available at www.sciencedirect.com

ScienceDirect

journal homepage: www.ejconline.com

Current perspective

Two decades at the cross-roads of biology, physics and epidemiology: Lessons learned in [18F-]FDG positron emission tomography in oncology

Saiyada N.F. Rizvi[a], *Emile F. Comans*[a], *Ronald Boellaard*[a], *Harm van Tinteren*[b], *Otto S. Hoekstra*[a*]

[a] *Department of Nuclear Medicine and PET Research, VU University Medical Centre, De Boelelaan 1117, 1081 HV Amsterdam, The Netherlands*

[b] *Comprehensive Cancer Centre Amsterdam, Plesmanlaan 125, 1066 CX Amsterdam, The Netherlands*

ARTICLE INFO

Article history:
Received 1 April 2010
Received in revised form 6 May 2010
Accepted 19 May 2010
Available online 17 June 2010

Keywords:
FDG
Positron emission tomography
PET–CT
Oncology

ABSTRACT

[18F-]FDG PET(-CT) is a primarily quantitative imaging technology that is rapidly gaining ground in clinical oncology; initially for staging and diagnosis, and now increasingly as a biomarker of response to therapy. In spite of 20 years of clinical research, there is discussion about its implementation among clinicians, decision-makers and other parties about its implementation. To some extent, this relates to heterogeneity of the PET results and of trial designs, but also to differences in levels of evidence required by various parties. With PET, biological and quantitative imaging is entering the clinical domain. The current subjective perspective reviews these aspects to help clinicians understand biological and physical elements underlying [18F-]FDG PET to increase the clinical awareness of its potential and limitations.

1. Introduction

Around 1992 whole body PET set foot on clinical ground. Since then, PET has become an important tool in oncological practice. Since 2005, at least 9 comprehensive reports and over 40 systematic reviews on clinical ([18F-]FDG) PET and its specific oncological indications have been published (for an overview, see e.g. [1]). Its role as a 'last resort' tool in complicated clinical situations is undebated but guidelines for the mainstream of patient care tracks can be different (Table 1). In part, this may have to do with local practice and performance of care providers in specific diseases, but it may also relate to different interpretations of the available evidence. In some aspects, the current debate on requirements for imaging biomarkers has similarity with that on the required level of evidence required for PET–CT in practice. In this paper, we will discuss the current state of affairs, with emphasis on how innovations in clinical and PET domains have interacted.

After the appreciation that the Warburg effect could be capitalised with [18F-]FDG-PET, most research focussed on improvement of diagnosis and staging, using visual assessment of whole body images. In this domain, added value of PET mainly relates to its superior target-background contrast of relevant pathology compared to CT. Later, biological

* *Corresponding author:* Address: Department of Nuclear Medicine and PET Research (4F011), VU University Medical Centre, De Boelelaan 1117, 1081 HV Amsterdam, The Netherlands. Tel.: +31 20 4444214; fax: +31 20 4443090.
E-mail address: os.hoekstra@vumc.nl (O.S. Hoekstra).
0959-8049/$ - see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.ejca.2010.05.018

Table 1 – Recent recommendations and assessments of [18F-]FDG PET in oncology (for methodological details see Appendix).

Tumour type	Facey 2007		KCE 2009		Fletcher 2008		CMS 2009	
	D&S	R&M	D&S	R&M	D&S&R	Benefit	D&S	R&M
Colorectal	S3	R1M3	S1	1	D2S1R1	D3S2R1	1	1
Oesophagus	D3S1	M3	S1	M3	D2S1R2	S1	1	1
Head and neck	1	R1	S1	1	D2S1R1	D3S1R1	1	1
Lymphoma	S1	1	S1	R3M1	D2S3R1	S1R1	1	1
NSCLC	S1	1	S1	3	D2S1R2	S1R3	1	1
Ovary			D3	1			1	1
Brain				3			1	CED
Cervix			S1	R3			D2S1	1
Small cell lung	D3S1	R2			D2S3R2	S3	1	CED
Soft tissue sarcoma					2	D3S3	1	CED
Pancreas			D1[e] S3	2	D1[d]S2R2	D1	1	CED
Testes			2	2			1	CED
Breast	S3	R3	2	2	D2S2R1	D3S2R1	D2S1[a]	1
Melanoma	S3	R3	S1	R1M3	D2S1R2	S1	D1S2[b]	1
Prostate			2	2			1	CED
Thyroid		R1[c]			D2S2 R1[c]	R1[c]	1	1/CED[c]
All other solid							1	CED
Myeloma							1	1
All other cancers not listed							CED	CED
Unknown primary			D1		D1S2R2	D1		
GIST				M1				

D, diagnosis; S, staging; R, restaging/recurrence; M, monitoring; CED, coverage with evidence development.

