[a,b]	Standard CRT	• Baseline • 17 ± 5 days after baseline scan (after 1 cycle of chemotherapy and 10–20 Gy of radiation)	Prospective cohort	Integrated	TBR 1.2 (muscle)	Hypoxic subvolumes can be identified by dynamic ^{18}F-FMISO-PET in patients with normoxic tumours on static PET
Boeke et al. (2017)[141]	54 (90)[a,b]	Standard CRT or ^{18}F-FMISO-PET-guided dose-escalated CRT	• Baseline • At LRF	Prospective cohort	Integrated	TBR 1.4 (muscle) and SUV_{mean}	Locoregional recurrences after CRT correlate with initial GTV hypoxic subvolumes
Welz et al. (2017)[71]	25 (90)[a,b]	CRT (standard versus ^{18}F-FMISO-PET-guided dose escalated)	• Baseline • Week 3	Prospective cohort and feasibility	Integral	Voxels M ≥1.0	No hypoxia on dynamic PET correlates with better LRC; hypoxia imaging-guided dose escalation does not affect toxicity rates
Zips et al. (2012)[142]	25 (60)[a,b]	Standard CRT	• Baseline • Week 1 (after 8–10 Gy) • Week 2 (after 18–20 Gy) • Week 5 (after 50–60 Gy)	Prospective cohort	Integrated	TBR 1.4, 1.6, 1.8, or 2.0 (muscle)	Higher HV and TBR_{max} (>median 1.93) at 1 and 2 weeks are associated with local recurrence
Löck et al. (2017)[143]	25 (60)[a]	Standard CRT	• Baseline • Week 1 • Week 2 • Week 5	Prospective cohort	Integrated	TBR 1.6 (muscle); SUV_{peak} tumour versus SUV_{mean} background	Lower HV and TBR_{peak} on ^{18}F-FMISO-PET at baseline, week 1, and week 2 correlated with higher LRC rates

^{18}F-FMISO, 1H-1-(3-[^{18}F]fluoro-2-hydroxy-propyl)-2-nitro-imidazole; CRT, chemoradiotherapy; DFS, disease-free survival; DMFS, distant metastases-free survival; DSS, disease-specific survival; FLAB, fuzzy locally adaptive Bayesian; GTV, gross tumour volume; HIF1α, hypoxia-inducible factor 1α; HNSCC, head and neck squamous cell carcinoma; HR, hazard ratio; HV, hypoxic volume; LRF, locoregional failure; LRC, locoregional tumour control; M, malignancy value derived from voxel-based parameters for tumour perfusion and hypoxia; NA, not applicable; NAC, neoadjuvant chemotherapy; NR, not reported; OS, overall survival; PFS, progression-free survival; RCT, randomized controlled trial; RPFS, regional progression-free survival; RT, radiotherapy; SUV, standard uptake value; T/B_{max}, pixel or voxel with the highest tumour-to-blood ratio; TBR, tumour-to-background ratio. [a]The number in parentheses refers to the total number of patients mentioned in the study register, whereas the first number indicates the number of patients included in the reported study. Multiple reports refer to substudies of the same studies, probably involving particular, potentially overlapping, patient subgroups. [b]Ongoing trial.

different stakeholders, the multidisciplinary expertise required, high costs, the amount of time that has to be invested, and the complexity of multicentre imaging studies (TABLE 2).

The relevance of molecular nuclear imaging is growing as researchers and practitioners increasingly acknowledge the clinical implications of heterogeneity within and between tumour lesions, as well as the complexity of the biological factors relating to the rapidly expanding field of immuno-oncology. In particular, these technologies could be used to support decision-making in immuno-oncology

Table 2 | **Summary of hurdles and solutions to implement innovative molecular imaging in oncology trials**

Hurdles	Solutions
Requirement for diverse knowledge and skills	
• Necessitates the involvement of a multidisciplinary team with wide ranging expertise • An appropriate biological understanding and rationale are required • Complex analyses of imaging and pharmacokinetic data necessitate team members with a specific skillset • Involvement of oncologists early during trial design is often limited, potentially reducing the clinical relevance of questions investigated and thus the study findings	• Create dedicated multidisciplinary teams, including oncologists, pharmacists, nuclear medicine physicians, chemists, radiologists, biomedical researchers, and experts in health technology assessment (HTA) • Define relevant clinical problems to be solved using imaging as well as the particular tracer • Promote dedicated multidisciplinary training programmes, fellowships, and exchange programmes
Access to PET and/or SPECT tracers	
• The half-life of tracers is often very short creating time and production pressures (such as the requirement for local isotope and/or tracer production or a readily accessible cyclotron) • The availability of cyclotrons is limited • Complex, specialized radiochemistry and radiopharmacy skills are needed • Scalability to GMP-compliant facilities is lacking • Regulatory and legal issues can be a barrier to sharing of tracers produced in academic facilities • Concerns over safety, mainly relating to radiation, requiring safe handling procedures and the associated additional training and costs	• Consider purchasing commercially available tracer, or non-commercial tracer from another centre • For local tracer production: ensure facilities and expertise are in place or shared with others, including radiochemistry department, cyclotron (if local isotope production is needed, for example, owing to a short half-life), and animal PET facilities; collaborate with pharmaceutical companies or other institutions if purchasing molecule or drug of interest and/or radioisotope for labelling (assuming permissive isotope half-life); organize preclinical validation, quality system, and safety testing, GMP production, and writing of Investigational Medicinal Product Dossier (IMPD) • Adhere to regulatory framework by developing and/or exchanging checklists, protocols, and standard operating procedures (SOPs)
Complexity	
• Understanding of the disease biology is often limited • Advanced technologies are needed • Validation, quality testing, and standardization procedures are often lacking • Tracer production is complex and often problematic • Thus, multicentre trials are difficult to perform	• Education to drive scientific and technological advances and provide the expertise necessary for the development and clinical testing of advanced imaging modalities • Data sharing, the development and/or use of dedicated software as well as artificial intelligence and radiomics methodologies, and formal guidelines for quantification can optimize the handling of large amounts of data and improve image analysis • Serial and multiplexed imaging can provide novel information on biological changes and several tumour characteristics • Use appropriate data-analysis methods for clinically relevant molecular imaging interpretation
Clinical barriers	
• Patient accrual is often difficult when no extra treatment is involved, making trial participation less attractive • Limited consideration of heterogeneity in tumour tracer uptake • The design and data quality of imaging trials is often suboptimal • Level of evidence required can require many patients in a study • Limited scope for input from patients	• Design 'rewarding' trials whereby participation does not only mean that the patient undergoes an imaging procedure but coincides with receipt of a treatment (experimental or otherwise) • Consider crowd-sourcing of patients for enrolment in clinical trials • Create multilayer platform with simultaneous tumour biopsies to enable detailed analyses, blood samples for evaluation of circulating tumour DNA, and other biomarker, genomic, and radiomic elements to maximize research interest and enable testing of multiple hypotheses • Prioritize appealing clinical trials incorporating molecular imaging, and feedback results to clinicians, including rapid turnaround time of imaging test results including the original (key) images • Design studies and analyses to address heterogeneity; learn from rare tumour types; perform in silico analyses of small imaging studies; power trials to prove clinical meaningful end points • Ensure harmonization and standardization of procedures in multicentre studies • Create data warehouses, re-use imaging data, and perform meta-analyses • Involve patients in trial design and provide patient information folder including visual aids and advertising of studies that are open for enrolment
Cost issues	
• Costs of tracer production (according to GMP requirements) are considerable • PET imaging is costly • Limited support from pharmaceutical, medical diagnostics equipment, and tracer companies • Investors are often reluctant to finance imaging studies because they want to be sure that the product can ultimately be commercialized • Reimbursement issues for the costs of novel imaging	• More public–private partnerships, such as the Innovative Medicines Initiative (IMI) are required • Collaborations with the pharmaceutical industry should be pursued • Collaboration with insurance companies might be possible for studies analysing cost effectiveness or pattern of care studies • Conduct health technology assessment as early as possible: cost-effective is increased when imaging provides a good biomarker for decision-making on the use of expensive drugs • Motivate investors by explaining the potential utility of new tracers and scan indications

GMP, good manufacturing practice; SPECT, single photon emission CT.

research, which is likely to be important considering that approximately 940 immuno-oncology agents are currently being investigated in >3,000 clinical trials with a combined target enrolment of >577,000 patients[16,17]. To enable a broader use of molecular imaging in clinical trials, improved access to tracers, harmonization of procedures, data sharing, sophisticated methods of data analysis, and novel trial designs will all be crucial; validation and cost reductions will also be essential for integrating imaging biomarker assessments in daily clinical practice (TABLE 2).

Herein, we provide an overview of the state-of-the-art response measurements and the process of developing new tracers for PET imaging — from the laboratory to the analysis of clinical trial data. We discuss the challenges that have restricted the use of molecular imaging in clinical oncology research and highlight the numerous opportunities to exploit the potential of molecular imaging in future research. Finally, we describe how PET-based molecular imaging could streamline the biomarker and drug development process and outline the advances needed to support the clinical implementation of these modalities.

Tumour response measurements

Currently, tumour response to treatment is typically assessed using anatomical imaging approaches, especially CT and MRI; the role of molecular imaging in response assessment in solid tumours is currently limited. FDG-PET can be used to support a diagnosis of progressive disease during systemic treatment by detecting new lesions[18]. FDG-PET also has established roles in evaluating treatment responses in patients with lymphoma: the Deauville criteria based on FDG avidity of a lymphoma mass were developed by panels of experts following international meetings and have been proved in subsequent studies in large cohorts of patients to provide clinically meaningful information[19,20].

Indeed, before molecular imaging is used in daily practice, we advocate the same robust process of evaluation that has been applied to anatomical imaging modalities. The development of the Response Evaluation Criteria in Solid Tumors (RECIST), which are the standard criteria for anatomical imaging-based tumour response assessment, provides a prime example of the importance of validation, evaluation, and modification of imaging assessments, as well as the ongoing process of refinement for use with both new treatment and imaging modalities (including FDG-PET)[18].

Changes in tumour burden are used as surrogate end points of treatment efficacy during drug development and support drug registration; thus, validated and consistent criteria are required to define tumour response, and the FDA and EMA use the RECIST criteria as either primary or supportive data for regulatory approval of new therapeutics[21–23]. The RECIST working group initially simplified the 1979 WHO tumour response criteria after modelling and validation using a warehouse of imaging-based response data from studies involving chemotherapy[24]. These criteria were subsequently refined in RECIST version 1.1 on the basis of an assessment of a data warehouse containing information on >18,000 target lesions in >6,500 patients, simulation studies, and literature reviews[18]. Applicability of RECIST has also been demonstrated in studies of targeted agents[25]. FDG-PET was incorporated into version 1.1 of RECIST, although only as an adjunct to anatomical imaging for a more accurate determination of tumour progression. An expanded role for FDG-PET in RECIST-based response assessment is pending while FDG-PET data obtained according to generally accepted standard procedures emerge from clinical trials.

During treatment with immune-checkpoint inhibitors, increased tumour size is not always synonymous with worsening of the disease: some patients experience an initial increase in the size of lesions owing to T cell infiltration, followed by a decrease in tumour size as the anticancer immune response subsides — an observation termed 'pseudoprogression'[26]. Consequently, the conventional RECIST guidelines might not accurately capture responses to immunotherapies. An adapted set of guidelines, iRECIST, has been developed for collecting and ultimately validating response criteria in trials involving such agents[27].