D & S & R & M: 1 = recommended; 2 = not recommended; 3 = uncertain.Benefit: 1 = yes; 2 = no; 3 = uncertain.

[a] Not covered for initial staging and/or staging of lymph nodes. Covered for initial staging of metastatic disease.

[b] Not covered for initial staging of regional lymph nodes.

[c] Covered for subsequent treatment strategy of recurrent or residual thyroid cancer of follicular cell origin previously treated by thyroidectomy and radioiodine ablation and have a serum thyroglobulin >10 ng/ml and have a negative I-131 whole body scan. All other indications for subsequent treatment strategy are CED.

[d] Only for CT inconclusive cases.

[e] Differentiation between chronic pancreatitis and pancreatic cancer and between benign and malignant pancreatic cysts.

features of [18F-]FDG-PET came into play using it as a predictive marker, to serve as a read-out of response to therapy. In a further attempt to help with stratification of patients, [18F-]FDG-PET also entered the 'prognostic' arena, again stimulated by biomolecular insights suggesting that glucose metabolism was linked to some key-factors of biological aggressiveness.

2. Technical developments

Since 2000, hybrid PET–CT scanners have become the standard technology, and this development made PET accessible for standard patient care. In the Netherlands, the number of PET scanners has increased from 2 to over 35 in a few years time. Many hospitals can afford to run PET–CT scanners by using the scanner partly for PET(-CT), partly for CT only. PET–CT has several advantages over standard PET: acquisition time and [18F-]FDG-dose (one of the main cost-drivers) have been reduced by at least 50%. Most scanners provide excellent CT quality without compromising the quality of PET. There are only few well-designed studies evaluating the diagnostic benefit of combined PET–CT versus PET alone with co-reading of the CT scan. However, the issue has become an academic one since scanner vendors simply stopped producing PET-only scanners for clinical use. Conceptually, one would expect that especially specificity and localisation performance would benefit from PET–CT fusion. Time-of-flight technology might have some impact on sensitivity by improving signal-to-noise ratios, especially in obese patients. With combined PET and CT readings, the challenge is to interpret the scans with knowledge of strengths and limitations of either method.

3. On detection limits and specificity

[18F-]FDG is not tumour specific. It has a favourable biodistribution with low background activity except in brain and the urinary tract. Detectability is a function of the level of cellular tracer uptake and cell density versus background. Here is where physics and biology meet.

From the former perspective, like with other imaging technologies, one has to account for partial volume effects with small lesions (i.e. diameters < twice the post-reconstruction image resolution): these effects will tend to underestimate the actual [18F-]FDG-uptake/gram tissue. Apparent target-background contrast can further be reduced by blurring of the target signal due to e.g. respiratory motion (such as with lesions in the basal lung fields). Together with less important effects of somewhat higher background activity (cf. the zones of West[2]) and scatter from the adjacent liver, these effects explain why small lung metastases are less well depicted by PET than CT. Obviously, this does not compromise the positive predictive value of PET in such lesions.

On the biological side, it has been shown that different tumour types have different [18F-]FDG affinity. Several typically have low affinity, like well-differentiated thyroid cancer, low grade glioma; others have variable uptake, like renal cell, pancreatic and broncho-alveolar cancer (BAC), whereas cancers of lung, head and neck, cervix, colorectum and melanoma typically are quite [18F-]FDG avid. This variability can be related to lack of glut-transporters on the cell membrane (BAC), predominantly mucinous content of cell cytoplasm (mucinous pancreatic cancers), but mostly this is not readily explained. On average, [18F-]FDG uptake in breast cancer is about half of that in non-small cell lung cancer (NSCLC[3]). However, it is clear that heterogeneity prevails within breast cancers, with extremely high uptake in many patients. Certainly, there are associations between differentiation grade, histological subtype, proliferation rates and [18F-]FDG uptake, but rarely this is a one to one relationship.[4–6] In breast cancer, lobular cancers will generally have lower uptake than ductal ones; in lung cancer, highest uptake is observed in squamous cell type, with larger variation in adenocarcinomas. Similar, and again empirical, evidence of inverse relations between [18F-]FDG avidity and outcome has emerged from thyroid and prostate cancers. Even though the pharmacokinetic model of [18F-]FDG uptake is relatively simple, drivers of glucose metabolism are multifactorial and [18F-]FDG uptake likely also carries information at the tumour-host level (e.g. hypoxia).[6,7] Even though our understanding at this level may not be complete, there is phenomenological proof of principle that [18F-]FDG uptake carries prognostically relevant information.[8] Confounders, such as underestimated uptake due to partial volume effects in small tumours,[9] need to be identified, and again, this is where physics and biology meet.

4. Clinical applications

PET has been investigated for use in diagnosis, staging, restaging and follow-up. A general rule is that the yield of PET will depend on the level of [18F-]FDG uptake by the primary tumour: low uptake in the primary tumour will unlikely lead to high sensitivity in lymph nodes and distant metastases. Based on the current evidence (see below), there is no generally applicable quantitative translation for this qualitative statement, even though single centre studies have shown proof of principle.[10,11] Obviously, in the context of variable biology, it is not realistic to expect sharply defined detection limits for PET; however, as a rule of thumb, it has become clear that for [18F-]FDG avid tumours, PET detects tumour below the centimetre threshold, usually in the 5–10 mm range. Initially, data were presented suggesting similar negative predictive values for PET and sentinel node biopsy in breast cancer,[12] but this overestimation likely related to the use of different, less sophisticated histopathological techniques in the PET studies.

5. Levels of evidence

During the first decade of clinical PET many studies in many types of cancer focussed on defining accuracy: to characterise solitary pulmonary nodules, breast tumours, lymph node staging and detection of distant metastases. During the same time frame, there was increasing awareness[13] that top-level accuracy studies are not trivial: e.g. in the pivotal study on the accuracy of PET in mediastinal lymph node staging in NSCLC,[14] distant metastases apparent on PET had to be ignored to meet such demands. Moreover, the apparent simplicity of PET interpretation led to imprecise or even lacking criteria of test positivity.[15]

It has become clear that setting the indication for tests that screen for locoregional and distant metastases should incorporate some estimate of pre-test probability of such metastases. Too often, accuracy studies involved almost the full spectrum of patients with a specific tumour rather than focussing on the ones who are really at risk of disseminated disease. Apart from waste of resources, the dark side of this moon is that one has to deal with (the clinical consequences of) considerable false positivity rates in low-prevalent conditions. Not unexpectedly, it has been shown that PET is useful for screening for (especially distant) metastases in T3–4 oesophageal, cervix and locally advanced breast cancer, and in patients with head and neck cancer at considerable risk of distant recurrence.[16,17]

However, the buck does not always stop with accuracy studies: the challenge is to produce reliable data on the impact on management in daily practice and on patient outcomes. For example, the accuracy of [18F-]FDG-PET to characterise primary lung lesions $\geqslant$8–10 mm is well established.[18] In clinical practice, pre- and post-test probabilities should be weighed for effectiveness and costs. This implies that estimation of pre-test (here: pre-PET) probabilities should be in place, and there are internationally validated algorithms to do so.[19–21] Apart from endemic diseases lowering the specificity of [18F-]FDG-PET (e.g. histoplasmosis and tuberculosis), local expertise with invasive procedures may also have to be considered when designing guidelines. So far, this is still the basic accuracy domain. However, the ultimate challenge for the clinician in charge of individual patients is to have a notion of the acceptable level of false negativity of the entire diagnostic trajectory. This involves totally different issues such as the quality of life of patients and physicians in continued situations of anxiety and uncertainty. To some extent, comprehensive algorithms can help to quantify residual uncertainties and help to set the indication for additional tests. The 2007 ACCP guideline recommends that [18F-]FDG-PET be performed to characterise solitary pulmonary nodules in patients with radiologically indeterminate nodules of at least 8 mm, with a 5–60% pre-test probability of malignancy.[22]

Recognising that this approach should not be seen as a panacea, we discuss below the process of evidence building for [18F-]FDG-PET to select NSCLC patients for surgery. This selection is a multidisciplinary step-by-step process, each diagnostic test contributing information. In accuracy studies test results are usually expressed in a 2×2 table: the test is either positive or negative and the disease is present or absent. In reality, however, unless the test result is pathognomonic (e.g. multiple bone metastases at bone scintigraphy), test results are less 'black-and-white' (a radiologist who had been forced to express himself in a language that was not his own, expressed it as 'Sitting on the fence – a radiologist's stock in trade – necessitates using words for balance,

weighing diagnostic probabilities, and leaning toward the heavier side. But because I couldn't use the subjunctive mood, I was forced into the realm of apparent diagnostic certainty'[23]).