Several other criteria for classifying patient responses to treatment are available, including the PERCIST[28] and Choi criteria[29]. However, these guidelines are partly based on expert opinion and/or data from small cohorts of patients and, therefore, have a lower evidence base and, correspondingly, are less commonly used than the RECIST criteria.

Quantification and imaging biomarkers

In routine care, the imaging features on diagnostic PET scans are reported primarily based on qualitative visual evaluation. However, tracer uptake is increasingly measured quantitatively according to the standardized uptake value (SUV), thus decreasing inter-observer and intra-observer subjectivity and improving the comparability of the findings. SUV values can serve as imaging biomarkers.

Radiomics is an approach that has been developed to extract even more information from imaging data[30,31], not only from anatomical CT and MRI assessments, but also from functional PET scans[32]. This approach can enable detailed quantification of tissue characteristics thereby capturing multiple features of all tumour lesions within a patient. These features include quantitative parameters of image intensity (such as CT-derived tissue density and functional information from MRI and/or PET radiotracer uptake), image intensity variability across the tumour, tumour shape and volume, and texture characteristics[32,33]. The use of semiautomated image-analysis tools can generate these quantitative metrics for tumour phenotyping using conventional imaging data, at essentially no additional cost. If clinical utility — for example, predictive or prognostic value — of the generated imaging parameters is demonstrated, these parameters can be regarded as imaging biomarkers. The first examples, comprising very small studies of radiomics analyses in relation to treatment outcome data, are now available[34–36]. These examples indicate that textural analysis of baseline FDG-PET–CT images can provide predictive biomarkers of survival outcomes in patients with locally advanced rectal cancer, that breast MRI-based radiomics has potential in image-based phenotyping to assess the risk of breast cancer recurrence, and that radiomic analyses of baseline chest CT images might help to predict which patients might develop immunotherapy-induced pneumonitis.

Developing new PET tracers

Despite the fact that molecular imaging is currently performed predominantly with FDG-PET, the adaptability of PET to detect different tracers provides potential opportunities for imaging of diverse characteristics of cancers[11,12]. A multidisciplinary effort is, however, essential to bringing new PET tracers to the clinic. Indeed, multiple stakeholders with diverse experience and expertise should ideally be involved in the various phases of this process, which include identification of a relevant clinical question to be addressed with molecular imaging, tracer development, preclinical and clinical studies, and data analysis (FIG. 1). The capacity to produce tracers according to good manufacturing

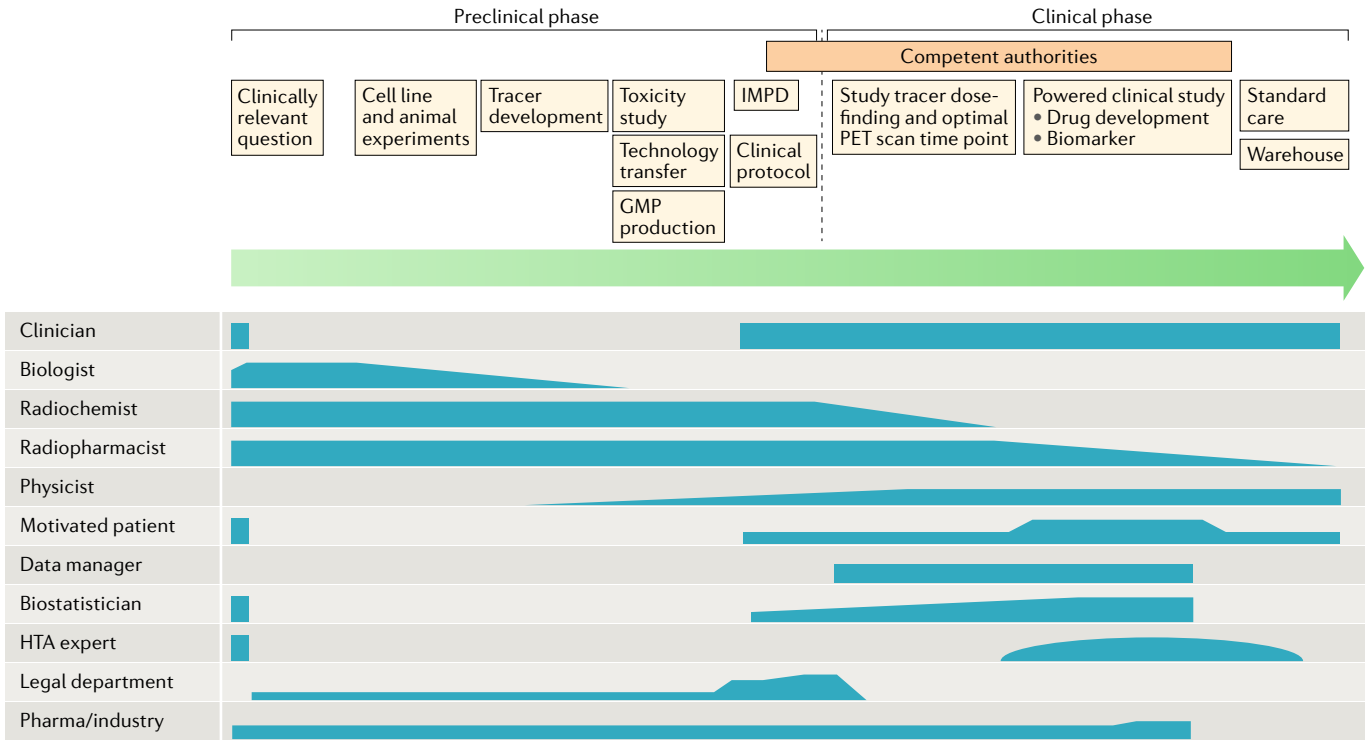

Fig. 1 | **Involvement of various participants in the development of molecular imaging-based biomarkers.** On the basis of our personal experience, a range of stakeholders with diverse expertise and experience have important roles in the various stages of imaging biomarker development, from conception, through tracer and protocol development, to implementation in the clinic. GMP, good manufacturing practice; HTA, health technology assessment; IMPD, Investigational Medicinal Product Dossier.

practice (GMP) guidelines and to assure standardization and harmonization of imaging procedures is also of critical importance to this process.

The production of a clinical grade radiopharmaceutical requires almost the same level of scrutiny as the development of a new drug[37,38]. Unlike drugs, however, new radiopharmaceuticals are rarely developed by (pharmaceutical) companies and instead are usually pursued in academic institutions on a much smaller budget. Reasons for this trend include the small and uncertain market for radiopharmaceuticals, which can limit profitability, and the radioactive nature of these agents, which results in additional infrastructure, regulatory and licencing requirements, complex logistics, and the need for frequent production of small batches. Furthermore, well-designed clinical validation studies necessitate multicentre collaboration to ensure enough participants, and specific logistics are required for such studies owing to radioactive decay (especially relating to the short half-life of most radionuclide used in PET tracers) and regulatory issues concerning tracer quality that must comply to release specifications across potentially different sites of manufacturing.

Tracer development for a first-in-human clinical trial consists of several steps. This process starts with radiochemical synthesis in the research lab[39], with radionuclide production in an on-site cyclotron in case of a radioisotope with a short half-life; preclinical evaluations, including in vivo proof-of-concept imaging studies; development of analytical methods; and pharmaceutical development, including purification, formulation and defining release specifications. These steps are followed by technology transfer from a research to a GMP environment, validation of analytical and manufacturing processes, and finally safety testing, if necessary, and preparing and submitting the required documentation[37]. For trials to be conducted in the EU, the Investigational Medical Product Dossier (IMPD) must be submitted to the competent authorities together with the clinical protocol[38]; the equivalent process in the USA involves filing of an Investigational New Drug (IND) application, which contains the same information and also includes details of the clinical protocol for FDA review.

A state-of-the-art PET facility requires a complex infrastructure and ideally a separate research and development and GMP environment with a cyclotron, clean rooms, hot cells, and automated synthesis modules; specific knowledge to safeguard radiation safety for personnel and product safety for patients is also essential. Owing to radioactive decay, the time from tracer manufacturing (including release quality control (QC)) to administration into patients is often restrictive. Radiolabelling of small molecules (<1 kDa), such as tyrosine kinase inhibitors (TKIs), is particularly challenging and often time-consuming because a drug-specific labelling strategy using a specific precursor and organic radiochemistry is required[40]. Use of a radionuclide with a half-life that matches the half-life of the drug is preferable for tracer labelling. For PET with labelled TKIs, therefore, radionuclides with a short half-life are generally used, such as ^{11}C (half-life of 20 minutes) or ^{18}F (half-life of 120 minutes). Several approved TKIs have been radiolabelled, including ^{11}C-imatinib, ^{11}C-erlotinib, ^{18}F-dabrafenib and ^{18}F-lapatinib (Supplementary Table 1), but only a few studies with such tracers have been performed in patients to date. The use of these tracers has, however, provided important information, for example, on the penetration of TKIs into brain

metastases[41,42]. Moreover, ^{18}F-MPG (*N*-(3-chloro-4-fluorophenyl)-7-(2-(2-(2-(2-[^{18}F]fluoroethoxy)ethoxy)ethoxy)ethoxy)-6-methoxyquinazolin-4-amine), a PET tracer with high specificity for EGFR proteins harbouring activating kinase mutation, has been developed to quantify EGFR mutation status[43]. This minimally invasive imaging biomarker approach is clinically relevant because these mutations predict efficacy of EGFR TKIs in patients with NSCLC.

Radionuclide generators are increasingly being used as an alternative approach for local radionuclide production at centres that lack an on-site cyclotron. Major improvements, especially in ^{68}Ge–^{68}Ga generators, have been achieved in the past few years, resulting in wider application of generator-based ^{68}Ga-labelled radiopharmaceuticals, such as ^{68}Ga-DOTA-octreotide derivatives and ^{68}Ga-prostate-specific membrane antigen (PSMA)[44]. Options for transportation of this radionuclide are limited by its short half-life of 68 minutes. Therefore, an on-site GMP facility for generator elution and tracer synthesis is still required. Moreover, the radionuclide yield is low and depends on the generator lifespan, and the associated costs are substantial[44].

Radiolabelling of antibodies is generally much simpler than radiolabelling of TKIs. In our experience, the average development time of clinically suitable antibody tracers is 4 months versus 4 months to 2 years with TKI tracers. Antibodies are generally labelled with radionuclides with a relatively long half-life, such as ^{89}Zr (half-life of 78.4 hours). A number of studies have been completed or are ongoing with investigational and registered antibodies, involving 26 ^{89}Zr-labelled antibody tracers in at least 52 oncology clinical trials (Supplementary Tables 1 and 2). During drug development, studies using these tracers have provided insight into issues including heterogeneity of target expression and penetration and saturation of the tumour by the drug[45–49].

Antibody-derivatives, such as diabodies, minibodies, and nanobodies, as well as small proteins, such as IL-2 and affibodies, are also attracting interest as potential tracers owing to their faster pharmacokinetic parameters, resulting in lower radiation exposure and short time-to-screen durations[50]. The tumour-to-background ratio of such agents seems favourable, although absolute tumour uptake is much lower than with full antibodies, mainly due to the lower internalization and faster clearance from circulation compared with antibodies[51,52]. Ultimately, the choice of tracer depends on the question that needs to be answered. When rapid screening for expression of a target in the tumour is the goal, an antibody derivative might be the preferred option. Labelling the full antibody is perhaps more relevant, however, when the question is related to the biodistribution of a therapeutic antibody. Obtaining greater insight into the difference between antibody and antibody-derivative PET tracers requires head-to-head comparisons of the various tracer types, and such comparisons have not yet been reported.