Since tests (and the people who interpret them[24]) are imperfect, and when clinical decision-making is multifactorial, improved diagnostic accuracy provided by one component of a diagnostic test sequence may not necessarily translate into relevant management changes. Verification of decisive PET findings remains crucial in cases without pathognomonic metastatic spread at PET,[25,26] and the success or failure to do so will affect the ultimate yield of the technique.

Our strategy to evaluate PET in NSCLC staging involved:

(1) *measurement of the residual inefficiency without PET*: in a 2 years multicentre cohort we found that about 50% of thoracotomies in patients with presumed operable NSCLC were futile[27];
(2) *modelling analysis*[28]: to explore at which stage of the diagnostic process PET might be most cost-effective;
(3) *multicentre randomized controlled clinical trial (RCT)*: demonstrating that PET reduced the clinical problem by 50%[25]. The 'number needed to PET' in order to prevent one futile thoracotomy was 5 (95%CI 3–14); identical estimations were found in a recently published Danish trial, now with PET–CT.[29] That objective criteria for endpoints and clinical consensus about management of patients after diagnosis are important is illustrated by another RCT[30] in which the surgeon's decision was taken as the gold standard without validation against follow-up information (e.g. early recurrence);
(4) *cost-effectiveness*: the RCT provided direct data for comparison of costs in relation to diagnosis and therapy. Scenario analyses included various hospital settings, tracer accessibility, sensitivity analyses on cost and accuracy of PET and scenarios for PET usage[31] and showed that the results were robustly in favour of PET;
(5) *population-based analysis*: data from the Regional Cancer Centre Registry of the same region in which the initial RCT had been conducted showed that after implementation of a guideline including PET the number lung resections had dropped with an absolute 20% (corresponding to the predicted 50% reduction in unnecessary thoracotomies.[32])

Even though this circle of evidence building was completed in relatively short notice, it is clear that this is impossible to achieve for all potential indications. The window of opportunity is limited: our first trial included 65% of the eligible patients but this accrual rate dropped to only 20% in a second.[33] Apparently, PET had diffused into practice impeding experimental settings which included a 'non-PET'-arm. Results from such trials are especially useful and lasting if management upon staging remains undisputed for some time. A recent multicentre RCT revealed a 38% relative risk reduction of futile laparotomies in patients with presumed resectable liver metastases of colorectal cancer.[34] Since the start of this trial, the clinical approach towards metastasized colorectal cancer is changing, so that the definition of 'futile procedures' may have to be recalibrated over time. In other indications, such as preoperative oesophageal,[35] head-and-neck cancer[17,36] and melanoma staging,[37] nation-wide multicentre observational study designs with cost-scenarios were applied.

A 'clinical value' study is an alternative approach to acquire evidence on impact on management in daily practice beyond the level of accuracy. This approach has been used for [18F-]FDG-PET[38–40] and more recently on a much larger scale, for PET–CT.[41] In the latter prospective US survey of almost 23,000 PET scans in a broad spectrum of oncological cases, physicians changed their intended management in 36.5% (95% CI, 35.9–37.2) after PET. In retrospect, this was comparable to our experience with 600 consecutive patients in 1997.[42] Even though the strength of such analysis is also its weakness (real life perception versus subjectivity), it provides at least a flavour of the impact of PET in daily practice. Apart from diagnosis and staging at primary presentation, PET has also been applied during follow-up of patients, such as with rising serum markers and clinical signs and symptoms suggestive of recurrent disease without anatomical substrate, and as such it can be of major help (Table 1).