For peptide-based or protein-based radiopharmaceuticals, the chemistry, manufacturing, and control (CMC) procedures, especially formulation development and stability testing of the tracer vehicle, can be cumbersome and time-consuming. Most radiolabelled monoclonal antibodies are not formulated per se and, instead, are essentially diluted in normal saline. Formulation is not always necessary because of the radioactive decay and consequent short shelf life of the tracer[53]. In case of aggregate formation of protein tracers, however, it can be necessary to increase their stability using excipient buffers, salts, surfactants, polyol, disaccharides, or polysaccharides, amino acids, or antioxidants, with optimal pH being of critical importance. As a result, formulation development can be a bottleneck in tracer development.

After a tracer has been tested as a potential biomarker in a single-centre proof-of-concept study, additional multicentre validation studies are required to demonstrate reproducibility, and clinical utility. This stage requires large-scale production and distribution of the tracer, possibly followed by market authorization and reimbursement to make the tracer available to a large population of patients. Academic centres are usually not focused on nor equipped for large-scale commercial tracer production. Thus, the gap between novel technology development and commercialization has to be bridged, but without restricting speed, expertise, and flexibility in novel tracer development. Moreover, differences in the requirements relating to PET tracers exist between countries. For example, the European Union (EU) has regulations applying to the production of radionuclide tracers, but local authorities also have a role in regulating this process[38]. For diagnostic tracers, this framework means that GMP production is not mandated EU-wide and rather regulations are determined by the health inspectorates of the individual countries. In the Netherlands, the Health and Youth Care Inspectorate still requires GMP production of PET tracers for clinical use.

Ultimately, radiopharmaceutical manufacturing services and academia–pharma hubs, or other centres with relevant expertise, are necessary for large-scale production of PET tracers[39,54–56]. When drugs themselves are radiolabelled for use in phase I/II trials to enhance pharmacological knowledge, the tracer is often produced in only a limited number of batches for use at a single centre or a few centres. In such situations, development and use of the tracer early in the drug development process is crucial to ensuring that the molecular imaging data can be considered in early decisions[56].

Standardization and harmonization

To enable robust validation of a biomarker, the same assessment methodology should be applied repeatedly in a standardized way. During the initial phases of validation, biomarkers are often tested at a single centre. To be clinically validated, however, a biomarker should be tested in various settings and hospitals, independent of variables relating to the available facilities, equipment, or human resources[15,57,58]. Therefore, validation requires standardization and harmonization to eliminate differences between centres and thereby enable reliable comparisons of results. Several groups have published guidelines to harmonize scanning and analysis procedures[57,59–61], including the European Association of Nuclear Medicine (EANM) initiative to promote multicentre nuclear medicine and research: EANM Research Ltd (EARL). Centres that implement the EARL guidelines can obtain formal accreditation from the EANM. Subsequent to FDG-PET, harmonization of ^{89}Zr-antibody PET has also been established[62], with accreditation available through EARL. Similar harmonization initiatives have been launched by the American College of Radiology (ACR) Accreditation programme, the Society of Nuclear Medicine and Molecular Imaging (SNMMI) Clinical Trials Network, and the Quantitative Imaging Biomarkers Alliance (QIBA) of the Radiologic Society of North America (RSNA).

Insight generated by molecular imaging

Nuclear molecular imaging can be used to generate biomarkers and to support drug development and can also provide insight into the increasingly acknowledged existence and clinical relevance of tumour heterogeneity. In addition, given major progress in the field of immuno-oncology,

interest in visualizing the tumour immune microenvironment is growing rapidly.

Molecular imaging to generate biomarkers. Many potential nuclear imaging biomarkers are now available for use in clinical oncology research, although few have been validated and formally endorsed by societies or regulatory bodies for use in clinical practice. Importantly, biomarkers must be validated appropriately in order to ensure clinical utility. In this journal, a European Organisation for Research and Treatment of Cancer (EORTC) and Cancer Research UK (CRUK) consensus group has previously provided recommendations for parallel (rather than sequential) tracks of technical (assay) validation and biological and/or clinical validation of imaging biomarkers[12]. In addition, the FDA has issued guidance for biomarker development in their Critical Path Initiative[63]. Consistent with the classic drug development paradigm of performing two carefully controlled clinical trials to demonstrate clinical utility, the FDA's guidance recommends an initial verification clinical trial to determine the accuracy of the biomarker in a small cohort of patients, followed by a multicentre validation study with a large sample size and comparability end points. Additional approaches to validating imaging biomarkers are discussed in the 'Generating proof of clinical utility' section. Biomarkers should preferably be tested, and results reported, according to the Reporting Recommendations for Tumour Marker Prognostic Studies (REMARK) criteria[64], although analyses have revealed that the prevalence of REMARK guideline endorsement of biomarkers used in fields such as pathology remains limited[65,66].

Integration of molecular imaging into a clinical trial protocol can generate information about the potential clinical utility of the imaging biomarkers investigated. To validate the prognostic or predictive value of the imaging biomarker, nuclear imaging can be incorporated in a trial only to obtain information and not to influence the trial interventions, as an 'integrated' biomarker assessment. Thus, trials with integrated biomarker assessments can identify or validate biomarkers that can be used in future studies. One example is PET scanning using the ^{89}Zr-labelled anti-HER2 antibody trastuzumab at baseline in 56 patients with advanced-stage HER2-positive breast cancer participating in a treatment trial with the antibody–drug conjugate trastuzumab-emtansine (T-DM1)[45]. In total 29% of the patients had HER2-negative lesions on ^{89}Zr-trastuzumab PET. Median time to treatment failure was 11.2 months (95% CI 8–15 months) in the HER2-imaging-positive group versus 3.5 months in the HER2-imaging-negative group (95% CI 1.4–7.6 months).

A validated biomarker with proven prognostic or predictive value can subsequently be incorporated in a trial as an 'integral' biomarker, whereby the biomarker findings serve to direct trial procedures — typically to select patients for, or to guide adaptations in, therapy. An example is provided by the use of ^{111}In-octreotide SPECT scanning to select patients with midgut neuroendocrine tumours for inclusion in the phase III NETTER-1 trial of ^{177}Lu-DOTATATE (DOTA-octreotide) plus octreotide long-acting repeatable (LAR) versus octreotide LAR alone[67]. Patients were eligible to participate in this trial only when ^{111}In-octreotide tumour uptake was grade ≥2 in the lesion with the highest uptake, whereby grade 2 uptake is equal to that observed in the liver[67,68].

The development of tumour hypoxia imaging with ^{18}F-FMISO-PET illustrates that drawing firm conclusions about the prognostic or predictive value of an imaging biomarker — ultimately reflecting clinical utility — can be very difficult, even after numerous studies have been performed (TABLE 1). Tumour hypoxia is associated with therapy resistance, immunosuppression, and poor clinical outcomes[69]; therefore, imaging of tumour hypoxia has been extensively investigated, especially in the context of head and neck squamous cell carcinoma (HNSCC). ^{18}F-FMISO is the most frequently used tracer for hypoxia imaging. ^{18}F-FMISO has a nitroimidazole structure and can freely diffuse through cell membranes. Intracellularly, this tracer is reduced, which is reversible in normoxic conditions but not in the presence of hypoxia; the reduced molecules bind covalently to various intracellular molecules and in this way get trapped in hypoxic cells[70]. In the past decade or so, 19 studies with ^{18}F-FMISO-PET as an integrated biomarker assessment have been performed in the setting of HNSCC, involving a total of nearly 500 patients (TABLE 1). However, the lack of harmonization — the use of different parameters for quantification, different reference tissues, different treatment regimens, and variable timing of follow-up imaging — have complicated interpretation of the results and prevented robust conclusions. Sharing of raw imaging data from the baseline scans together with the outcome parameters of these 500 patients would enable meta-analyses of individual patient data with differences between centres accounted for mathematically. This approach might provide solid evidence for the prognostic value of ^{18}F-FMISO-PET-derived biomarkers and a strong basis for using ^{18}F-FMISO uptake as an integral biomarker in future therapeutic trials. A meta-analysis of all available data could have better informed subsequent studies, for example, with regard to optimal interpretation of the ^{18}F-FMISO results in order to identify prognostically relevant subgroups.

Currently, two studies in which ^{18}F-FMISO is being used to direct therapy are ongoing (TABLE 1) without the benefit of such information. In one study[71,72], patients with locally advanced HNSCC with tumour hypoxia, as determined using ^{18}F-FMISO-PET, are being randomly assigned to receive standard chemoradiotherapy or chemoradiotherapy with an increased radiation dose to hypoxic tumour volumes. In the second study[73,74], a reduced radiotherapy dose to lymph nodes will be delivered to patients with a favourable prognosis, defined as those with human papillomavirus-positive tumours and either no hypoxia evident on a baseline ^{18}F-FMISO-PET scan or early resolution of tumour hypoxia during radiotherapy. These studies illustrate how ^{18}F-FMISO evolved from an integrated to an integral biomarker in clinical trials.

Molecular imaging for drug development. Molecular imaging is also increasingly being used during early drug development of both small-molecule and large-molecule anticancer agents, such as TKIs and antibodies, by labelling the drugs themselves[75] (Supplementary Tables 1 and 2). To support anticancer drug discovery and development, trialists at The Institute of Cancer Research (London, UK) have proposed the Pharmacological Audit Trail, which contains crucial aspects to be considered: the target patient population; pharmacokinetic characteristics; evidence of target engagement, pathway modulation, and biological effect with proof-of-concept pharmacodynamic biomarkers; intermediate biomarkers of response; mechanisms of tumour response and resistance; and strategies to overcome resistance through combination or sequential therapy and/or new target and drug discovery[76]. Importantly, molecular imaging can provide additional information relevant to all parts of this audit trail.

The feasibility and relevance of the use of molecular imaging in this setting can be illustrated using several examples.

Visualization of target engagement and pathway modulation has been demonstrated with ^{18}F-fluoroestradiol (FES)-PET imaging of patients with metastatic breast cancer during treatment with the selective oestrogen receptor (ER) degrader fulvestrant; the median change in FES uptake was −85%, but varied widely. In addition, residual tumour FES uptake, indicating incomplete ER occupancy by fulvestrant, was associated with drug resistance and early disease progression[77]. The visualization of the pharmacodynamic effect of a heat shock protein inhibitor in patients with metastatic HER2-positive breast cancer is another example: early changes in tumour ^{89}Zr-trastuzumab uptake on PET scans were positively associated with changes in the size of individual lesions on a CT scan[49]. A PET study using radiolabelled docetaxel in patients with NSCLC demonstrated unintended pharmacokinetic interactions, with reduced delivery of ^{11}C-docetaxel into tumours for up to 4 days after administration of the anti-VEGFA antibody bevacizumab[78]. Finally, an example of how PET imaging can be used to verify strategies to overcome treatment resistance is provided by the use of ^{124}I or ^{131}I PET to demonstrate restoration of iodine uptake in patients with radioactive iodine-refractory metastatic thyroid cancer after treatment with a MEK inhibitor[79] or BRAF inhibitor[80,81]. These examples emphasize the potential of integrated molecular imaging studies to expedite the development and improve the use of anticancer therapies.

Visualization of tumour heterogeneity. Heterogeneity in tumour characteristics within and across lesions is a consequence of tumour evolution over time[82]. In addition, the delivery, uptake, and accumulation of drugs in tumours can be affected by several microenvironmental factors, including the structure and function of the tumour vasculature[83]. Tumour heterogeneity is increasingly acknowledged to pose a major challenge to defining the optimal treatment for individual patients[84]. Even analysing multiple tumour biopsy samples, circulating tumour DNA, and circulating tumour cells would not necessarily provide sufficient details of the specific heterogeneous characteristics of all metastases that are present in a patient. In this respect, whole-body PET imaging can clearly support findings relating to tumour heterogeneity. Currently, at least 35 PET imaging studies specifically addressing tumour heterogeneity are registered in the ClinicalTrails.gov database, of which 19 are actively recruiting patients. In addition to 18 studies using FDG-PET, 14 other tracers are being used across 17 studies (Supplementary Table 3).

Published results from several molecular imaging studies have indicated how findings relating to tumour heterogeneity can potentially affect clinical decision-making. For example, substantial heterogeneity has been observed for FES uptake indicating variation in ER expression across tumour lesions present in individual patients[85]. In 91 patients with metastatic breast cancer, of the 1,617 metastases detected on imaging, 11.2% were visible on CT only, 56.6% on FES-PET only, and 32.2% were visible with both imaging modalities[85]. Furthermore, median tumour FES uptake values varied greatly between patients[85]. Interestingly, the level of FES uptake of bone metastases was higher than that of lymph node and lung metastases[85]. The findings of these studies might in the future stimulate studies to analyse the role of FES-PET in treatment decision, such as the allocation of hormone therapy.

In another study[86], ^{18}F-fluorodihydrotestosterone (FDHT)-PET and FDG-PET were used to investigate the heterogeneity of androgen receptor (AR) expression and glycolytic activity, respectively, in a cohort of 133 patients with metastatic prostate cancer. Tumours were biopsied to enable correlation of imaging phenotypes with histological findings[86]. Interestingly, imaging characteristics revealed clinically relevant heterogeneity in AR expression and glycolysis on a lesion and individual patient basis, which correlated with prognosis[86]. Most lesions expressed AR, although 49% of the patients had at least one AR-negative lesion with detectable FDG uptake, and this imaging phenotype was strongly associated with short survival durations, possibly owing to anti-androgen resistance[86]. The authors suggest the consideration to biopsy AR-negative lesions to check for the presence of non-prostatic cancer or to identify actionable mutations.

Importantly, the presence of a drug target in the tumour, as assessed using immunohistochemistry, does not mean that the drug can reach the target. In patients with HER2-positive metastatic breast cancer (according to immunohistochemistry or fluorescence in situ hybridization analysis of the primary tumour or, when not available, a metastasis) who underwent whole-body PET imaging with ^{89}Zr-trastuzumab before treatment with T-DM1, no tumour tracer uptake was detected in around 30% of the cohort[45]. This patient subgroup had a shorter duration of treatment benefit than the subgroup with detectable tumour ^{89}Zr-trastuzumab uptake. This example illustrates that accessibility of the tumour for the drug might determine the outcomes of treatment and that molecular imaging can provide clinically relevant insights into tumour targeting and penetration.

Heterogeneity has been observed not only for tracers targeting receptors expressed by tumour cells, but also for tracers that bind to targets in the tumour microenvironment, including angiogenic growth factors, such as VEGFA. Indeed, PET imaging with the VEGFA-binding tracer ^{89}Zr-bevacizumab revealed 125 evaluable tumour lesions in 22 patients with metastatic renal cell carcinoma[47]. The level of tumour ^{89}Zr-bevacizumab uptake was generally high but varied widely within and between patients[47]. This observation suggests remarkable interpatient and intrapatient tumour heterogeneity in VEGFA expression. Studies in larger cohorts of patients are needed to determine implications regarding susceptibility to anti-angiogenic therapy and potentially other antibody therapies (owing to issues of drug accessibility).

Imaging the tumour immune environment. The introduction of immuno-oncology drugs into the clinic has been an important breakthrough, which has invigorated interest in biomarkers that capture the dynamics and heterogeneity of the immune response, especially within the tumour microenvironment[16,87]. When assessed immunohistochemically, expression of programmed cell death 1 ligand 1 (PD-L1) can be used to enrich certain patient populations with responders to anti-programmed cell death protein 1 (PD-1) or anti-PD-L1 antibodies, but not to precisely identify which patients will benefit from these agents. Extensive studies have been performed in animal models to study the tissue distribution of anti-PD-1 and anti-PD-L1 antibodies. Using immuno-PET with various tracers, the presence of PD-L1 in tumours, lymph nodes, spleen, thymus and brown tissues has been demonstrated in mice[88–91]. Reports on PET imaging with ^{89}Zr-labelled atezolizumab (anti-PD-L1 antibody) or nivolumab (anti-PD-1 antibody) have described considerable heterogeneity in tumour uptake of these tracers in patients[46,92]. Interestingly, PD-L1 is expressed by tumour cells as well as immune cells, including T cells and macrophages, whereas PD-1 is predominantly expressed by T cells[93,94]. In the study of ^{89}Zr-atezolizumab[46], 22 patients

with triple-negative breast cancer, NSCLC, or bladder cancer underwent PET imaging before atezolizumab treatment, revealing high signal intensities in lymphoid tissues and at sites of inflammation. No tracer uptake in thymus and brown adipose tissue was detected, probably owing to the fact that adults have no thymus tissue and relatively little brown fat[46]. PD-L1 signal intensity was generally also high, but heterogeneous, in tumours, varying within and between lesions, patients, and tumour types[46]. Intriguingly, clinical responses in patients included in this study were better correlated with pretreatment PET signal than with immunohistochemistry-based or RNA sequencing-based predictive biomarkers, encouraging further development of molecular PET imaging for assessment of PD-L1 status and clinical response prediction.

Currently, >90 bispecific antibodies have been developed, often with intended applications in anticancer immunotherapy[95]. The development of such agents for clinical use is more challenging than with conventional monoclonal antibodies, as very limited information on the biodistribution of bispecific antibodies is available[96,97]. This deficit is partly explained by the fact that the two antigen-binding sites of a bispecific antibody recognize different targets, which might result in substantial variability in biodistribution. In transgenic immunocompetent mice expressing human CD3 and bearing murine tumours transfected with human HER2, the distribution of a full-length bispecific anti-HER2–CD3 antibody was predominantly determined by the CD3 arm[98]. Using CD3 affinity variants, the investigators showed that the high affinity of this agent for CD3 reduced the systemic exposure and shifted the distribution of the antibody, as measured with SPECT, away from the HER2-expressing tumours to T cell-containing tissues[98]. These findings illustrate that molecular imaging can serve to provide early information on how well a drug reaches the tumour and, therefore, on the potential effectiveness of such agents.

Furthermore, different tracers have been developed to visualize the various components of the local immune response in the tumour microenvironment during immunotherapy. Activated T cells can be visualized using tracers including ^{18}F-labelled IL-2 (REFS[99,100]) and ^{89}Zr-labelled anti-CD8 antibodies[101], and PET imaging studies of visualizing a target on monocytes and/or macrophages are being initiated[102,103]. Comprehensive immunohistological analyses of metastases from a patient with ovarian cancer have revealed that heterogeneity in CD8 expression can also be very extensive[104]. Given the fact that anticancer immune responses are highly dynamic and heterogeneous, using tracers that capture different elements of the response might be a very interesting approach for individualized decisions on treatment durations and/or combinations.

A promising approach to evaluating different aspects of the immune response involves the use of multiplexed imaging, which encompasses multimodality, multisignal, multiparametric, and theranostic imaging[105]. Current applications of multiplexed imaging involve combinations and fusion of imaging techniques, such as PET, CT, and/or MRI. FIGURE 2 illustrates this approach with imaging of VEGFA, which is known to be present in the tumour microenvironment, using ^{89}Zr-bevacizumab PET, together with anatomical imaging with CT and MRI. This approach can also be expanded with the use of multiple molecular imaging tracers, not limited to PET tracers. For example, clinical studies could also incorporate fluorescent tracers coupled with in vivo or ex vivo fluorescence imaging, which can provide information on tracer localization down to the subcellular level[106,107]. In this way, multiplexed imaging would provide the possiblity of obtaining serial insights on several tumour characteristics at the same time points in individual patients.

Data sharing for clinical relevance

Large, well-powered imaging studies are difficult to perform, although several studies with large cohorts of patients are ongoing (NCT01957332, NCT00606294, NCT0313426, and NCT03396874). Adherence to uniform imaging procedures is crucial to advancing the field, by enabling the storage of imaging data in repositories and permitting re-utilization and pooling of data for analyses. As mentioned previously, validation of the RECIST criteria was facilitated by the use of a warehouse filled with well annotated imaging and outcomes data from clinical trials. Moreover, a number of initiatives for repositories of molecular imaging data exist, including the RSNA's Quantitative Imaging Biomarkers Alliance (QIBA) and Radiology Informatics

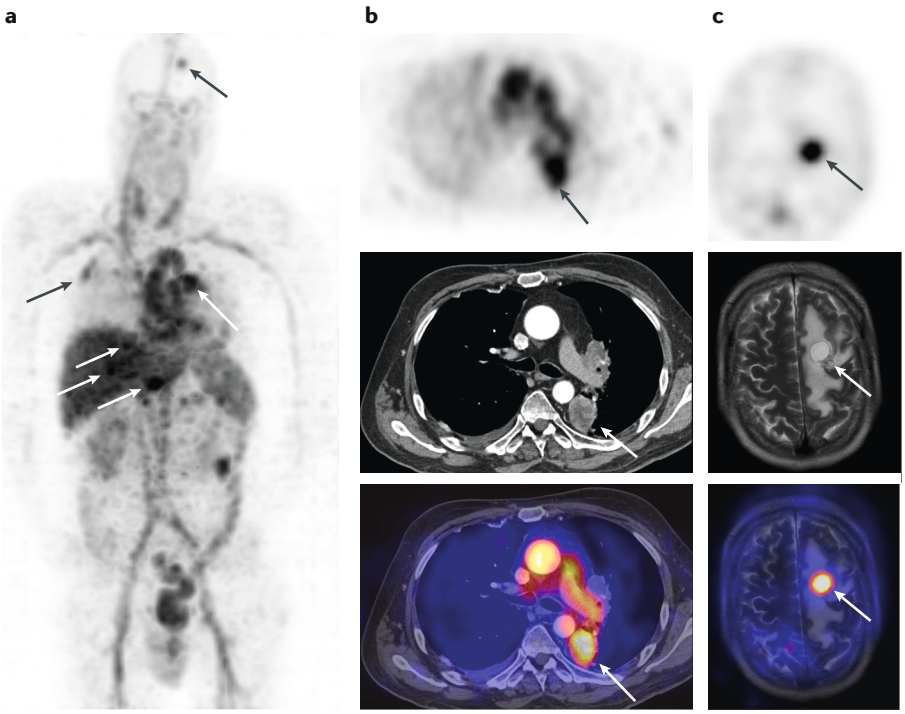

Fig. 2 | **Example of multiplexed imaging with ^{89}Zr-bevacizumab-PET, CT, and MRI.** **a** | ^{89}Zr-bevacizumab-PET maximum intensity projection image of a patient with metastatic renal cell carcinoma performed before first-line systemic treatment as part of a clinical trial to evaluate the feasibility of ^{89}Zr-bevacizumab-PET for VEGFA-based imaging biomarkers (NCT00831857). The image reveals metastases in the brain, lung, mediastinal lymph nodes, and liver (arrows). **b** | Transversal plane ^{89}Zr-bevacizumab-PET, CT and fusion images of the patient's chest providing anatomical imaging and ^{89}Zr-bevacizumab uptake data on the lung metastasis (arrow). **c** | Transversal plane ^{89}Zr-bevacizumab-PET, MRI, and fusion images of the patient's head provide analogous information on the brain metastasis (arrow).