In conclusion, evidence is mainly based upon accuracy studies, supported by clinical value – and cost-modelling studies. When new technology has settled and is widely available, this is often as far as it goes. It may not be surprising that various systematic reviews, technology assessment reports and reimbursement agencies (Table 1) provided different estimates of clinical [18F-]FDG-PET(-CT), with discrepancies partly relating to variable appreciations of the evidence, as well as to the required levels of evidence.[1,16,43,44] This situation is not unique for PET but has also been observed with the introduction of CT[45] and MRI.[46]

6. [18F-]FDG-PET as prognostic and predictive biomarker

Metabolic information obtained by [18F-]FDG-PET adds prognostic information within clinical stages[8,47]: high uptake is prognostically unfavourable compared to tumours with lower uptake. The biological rationale has not been fully explained even though our knowledge of signal transduction in cancer has increased.[7] However, while digging deeper into cancer biology, cartoonish perceptions of complex systems may likely prove to be oversimplifications, even if correct at the level of individual cells.

Even though the lack of standardised quantitative PET-procedures impairs meta-analysis of individual studies, the conceptual point has been made. How this information may be combined with other prognostic markers to develop strategies to improve patient outcomes remains to be validated, but a role in studies stratifying patients for adjuvant therapy seems logical.

There is an increasing interest to use [18F-]FDG-PET as a biomarker to assess therapeutic efficacy in clinical trials, drug development and patient management.[48] The clinical need for alternative response read-outs increased since especially with 'targeted' agents size changes prove to dissociate from patient outcomes.[49] There is ample proof of principle that with standard chemotherapy changes of [18F-]FDG uptake reflect therapeutic effect. Most studies were explorative and

tested the hypothesis that PET added predictive value to volume assessments. More recently, prospective validation of proposed quantitative thresholds of response in the same clinical setting has also been performed (e.g. [50,51]). Studies often were performed in neoadjuvant therapy of patients with locally advanced disease in e.g. oesophageal, breast, NSCLC and head and neck cancer, as well as in disseminated NSCLC, using histopathological assessment of surgical specimens and/or survival measures as end-points. The overall impression is that (changes of) the [18F-]FDG signal relate to (changes of) viable tumour load. When measured early during therapy such changes also carry predictive value, which is relevant in case alternative therapeutic strategies prevail and in drug development.

In 1999, and after review of the limited data at that time, consensus recommendations were formulated for a categorical response classification of [18F-]FDG-PET,[52] using semantic classifications similar to the volume-based ones and comprising a combination of quantitative and qualitative PET measures. However, implementation of these recommendations has not been fulfilled to the extent that meta-analyses were possible. The next challenge beyond accuracy is to demonstrate that adding the PET results to the management of patients improves outcomes compared to strategies without PET. Again, this can only be investigated with randomized trials, using prospectively defined criteria of PET responsiveness; studies with such design are now ongoing.[53]

In fact, methodological heterogeneity from various sources (clinical, PET-methodology, PET criteria and clinical end-points[54]) prevails, and as a result meta-analysis is usually impossible. As a result, [18F-]FDG-PET still has a limited role in the new RECIST1.1 system.[55] To reduce heterogeneity, guidelines for interpretation of [18F-]FDG-PET in first-line therapy for malignant lymphoma have been developed,[56] which were published together with new guidelines on clinical response assessment now incorporating PET.[57] For post-therapy evaluation of first-line lymphoma therapy, and e.g. that of GIST and imatinib therapy,[58] visual PET assessment appears sufficient, but with solid tumours, and especially with early response evaluation, quantification is crucial.

PET is a unique quantitative technique but this quantitative power can only become clinically productive with standardised imaging protocols and quality control and assurance. In the past few years, a comprehensive picture has been obtained of factors that may affect the result of quantification,[59,60] and these relate to a spectrum of factors: scanner calibration, patient preparation, duration of the interval between [18F-]FDG injection and scanning, image reconstruction settings, region of interest definition, etc. To quantify the [18F-]FDG signal, most studies use some form of SUV ('Standardized Uptake Value'); basically the ratio of counts in a tumour over the injected dose normalised for volume of tracer distribution. SUV is used instead of the standard PET method of kinetic modelling, because of its simplicity and ability to measure any lesion in one whole body scan. It has correctly been pointed out that the term 'SUV' is an oxymoron[61]; so far it has involved normalisation rather than standardisation. Finally, use of SUV instead of kinetic modelling also requires validation of its underlying assumptions, when applied with the newer 'targeted' agents which may affect [18F-]FDG pharmacokinetics.[60] A generic guideline (e.g. [59]) will assist physicians in properly collecting, analysing and interpreting quantitative PET studies.