Committee (RIC) Quantitative Imaging Data Warehouse (QIDW), the National Biomedical Imaging Archive (NBIA), the American College of Radiology Imaging Network (ACRIN), and The Cancer Imaging Archive (TCIA). In the USA, the National Oncologic PET Registry (NOPR; NCT00868582) was developed in response to a proposal from the Centers for Medicare and Medicaid Services to expand coverage for FDG-PET to include cancers and indications currently ineligible for Medicare reimbursement[108,109]. Patient registration for FDG-PET was performed between 2006 and 2013, and for sodium fluoride PET, between 2011 and 2017. This registry thus contains a wealth of observational data reflecting actual clinical practice with PET–CT systems derived from large cohorts of patients.

The development of a warehouse for tracers that are used less frequently than FDG, for example, could also be very helpful. Each of the studies performed with such tracers might be underpowered to make credible conclusions, as emphasized above for the studies using ^{18}F-FMISO (TABLE 1). When performed according to similar protocols, however, data from studies in small cohorts of patients can be combined and thus more robust conclusions can be drawn. For example, data from studies with ^{89}Zr-labelled antibodies are now being combined in a warehouse[110]. The usefulness of these warehouses would be even greater if they could be expanded to include relevant phenotypic and genomics data in combination with data generated through radiomics research.

It is becoming increasingly clear that a single biomarker will not provide sufficient information to support treatment decisions for most patients with cancers, and data-sharing warehouses could be very helpful in overcoming this problem. For example, a minimal set of features that has substantial predictive power, which might include more than one imaging biomarker or an imaging biomarker combined with biomarkers based on blood or tumour tissues, could be identified using the data warehouse. Ideally, generating this set of features will require few financial and manpower resources (because the process will mostly be computational), which increases the likelihood of fast translation to clinical practice on a broad scale. Such data warehouses will also enable the use of machine learning models or deep learning methods, which are generally data-intensive[111] but can yield important new insights and predictive algorithms that can then be tested prospectively in pivotal studies.

Generating proof of clinical utility

To generate evidence of clinical utility, empirical data is needed from high-quality studies appropriately designed to address a specific clinical question; however, evaluation of a molecular imaging biomarker together with treatment in a randomized trial, which is the default design for predictive biomarkers, is not necessarily required for biomarker-informed treatment decisions. Well-conducted observational cohort studies in patients treated with the relevant agent can also yield convincing evidence of clinical utility for biomarkers. This approach is especially appealing when refinement of patient selection is needed for an established treatment indication that is associated with suboptimal clinical benefit rates; for example, if a subgroup of patients has rapid disease progression after such treatment. The aim would then be to predict which patients will not have any clinical benefit — upfront or early following treatment initiation — so that alternative treatment modalities can be prescribed or investigated for these patients in the future. Within an observational study, a biomarker signature predicting rapid progression can be developed or tested in a cohort of treated patients. The findings of such observational studies cannot prove that the therapy was entirely ineffective in patients with rapid progression; however, if the biomarker can be used to identify a subgroup with an overall dismal disease course irrespective of treatment with a specific agent, omitting the therapy for these patients would be a clinically sound decision.

When incorporated in a randomized treatment trial, molecular imaging can be used more explicitly to identify patients benefiting from the treatment. Over the past decade, various randomized trial designs formally incorporating biomarkers have been proposed and used, including adaptive trials[112]. In many such designs, both the predictive and the prognostic value of molecular imaging can be assessed prospectively. Despite the anticipated excellent performance of imaging biomarkers, randomized studies still require sample sizes that are infeasible for most molecular imaging studies. Current efforts to improve the statistical design and data analysis of trials involving patients with rare diseases might, therefore, prove very relevant in decreasing the amount of data necessary to generate sufficient evidence for molecular imaging biomarkers[113]. Such efforts include the improvement and evaluation of methodology ranging from best randomization practices in small populations to evidence synthesis methods that incorporate lower-level evidence than that generated through randomized trials, as well as the use of within patient data modelling to improve statistical efficiency.

In many molecular imaging studies, the readout is a single aggregated measure of tracer uptake, such as the maximal SUV (SUV_{max}), averaged over all metastases present in an individual patient; although informative, this approach does not exploit the full potential of molecular imaging. Alternatively, a large array of potentially relevant features can be extracted from a single PET scan[114]. Determining how to combine this wealth of information, possibly also incorporating other biomarkers and tests in a signature, is not a trivial process and necessitates statistical methods capable of identifying robust prognostic or predictive signatures using only small studies. After signature development, appropriate validation steps will always be necessary.

Decision modelling (also using economic criteria) early in the evaluation of molecular imaging biomarkers could facilitate efforts to obtain evidence of clinical utility. Such decision modelling would link evidence from molecular imaging studies with outcomes data from treatment trials and other information, including data from repositories. This approach can enable projections of efficacy and cost-effectiveness on the basis of the existing evidence and can facilitate the prioritization of research agendas to generate new evidence to address the most critical sources of uncertainty[115]. Furthermore, such models can help to identify optimal thresholds for molecular imaging biomarker signatures. Basically, by evaluating the cost-effectiveness at different thresholds for biomarker signature positivity, yielding different pairs of sensitivity and specificity for each threshold and therefore different projected cost and effect consequences, the optimal operating point can be determined from a cost-effectiveness point of view[116]. Determining the best study design and data-analysis approach to efficiently generate the necessary level of evidence for a particular molecular imaging application at a given time is not always a straightforward process. In this context, decisions can be supported by evidence gleaned from modelling studies using simulated data representative of actual patients, as informed by data from earlier studies (for example, using international repository initiatives) or from patients included in an earlier phase of the same clinical imaging study while it

remains ongoing. By performing hundreds of thousands of in silico studies with various designs, sample sizes, and analysis techniques, the most effective approach with the greatest likelihood of developing meaningful new evidence can be identified. Such simulations have been successfully used in various related fields of research, for instance, for identifying the optimal data-analysis approach for developing classifiers based on gene-expression data[117] or for identifying the optimal randomized controlled clinical trial design for a particular intervention[118], and also in the context of cancer and (imaging) biomarker studies[119–121].

When multiple lesions are present in individual patients, per-lesion analyses of molecular imaging features in relation to their disease course, while taking within-patient clustering of lesions with particular features into account, will improve statistical efficiency and thus the power of a study that investigates the prognostic and/or predictive power of imaging biomarkers, and will provide supporting mechanistic information. The biological insights from these analyses could ultimately be used to improve patient outcomes. For instance, in ER-positive metastatic breast cancer, any ER-negative tumour lesions detected using FES-PET could potentially be targeted with radiotherapy, while the other ER-positive lesions would be expected to respond to systemic endocrine treatment. Thus, randomized trials could be designed to test the hypothesis that patients with a heterogeneous ER status on FES-PET will benefit from local treatment of ER-negative tumours added to standard hormone therapy.

Costs and cost-effectiveness

Limited information is available on the roles of molecular imaging in the early stages of drug development, although early and rational go–no-go decisions are clearly cost-effective. For example, development of an antibody-based drug intended to target pancreatic cancer cells was stopped early after it was demonstrated to be fatally toxic in animals, with molecular imaging of a labelled version of drug revealing that the agent was taken up in the bone marrow where it depleted white blood cell precursors resulting in severe cytopenias[48]. This example shows that the results of molecular imaging with a radiolabelled antibody, which perhaps cost US$30,000 for radiopharmaceutical development, prevented not only substantial further drug development costs but also unnecessary exposure of patients to the toxic drug, clearly making the no-go decision highly cost-effective. The costs of clinical drug and/or biomarker tracer development can be covered by pharmaceutical companies or grants from non-profit organizations. Robust trials are required to prove the relevance of the tracer in the clinic, for which major funding is necessary and is unlikely to be provided other than by non-profit organizations.

In the clinic, current costs of a standard FDG-PET–CT scan range from approximately €1,500 in the Netherlands to US$7,000 or more[122] in the USA, including production of the tracer, scanning, and reporting by the nuclear medicine physician. The introduction of new PET tracers and especially serial imaging will rapidly increase costs; therefore, clinical utility and cost-effectiveness have to be demonstrated before implementing new molecular imaging techniques or new tracers in routine clinical practice. To this end, health technology assessment (HTA) can be incorporated in validation studies to analyse the economic implications in parallel with clinical utility of the imaging biomarker. At the same time, costs will probably reduce with wider use of PET imaging owing to economies of scale. Remarkably few studies have addressed the cost-effectiveness of PET scanning, and most of the available studies were focused on FDG-PET. The Choosing Wisely initiative of the American Board of Internal Medicine in partnership with Consumer Reports seeks to advance a dialogue in the USA on avoiding wasteful or unnecessary medical tests, treatments, and procedures. This initiative provides advice on when FDG-PET is not relevant for the patient. If PET imaging with FDG or other tracers can replace other diagnostics or lead to better treatment choices, however, such assessments could potentially be cost-effective. Of note, a computer simulation has been used to evaluate the effect of using PET and PET–CT with FES or FDG as an upfront imaging test for diagnosing ER-positive metastatic breast cancer on the number of performed biopsies and associated costs, compared to the standard clinical work-up[123]; fewer biopsies were performed using the FES-PET–CT strategy at an incremental cost-effectiveness ratio (ICER) of €12,100 ± 3,400 per biopsy avoided. In addition, if PET imaging could be used to select patients for treatment with expensive cancer drugs, such as immune-checkpoint inhibitors, the modality would immediately be cost-effective owing to the cost savings associated with preventing unnecessary expenditure on ineffective treatments and on managing the associated toxicities.

Conclusions

Molecular nuclear imaging has great potential to generate predictive biomarkers and can support early drug development in multiple ways. Unfortunately, the use of molecular nuclear imaging in clinical research and practice remains limited, but could be stimulated by increasing access to tracers, education of multidisciplinary teams, the use of novel trial designs, harmonization of procedures, and establishing initiatives to share and re-analyse data from clinical imaging studies (TABLE 2). In addition, the complexity and costs of molecular imaging studies require close collaboration between academia, pharmaceutical companies, and companies that provide tracers and/or imaging platforms.

Elisabeth G. E. de Vries[1], Laura Kist de Ruijter[1], Marjolijn N. Lub-de Hooge[2,3], Rudi A. Dierckx[3], Sjoerd G. Elias[4] and Sjoukje F. Oosting[1]*

[1]Department of Medical Oncology, University Medical Center Groningen, University of Groningen, Groningen, Netherlands.

[2]Department of Clinical Pharmacy and Pharmacology, University Medical Center Groningen, University of Groningen, Groningen, Netherlands.

[3]Department of Nuclear Medicine and Molecular Imaging, University Medical Center Groningen, University of Groningen, Groningen, Netherlands.

[4]Department of Epidemiology, Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht University, Utrecht, Netherlands.

**e-mail: e.de.vries@umcg.nl*

https://doi.org/10.1038/s41571-018-0123-y

Published online: 27 November 2018

1. Cottingham, M. D., Kalbaugh, C. A. & Fisher, J. A. Tracking the pharmaceutical pipeline: clinical trials and global disease burden. *Clin. Transl Sci.* **7**, 297–299 (2014).
2. GlobalData Healthcare. Number of pipeline drugs in the US by therapy area. *Pharmaceutical Technology* https://www.pharmaceutical-technology.com/comment/number-pipeline-drugs-us-therapy-area/ (2017).
3. Cherny, N. I. et al. ESMO-magnitude of clinical benefit Scale version 1.1. *Ann. Oncol.* **28**, 2340–2366 (2017).
4. Schnipper, L. E. et al. Updating the American Society of Clinical Oncology value framework: revisions and reflections in response to comments received. *J. Clin. Oncol.* **34**, 2925–2934 (2016).
5. Vivot, A. et al. Clinical benefit, price and approval characteristics of FDA-approved new drugs for treating advanced solid cancer, 2000–2015. *Ann. Oncol.* **28**, 1111–1116 (2017).
6. Tibau, A. et al. Magnitude of clinical benefit of cancer drugs approved by the US Food and Drug Administration. *J. Natl Cancer Inst.* **110**, 486–492 (2018).
7. Jaffee, E. M. et al. Future cancer research priorities in the USA: a Lancet Oncology Commission. *Lancet Oncol.* **18**, e653–e706 (2017).
8. Wolff, A. C. et al. Human epidermal growth factor receptor 2 testing in breast cancer: American Society of Clinical Oncology/College of American Pathologists clinical practice guideline focused update. *J. Clin. Oncol.* **36**, 2105–2122 (2018).
9. Menzies, A. M. & Long, G. V. Systemic treatment for BRAF-mutant melanoma: where do we go next? *Lancet Oncol.* **15**, e371–e381 (2014).
10. Hanahan, D. & Weinberg, R. A. Hallmarks of cancer: the next generation. *Cell* **144**, 646–674 (2011).

11. Alam, I. S., Arshad, M. A., Nguyen, Q. D. & Aboagye, E. O. Radiopharmaceuticals as probes to characterize tumour tissue. *Eur. J. Nucl. Med. Mol. Imag.* **42**, 537–561 (2015).
12. O'Connor, J. P. et al. Imaging biomarker roadmap for cancer studies. *Nat. Rev. Clin. Oncol.* **14**, 169–186 (2017).
13. National Comprehensive Cancer Network. NCCN Clinical Practice Guidelines in Oncology (NCCN Guidelines®): Non-Small Cell Lung Cancer. *National Comprehensive Cancer Network* https://www.nccn.org/professionals/physician_gls/pdf/nscl.pdf (2018).
14. Lordick, F. et al. Oesophageal cancer: ESMO clinical practice guidelines for diagnosis, treatment and follow-up. *Ann. Oncol.* **27** (Suppl. 5), v50–v57 (2016).
15. Liu, Y. et al. A risk management approach for imaging biomarker-driven clinical trials in oncology. *Lancet Oncol.* **16**, e622–628 (2015).
16. Tang, J., Shalabi, A. & Hubbard-Lucey, V. M. Comprehensive analysis of the clinical immuno-oncology landscape. *Ann. Oncol.* **29**, 84–91 (2018).
17. Pharmaceutical Research and Manufacturers of America. Medicines in development for immuno-oncology 2017 report. *PhRMA* http://www.phrma.org/medicines-in-development-immuno-oncology (2017).
18. Eisenhauer, E. A. et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). *Eur. J. Cancer* **45**, 228–247 (2009).
19. Meignan, M., Gallamini, A. & Haioun, C. Report on the first international workshop on interim-PET scan in lymphoma. *Leuk. Lymphoma* **50**, 1257–1260 (2009).
20. Barrington, S. F. et al. Role of imaging in the staging and response assessment of lymphoma: consensus of the International Conference on Malignant Lymphomas Imaging Working Group. *J. Clin. Oncol.* **32**, 3048–3058 (2014).
21. Food and Drug Administration. Guidance for industry: clinical trial endpoints for the approval of cancer drugs and biologics. *FDA* https://www.fda.gov/downloads/drugsGuidanceComplianceRegulatoyInformation/Guidance/UCM071590.pdf (2017).
22. Schwartz, L. H. et al. RECIST 1.1-update and clarification: from the RECIST committee. *Eur. J. Cancer* **62**, 132–137 (2016).
23. Janku, F. et al. Outcomes of phase II clinical trials with single-agent therapies in advanced/metastatic non-small cell lung cancer published between 2000 and 2009. *Clin. Cancer Res.* **18**, 6356–6363 (2012).
24. Therasse, P. et al. New guidelines to evaluate the response to treatment in solid tumors. European Organization for Research and Treatment of Cancer, National Cancer Institute of the United States, National Cancer Institute of Canada. *J. Natl Cancer Inst.* **92**, 205–216 (2000).
25. Litière, S. et al. Validation of RECIST 1.1 for use with cytotoxic agents and targeted cancer agents (TCA): results of a RECIST Working Group analysis of a 50 clinical trials pooled individual patient database [abstract]. *J. Clin. Oncol.* **35** (Suppl. 15), 2534 (2017).
26. Wolchok, J. D. et al. Guidelines for the evaluation of immune therapy activity in solid tumors: immune-related response criteria. *Clin. Cancer Res.* **15**, 7412–7420 (2009).
27. Seymour, L. et al. RECIST working group. iRECIST: guidelines for response criteria for use in trials testing immunotherapeutics. *Lancet Oncol.* **18**, e143–e152 (2017).
28. Wahl, R. L., Jacene, H., Kasamon, Y. & Lodge, M. A. From RECIST to PERCIST: Evolving considerations for PET response criteria in solid tumors. *J. Nucl. Med.* **50** (Suppl. 1), 122S–150S (2009).
29. Choi, H. et al. Correlation of computed tomography and positron emission tomography in patients with metastatic gastrointestinal stromal tumor treated at a single institution with imatinib mesylate: proposal of new computed tomography response criteria. *J. Clin. Oncol.* **25**, 1753–1759 (2007).
30. Lambin, P. et al. Radiomics: extracting more information from medical images using advanced feature analysis. *Eur. J. Cancer* **48**, 441–446 (2012).
31. Kumar, V. et al. Radiomics: the process and the challenges. *Magn. Reson. Imag.* **30**, 1234–1248 (2012).
32. Aerts, H. J. W. L. Data science in radiology: a path forward. *Clin. Cancer Res.* **24**, 532–534 (2018).
33. O'Connor, J. P. et al. Imaging intratumor heterogeneity: role in therapy response, resistance, and clinical outcome. *Clin. Cancer Res.* **21**, 249–257 (2015).
34. Lovinfosse, P. et al. FDG PET/CT radiomics for predicting the outcome of locally advanced rectal cancer. *Eur. J. Nucl. Med. Mol. Imag.* **45**, 365–375 (2018).
35. Li, H. et al. MR imaging radiomics signatures for predicting the risk of breast cancer recurrence as given by research versions of MammaPrint, Oncotype DX, and PAM50 gene assays. *Radiology* **281**, 382–391 (2016).
36. Colen, R. R. et al. Radiomics to predict immunotherapy-induced pneumonitis: proof of concept. *Invest. New Drugs* **36**, 601–607 (2018).
37. Campbell, M. G. et al. Bridging the gaps in ^{18}F PET tracer development. *Nat. Chem.* **9**, 1–3 (2016).
38. Decristoforo, C., Penuelas, I., Patt, M. & Todde, S. European regulations for the introduction of novel radiopharmaceuticals in the clinical setting. *Q. J. Nucl. Med. Mol. Imag.* **61**, 135–144 (2017).
39. Lange, R. et al. Untangling the web of European regulations for the preparation of unlicensed radiopharmaceuticals: a concise overview and practical guidance for a risk-based approach. *Nucl. Med. Commun.* **36**, 412–422 (2015).
40. van Dongen, G. A., Poot, A. J. & Vugts, D. J. PET imaging with radiolabeled antibodies and tyrosine kinase inhibitors: immuno-PET and TKI-PET. *Tumor Biol.* **33**, 607–615 (2012).
41. Weber, B. et al. Erlotinib accumulation in brain metastases from non-small cell lung cancer: visualization by positron emission tomography in a patient harboring a mutation in the epidermal growth factor receptor. *J. Thorac. Oncol.* **6**, 1287–1289 (2011).
42. Saleem, A. et al. Lapatinib access into normal brain and brain metastases in patients with Her-2 overexpressing breast cancer. *EJNMMI Res.* **5**, 30 (2015).
43. Sun, X. et al. A PET imaging approach for determining EGFR mutation status for improved lung cancer patient management. *Sci. Transl Med.* **10**, eaan8840 (2018).
44. Roesch, F. & Riss, P. J. The renaissance of the ^{68}Ge/^{68}Ga radionuclide generator initiates new developments in ^{68}Ga radiopharmaceutical chemistry. *Curr. Top. Med. Chem.* **10**, 1633–1668 (2010).
45. Gebhart, G. et al. Molecular imaging as a tool to investigate heterogeneity of advanced HER2-positive breast cancer and to predict patient outcome under trastuzumab emtansine (T-DM1): the ZEPHIR trial. *Ann. Oncol.* **27**, 619–624 (2016).
46. Bensch, F. et al. ^{89}Zr-atezolizumab imaging as non-invasive approach to assess clinical response to PD-L1 blockade in cancer. *Nat. Med.* https://doi.org/10.1038/s41591-018-0255-8 (2018).
47. Oosting, S. F. et al. ^{89}Zr-bevacizumab PET visualizes heterogeneous tracer accumulation in tumor lesions of renal cell carcinoma patients and differential effects of antiangiogenic treatment. *J. Nucl. Med.* **56**, 63–69 (2015).
48. Mestel, R. Cancer: imaging with antibodies. *Nature* **543**, 743–746 (2017).
49. Gaykema, S. B. et al. ^{89}Zr-trastuzumab and ^{89}Zr-bevacizumab PET to evaluate the effect of the HSP90 inhibitor NVP-AUY922 in metastatic breast cancer patients. *Clin. Cancer Res.* **20**, 3945–3954 (2014).
50. Sandberg, D. et al. Intra-image referencing for simplified assessment of HER2-expression in breast cancer metastases using the affibody molecule ABY-025 with PET and SPECT. *Eur. J. Nucl. Med. Mol. Imag.* **44**, 1337–1346 (2017).
51. Orlova, A., Wållberg, H., Stone-Elander, S. & Tolmachev, V. On the selection of a tracer for PET imaging of HER2-expressing tumors: direct comparison of a ^{124}I-labeled affibody molecule and trastuzumab in a murine xenograft model. *J. Nucl. Med.* **50**, 417–425 (2009).
52. Malmberg, J., Sandström, M., Wester, K., Tolmachev, V. & Orlova, A. Comparative biodistribution of imaging agents for in vivo molecular profiling of disseminated prostate cancer in mice bearing prostate cancer xenografts: focus on 111In- and ^{125}I-labeled anti-HER2 humanized monoclonal trastuzumab and ABY-025 affibody. *Nucl. Med. Biol.* **38**, 1093–1102 (2011).
53. Waaijer, S. J. H. et al. Molecular imaging of radiolabeled bispecific T cell engager ^{89}Zr-AMG211 targeting CEA-positive tumors. *Clin. Cancer Res.* **24**, 4988–4996 (2018).
54. Hanessian, S. Academic-industrial collaboration: toward the consilience of two solitudes. *ACS Med. Chem. Lett.* **7**, 6–9 (2015).
55. Bernard-Gauthier, V., Collier, T. L., Liang, S. H. & Vasdev, N. Discovery of PET radiopharmaceuticals at the academic-industry interface. *Drug Discov. Today Technol.* **25**, 19–26 (2017).
56. Saleem, A., Murphy, P., Plisson, C. & Lahn, M. Why are we failing to implement imaging studies with radiolabelled new molecular entities in early oncology drug development? *ScientificWorldJournal* **2014**, 269605 (2014).
57. European Society of Radiology (ESR). ESR statement on the stepwise development of imaging biomarkers. *Insights Imag.* **4**, 147–152 (2013).
58. Biomarker, Imaging, and Quality of Life Studies Funding Program. Biomarker study evaluation guidelines. *cancer.gov* https://www.cancer.gov/about-nci/organization/ccct/funding/biqsfp/2018-biomarker-study-eval-guide.pdf (2018).
59. Weinreb, J. C. et al. PI-RADS prostate imaging — reporting and data system: 2015, version 2. *Eur. Urol.* **69**, 16–40 (2016).
60. Sullivan, D. C. et al. Metrology standards for quantitative imaging biomarkers. *Radiology* **277**, 813–825 (2015).
61. Aide, N. et al. EANM/EARL harmonization strategies in PET quantification: from daily practice to multicenter oncological studies. *Eur. J. Nucl. Med. Mol. Imag.* **44**, 17–25 (2017).
62. Makris, N. E. et al. Multicenter harmonization of ^{89}Zr PET/CT performance. *J. Nucl. Med.* **55**, 264–267 (2014).
63. U.S. Food and Drug Administration. CDER biomarker qualification program. *FDA* https://www.fda.gov/Drugs/DevelopmentApprovalProcess/DrugDevelopmentToolsQualificationProgram/BiomarkerQualificationProgram/default.htm (2018).
64. McShane, L. M. et al. REporting recommendations for tumor MARKer prognostic studies (REMARK). *Nat. Clin. Pract. Oncol.* **2**, 416–422 (2005).
65. Caron, J. E., March, J. K., Cohen, M. B. & Schmidt, R. L. A survey of the prevalence and impact of reporting guideline endorsement in pathology journals. *Am. J. Clin. Pathol.* **148**, 314–322 (2017).
66. Sekula, P., Mallett, S., Altman, D. G. & Sauerbrei, W. Did the reporting of prognostic studies of tumour markers improve since the introduction of REMARK guideline? A comparison of reporting in published articles. *PLOS ONE* **12**, e0178531 (2017).
67. Strosberg, J. et al. Phase 3 trial of ^{177}Lu-dotatate for midgut neuroendocrine tumors. *N. Engl. J. Med.* **376**, 125–135 (2017).
68. Kwekkeboom, D. et al. Treatment with the radiolabeled somatostatin analog [^{177}Lu-DOTA^{0},Tyr^{3}] octreotate: toxicity, efficacy, and survival. *J. Clin. Oncol.* **26**, 2124–2130 (2008).
69. Vaupel, P., Mayer, A. & Hockel, M. Tumour hypoxia and malignant progression. *Meth. Enzymol.* **381**, 335–354 (2004).
70. Lopci, E. et al. PET radiopharmaceuticals for imaging of tumor hypoxia: a review of the evidence. *Am. J. Nucl. Med. Mol. Imag.* **4**, 365–384 (2014).
71. Welz, S. et al. Prognostic value of dynamic hypoxia PET in head and neck cancer: results from a planned interim analysis of a randomized phase II hypoxia-image guided dose escalation trial. *Radiother. Oncol.* **124**, 526–532 (2017).
72. US National Library of Medicine. *ClinicalTrials.gov* https://clinicaltrials.gov/ct2/show/NCT02352792?term=NCT02352792&rank=1 (2015).
73. Lee, N. et al. Strategy of using intratreatment hypoxia imaging to selectively and safely guide radiation dose de-escalation concurrent with chemotherapy for locoregionally advanced human papillomavirus related oropharyngeal carcinoma. *Int. J. Radiat. Oncol. Biol. Phys.* **96**, 9–17 (2016).
74. US National Library of Medicine. *ClinicalTrials.gov* https://clinicaltrials.gov/ct2/show/NCT00606294?term=NCT00606294&rank=1 (2018).
75. Lamberts, L. E. et al. Antibody positron emission tomography imaging in anticancer drug development. *J. Clin. Oncol.* **33**, 1491–1504 (2015).
76. Banerji, U. & Workman, P. Critical parameters in targeted drug development: the pharmacological audit trail. *Semin. Oncol.* **43**, 436–445 (2016).
77. van Kruchten, M. et al. Measuring residual estrogen receptor availability during fulvestrant therapy in patients with metastatic breast cancer. *Cancer Discov.* **5**, 72–81 (2015).
78. van der Veldt, A. A. et al. Rapid decrease in delivery of chemotherapy to tumors after anti-VEGF therapy: implications for scheduling of anti-angiogenic drugs. *Cancer Cell* **21**, 82–91 (2012).
79. Ho, A. L. et al. Selumetinib-enhanced radioiodine uptake in advanced thyroid cancer. *N. Engl. J. Med.* **368**, 623–632 (2013).
80. Rothenberg, S. M., McFadden, D. G., Palmer, E. L., Daniels, G. H. & Wirth, L. J. Redifferentation of iodine-refractory BRAF V600E-mutant metastatic

papillary thyroid cancer with dabrafenib. *Clin. Cancer Res.* **21**, 1025–1035 (2015).
81. Dunn, L. et al. Enhancing radioiodine (RAI) incorporation into BRA^{FV600}E-mutant, RAI-refractory thyroid cancer with the BRAF inhibitor vemurafenib: a pilot study [abstract]. *J. Clin. Oncol.* **34** (Suppl. 15), 6099 (2016).
82. McGranahan, N. & Swanton, C. Clonal heterogeneity and tumor evolution: past, present, and the future. *Cell* **168**, 613–628 (2017).
83. Heneweer, C., Holland, J. P., Diviliv, V., Carlin, S. & Lewis, J. S. Magnitude of enhanced permeability and retention effect in tumors with different phenotypes: ^{89}Zr-albumin as a model system. *J. Nucl. Med.* **52**, 625–633 (2011).
84. Dagogo-Jack, I. & Shaw, A. T. Tumour heterogeneity and resistance to cancer therapies. *Nat. Rev. Clin. Oncol.* **15**, 81–94 (2018).
85. Nienhuis, H. H. et al. $^{(18)}$F-fluoroestradiol tumor uptake is heterogeneous and influenced by site of metastasis in breast cancer patients. *J. Nucl. Med.* **59**, 1212–1218 (2018).
86. Fox, J. J. et al. Positron emission tomography/computed tomography-based assessments of androgen receptor expression and glycolytic activity as a prognostic biomarker for metastatic castration-resistant prostate cancer. *JAMA Oncol.* **4**, 217–224 (2017).
87. Ribas, A. & Wolchok, D. Cancer immunotherapy using checkpoint blockade. *Science* **359**, 1350–1355 (2018).
88. Ingram, J. R. et al. PD-L1 is an activation-independent marker of brown adipocytes. *Nat. Commun.* **8**, 647 (2017).
89. Lesniak, W. G. et al. PD-L1 detection in tumors using (^{64}Cu)atezolizumab with PET. *Bioconjug. Chem.* **27**, 2103–2110 (2016).
90. Chatterjee, S. et al. A humanized antibody for imaging immune checkpoint ligand PD-L1 expression in tumors. *Oncotarget* **7**, 10215–10227 (2016).
91. Hettich, M., Braun, F., Bartholomä, M. D., Schirmbeck, R. & Niedermann, G. Resolution PET imaging with therapeutic antibody-based PD-1/PD-L1 checkpoint tracers. *Theranostics* **6**, 1629–1640 (2016).
92. Niemeijer, A. et al. Whole body PD-1 and PD-L1 PET in pts with NSCLC [abstract 1305PD]. *Ann. Oncol.* **28** (Suppl. 5), v460–v496 (2017).
93. Chen, D. S. & Mellman, I. Elements of cancer immunity and the cancer-immune set point. *Nature* **541**, 321–330 (2017).
94. Ishida, Y., Agata, Y., Shibahara, K. & Honjo, T. Induced expression of PD-1, a novel member of the immunoglobulin gene superfamily, upon programmed cell death. *EMBO J.* **11**, 3887–3895 (1992).
95. Sedykh, S. E., Prinz, V. V., Buneva, V. N. & Nevinksky, G. A. Bispecific antibodies: design, therapy, perspectives. *Drug Des. Devel. Ther.* **12**, 195–208 (2018).
96. Carter, P. J. & Lazar, G. A. Next generation antibody drugs: pursuit of the 'high-hanging fruit'. *Nat. Rev. Drug Discov.* **17**, 197–223 (2017).
97. Brinkmann, U. & Kontermann, R. E. The making of bispecific antibodies. *MAbs* **9**, 182–212 (2017).
98. Mandikian, D. et al. Relative target affinities of T cell dependent bispecific antibodies determine biodistribution in a solid tumor mouse model. *Mol. Cancer Ther.* **17**, 776–785 (2018).
99. Glaudemans, A. W. et al. In vivo and in vitro evidence that ^{99m}Tc-HYNIC-interleukin-2 is able to detect T lymphocytes in vulnerable atherosclerotic plaques of the carotid artery. *Eur. J. Nucl. Med. Mol. Imag.* **41**, 1710–1719 (2014).
100. US National Library of Medicine. *ClinicalTrials.gov* https://clinicaltrials.gov/ct2/show/NCT02922283?term=NCT02922283&rank=1 (2016).
101. US National Library of Medicine. *ClinicalTrials.gov* https://clinicaltrials.gov/ct2/show/NCT03107663?term=NCT03107663&rank=1 (2017).
102. Blykers, A. et al. PET imaging of macrophage mannose receptor-expressing macrophages in tumor stroma using ^{18}F-radiolabeled Camelid single-domain antibody fragments. *J. Nucl. Med.* **56**, 1265–1271 (2015).
103. Movahedi, K. et al. Nanobody-based targeting of the macrophage mannose receptor for effective in vivo imaging of tumor-associated macrophages. *Cancer Res.* **72**, 4165–4177 (2012).
104. Jiménez-Sánchez, A. et al. Heterogeneous tumor-immune microenvironments among differentially growing metastases in an ovarian cancer patient. *Cell* **170**, 927–938 (2017).
105. Heinzmann, K., Carter, L. M., Lewis, J. S. & Aboagye, E. O. Multiplexed imaging for diagnosis and therapy. *Nat. Biomed. Eng.* **1**, 697–713 (2017).
106. Lamberts, L. E. et al. Tumor-specific uptake of fluorescent bevacizumab-IRDye800CW microdosing in patients with primary breast cancer: a phase I feasibility study. *Clin. Cancer Res.* **23**, 2730–2741 (2017).
107. Koch, M. et al. Threshold analysis and biodistribution of fluorescently labeled bevacizumab in human breast cancer. *Cancer Res.* **77**, 623–631 (2017).
108. Tunis, S. & Whicher, D. The National Oncologic PET Registry: lessons learned for coverage with evidence development. *J. Am. Coll. Radiol.* **6**, 360–365 (2009).
109. Lindsay, M. J. et al. The National Oncology PET Registry: expanded medicare coverage for PET under coverage with evidence development. *Am. J. Roentgenol.* **188**, 1109–1113 (2007).
110. Bensch, F. et al. Comparative biodistribution analysis across four different ^{89}Zr-monoclonal antibody tracers – the first step towards an imaging warehouse. *Theranostics* **8**, 4295–4304 (2018).
111. Beam, A. L. & Kohane, I. S. Big data and machine learning in health care. *JAMA* **319**, 1317–1318 (2018).
112. Wilhelm-Benartzi, C. S. et al. Challenges and methodology in the incorporation of biomarkers in cancer clinical trials. *Crit. Rev. Oncol. Hematol.* **110**, 49–61 (2017).
113. Hilgers, R. D., Roes, K. & Stallard, N. Directions for new developments on statistical design and analysis of small population group trials. *Orphanet J. Rare Dis.* **11**, 78 (2016).
114. Lambin, P. et al. Radiomics: the bridge between medical imaging and personalized medicine. *Nat. Rev. Clin. Oncol.* **14**, 749–762 (2017).
115. IJzerman, M. J., Koffijberg, H., Fenwick, E. & Krahn, M. Emerging use of early health technology assessment in medical product development: a scoping review of the literature. *Pharmacoeconomics* **35**, 727–740 (2017).
116. Doble, B., Tan, M., Harris, A. & Lorgelly, P. Modeling companion diagnostics in economic evaluations of targeted oncology therapies: systemic review and methodological checklist. *Expert Rev. Mol. Diagn.* **15**, 235–254 (2015).
117. Jong, V. L., Novianti, P. W., Roes, K. C. & Eijkemans, M. J. Selecting a classification function for class prediction with gen expression data. *Bioinformatics* **32**, 1814–1822 (2016).
118. Bajard, A. et al. An in silico approach helped to identify the best experimental design, population, and outcome for future randomized clinical trials. *J. Clin. Epidemiol.* **69**, 125–136 (2016).
119. Ventz, S., Alexander, B. M., Parmigiani, G., Gelber, R. D. & Trippa, L. Designing clinical trials that accept new arms: an example in metastatic breast cancer. *J. Clin. Oncol.* **35**, 3160–3168 (2017).
120. Wallstrom, G., Anderson, K. S. & LaBaer, J. Biomarker discovery for heterogeneous diseases. *Cancer Epidemiol. Biomarkers Prev.* **22**, 747–755 (2013).
121. Kurland, B. F., Doot, R. K., Linden, H. M., Mankoff, D. A. & Kinahan, P. E. Multicenter trials using ^{18}F-fluorodeoxyglucose (FDG) PET to predict chemotherapy response: effects of differential measurement error and bias on power calculations for unselected and enrichment designs. *Clin. Trials* **10**, 886–995 (2013).
122. The ABIM Foundation. PET scans after cancer treatment: when you need them — and when you don't. *choosingwisely* http://www.choosingwisely.org/patient-resources/pet-scans-after-cancer-treatment/ (2014).
123. Koleva-Kolarova, R. G. et al. The value of PET/CT with FES or FDG tracers in metastatic breast cancer: a computer simulation study in ER-positive patients. *Br. J. Cancer* **112**, 1617–1625 (2015).
124. Thorwarth, D. et al. Kinetic analysis of dynamic ^{18}F-fluoromisonidazole PET correlates with radiation treatment outcome in head-and-neck cancer. *BMC Cancer* **5**, 152 (2005).
125. Rajendran, J. G. et al. Tumor hypoxia imaging with [F-18] fluoromisonidazole positron emission tomography in head and neck cancer. *Clin. Cancer Res.* **12**, 5435–5441 (2006).
126. Rischin, D. et al. Prognostic significance of [^{18}F]-misonidazole positron emission tomography-detected tumor hypoxia in patients with advanced head and neck cancer randomly assigned to chemoradiation with or without tirapazamine: a substudy of Trans-Tasman Radiation Oncology Group Study 98.02. *J. Clin. Oncol.* **24**, 2098–2104 (2006).
127. Eschmann, S. M. et al. Hypoxia-imaging with ^{18}F-misonidazole and PET: changes of kinetics during radiotherapy of head-and-neck cancer. *Radiother. Oncol.* **83**, 406–410 (2007).
128. Dirix, P. et al. Dose painting in radiotherapy for head and neck squamous cell carcinoma: value of repeated functional imaging with ^{18}F-FDG PET, ^{18}F-fluoromisonidazole PET, diffusion-weighted MRI, and dynamic contrast-enhanced MRI. *J. Nucl. Med.* **50**, 1020–1027 (2009).
129. Nehmeh, S. A. et al. Reproducibility of intratumor distribution of ^{18}F fluoromisonidazole in head and neck cancer. *Int. J. Radiat. Oncol. Biol. Phys.* **70**, 235–242 (2008).
130. Lee, N. Y. et al. Fluorine-18-labeled fluoromisonidazole positron emission and computed tomography-guided intensity-modulated radiotherapy for head and neck cancer: a feasibility study. *Int. J. Radiat. Oncol. Biol. Phys.* **70**, 12–13 (2008).
131. Lee, N. Y. et al. Prospective trial incorporating pre-/mid-treatment [^{18}F]-misonidazole positron emission tomography for head-and-neck cancer patients undergoing concurrent chemoradiotherapy. *Int. J. Radiat. Oncol. Biol. Phys.* **75**, 101–108 (2009).
132. Kikuchi, M. et al. ^{18}F-fluoromisonidazole positron emission tomography before treatment is a predictor of radiotherapy outcome and survival prognosis in patients with head and neck squamous cell carcinoma. *Ann. Nucl. Med.* **25**, 625–633 (2011).
133. Yamane, T., Kikuchi, M., Shinohara, S. & Senda, M. Reduction of [^{18}F]fluoromisonidazole uptake after neoadjuvant chemotherapy for head and neck squamous cell carcinoma. *Mol. Imag. Biol.* **13**, 227–231 (2011).
134. Okamoto, S. et al. High reproducibility of tumor hypoxia evaluated by ^{18}F-fluoromisonidazole PET for head and neck cancer. *J. Nucl. Med.* **54**, 201–207 (2013).
135. Sato, J. et al. ^{18}F-fluoromisonidazole PET uptake is correlated with hypoxia-inducible factor-1α expression in oral squamous cell carcinoma. *J. Nucl. Med.* **54**, 1060–1065 (2013).
136. Henriques de Figueiredo, B. et al. Potential of [^{18}F]-fluoromisonidazole positron emission tomography for radiotherapy planning in head and neck squamous cell carcinomas. *Strahlenther. Onkol.* **189**, 1015–1019 (2013).
137. Bittner, M. I. et al. Exploratory geographical analysis of hypoxic subvolumes using ^{18}F-MISO-PET imaging in patients with head and neck cancer in the course of primary chemoradiotherapy. *Radiother. Oncol.* **108**, 511–516 (2013).
138. Wiedenmann, N. E. et al. Serial [^{18}F]-fluoromisonidazole PET during radiochemotherapy for locally advanced head and neck cancer and its correlation with outcome. *Radiother. Oncol.* **117**, 113–117 (2015).
139. Okamoto, S. et al. The reoxygenation of hypoxia and the reduction of glucose metabolism in head and neck cancer by fractionated radiotherapy with intensity-modulated radiation therapy. *Eur. J. Nucl. Med. Mol. Imag.* **43**, 2147–2154 (2016).
140. Grkovski, M. et al. Monitoring early response to chemoradiotherapy with ^{18}F-FMISO dynamic PET in head and neck cancer. *Eur. J. Nucl. Med. Mol. Imag.* **44**, 1682–1691 (2017).
141. Boeke, S. et al. Geometric analysis of loco-regional recurrences in relation to pre treatment hypoxia in patients with head and neck cancer. *Acta Oncol.* **56**, 1571–1576 (2017).
142. Zips, D. et al. Exploratory prospective trial of hypoxia-specific PET imaging during radiochemotherapy in patients with locally advanced head-and-neck cancer. *Radiother. Oncol.* **105**, 21–28 (2012).
143. Löck, S. et al. Residual tumour hypoxia in head-and-neck cancer patients undergoing primary radiochemotherapy, final results of a prospective trial on repeat FMISO-PET imaging. *Radiother. Oncol.* **124**, 533–540 (2017).

Acknowledgements
The work of the authors is supported by the European Research Council (ERC) Advanced Grant 2011 OnQView (293445), Innovative Medicines Initiative 2 (IMI2) TRISTAN grant 2016 (116106), and Dutch Cancer Society IMPACT (RUG 2012–5565) and POINTING (RUG 2016–10034) grants. The authors would like to thank C. Divgi and A. Glaudemans for their valuable input into this manuscript.

Author contributions
All authors made a substantial contribution to all aspects of the preparation of this manuscript.

Competing interests
E.G.E.d.V. has served on data safety monitoring boards or advisory boards for Pfizer and Sanofi and has received research funding from Amgen, Chugai, CytomX, Genentech/Roche, Nordic Nanovector, and Regeneron, with all reimbursements and funding made available to her institution (University Medical Center Groningen). The other authors declare no competing interests.

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

Reviewer information
Nature Reviews Clinical Oncology thanks J. O'Connor, N. Devoogdt, and other anonymous reviewer(s) for their contribution to the peer-review of this work.

Supplementary information
Supplementary information is available for this paper at https://doi.org/10.1038/s41571-018-0123-y.

RELATED LINKS

American Board of Internal Medicine (ABIM) foundation Choosing Wisely initiative: http://www.choosingwisely.org/
American College of Radiology Accreditation: https://www.acraccreditation.org/
American College of Radiology Imaging Network (ACRIN): https://www.acrin.org/HOME.aspx
ClinicalTrials.gov database: https://clinicaltrials.gov/
EANM Research Ltd (EARL): http://earl.eanm.org/cms/website.php
ESMO-Magnitude of Clinical Benefit Scale: http://www.esmo.org/Policy/Magnitude-of-Clinical-Benefit-Scale
European Association of Nuclear Medicine (EANM): https://www.eanm.org/
European Imaging Biomarkers Alliance: http://www.eibir.org/scientific-activities/joint-initiatives/eiball/
FDA Biomarker Qualification Program: https://www.fda.gov/Drugs/DevelopmentApprovalProcess/DrugDevelopmentToolsQualificationProgram/BiomarkerQualificationProgram/default.htm
FDA Clinical Trial Imaging Endpoint Process Standards Guidance for Industry: https://www.fda.gov/downloads/drugs/guidances/ucm268555.pdf
Innovative Medicines Initiative (IMI): http://www.imi.europa.eu/
National Biomarker Development Alliance (NBDA): http://nbdabiomarkers.org
National Biomedical Imaging Archive (NBIA): https://imaging.nci.nih.gov/ncia/login.jsf
National Cancer Institute (NCI) Cancer Imaging Program (CIP): https://imaging.cancer.gov/default.htm
NCI Clinical Trials Working Group (CTWG) Biomarker Study Evaluation Guidelines: https://www.cancer.gov/about-nci/organization/ccct/funding/biqsfp/2017-biomarker-study-eval-guide.pdf
NCI Quantitative Imaging Network (QIN): https://imaging.cancer.gov/programs_resources/specialized_initiatives/qin/about/default.htm
Quantitative Imaging Biomarkers Alliance (QIBA): https://www.rsna.org/QIBA/
Quantitative Imaging Data Warehouse (QIDW): https://www.rsna.org/QIDW/
RECIST criteria: http://recist.eortc.org/
Society of Nuclear Medicine and Molecular Imaging (SNMMI): http://www.snmmi.org/Research/ClinicalTrialsNetwork.aspx?ItemNumber=6831
The Cancer Imaging Archive (TCIA): http://www.cancerimagingarchive.net