On the biological side of the coin, it has become clear that [18F-]FDG is not a generic read-out of response to any therapy per se. In the neoadjuvant setting, PET studies typically follow the metabolic activity in the primary tumour, testing the hypothesis that this also reflects the effect of therapy on microscopic distant metastases. Combining radiotherapy and chemotherapy requires validation of this assumption, since radiotherapy may also add aspecific information to the signal, likely as a function of time (see e.g. [61]). Whether and how this affects the predictive strength of [18F-]FDG-PET is still the subject of study in several tumours. With some tyrosine kinase inhibitors, preclinical studies have shown that [18F-]FDG changes do not always directly reflect the actual viable tumour load, but predominantly indicate that a target of therapy has been hit.[62,63] Nevertheless, such acute changes also provided spectacular and useful predictive information.[58] With early response evaluation in non-Hodgkin's lymphoma patients treated with R-CHOP, it is not clear whether and how e.g. the addition of rituximab compromises its predictive value.[64,65]

Such observations underline the basic principle that PET markers of response require technical as well as biological validation (Fig. 1). In random order, critical elements are (1) exploration and validation of the association of [18F-]FDG change and patient-related outcome measures for a specific intervention, (2) definition and validation of PET-thresholds of responsiveness, (3) assessment of quantitative reproducibility, and (4) the validation of simplified quantitative measures like the Standardized Uptake Value (SUV).[60,66]

7. Indications for PET(-CT)

From the previous sections, it is clear that there is no generally accepted set of guidelines covering all applications of PET in oncology. In our institution, we advocate that PET referrals beyond local guidelines should be screened for appropriateness prior to submission of the referral, in a telephone call between attending clinician and nuclear medicine physician. Here, the

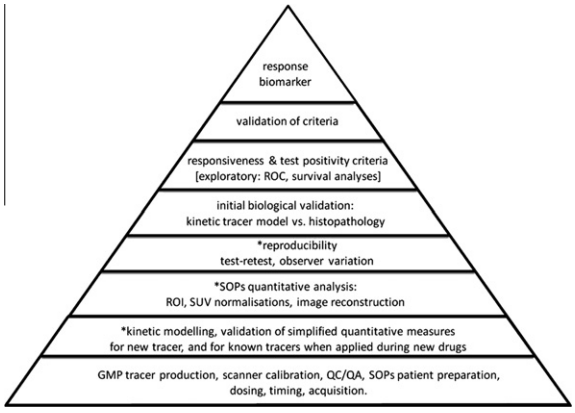

Fig. 1 – The 'Maslow pyramid of needs' for clinical PET to serve as quantitative biomarker of response. *: feasible in same study.

Table 2 – Internal guidelines for oncological [18F-]FDG PET-(CT) referrals.

Indicated[a]	Not indicated[b]	Potential[c]
Staging and diagnosis		
Potentially resectable NSCLCSPN (⩾8 mm, pre-test probability 5–60%)	Colorectal	Susp. recurrent laryngeal cancer after RT
ENT with risk factors[70] for distant metastases	Gastric	
Oesophageal cancer (cT3/4)	Pancreatic	
Malignant lymphoma (except low grade)	Ovarian	
Unknown primary tumour	Renal	
Locally advanced breast cancer (LABC)	Prostate	
Locally advanced cervical cancer	Breast (exc. LABC)	
Potentially resectable liver metastases, colorectal cancer		
Potentially resectable nodal metastases of melanoma		
Rising serum markers (FDG avid tumours) w/o substrate		
		Radiotherapy planning
		Prognostic value SUV
Therapy evaluation		
Malignant lymphoma (except low grade), GIST	Testicular cancer	Predictive value SUV in (early) therapy evaluation of solid tumours
		Interim PET in malignant lymphoma

[a] I: no ballotage required.
[b] N: only after ballotage.
[c] P: investigational (the list is not exhaustive).

key topic is to assess whether the PET result might improve patient management, covering aspects like [18F-]FDG avidity of the primary tumour, size and location of radiologically equivocal lesions, as well as the intended next clinical step if PET would not be performed. In this way, we try to avoid redundant or even counterproductive referrals, as a verbal variant of written educational material.[67] We recognise two categories of PET referrals: 'guidelines' (Table 2) and 'last resort'. The former primarily relates to the most prevalent tumour types in our hospital and is a balance of evidence and acceptance. The last resort category involves any clinical problem with an individual patient that passed the abovementioned 'ballotage' (most often, this concerns (re)staging of patients prior to local therapy with curative intent and/or in patients with significant comorbidity prior to intended intensive therapy). This is an attempt to accommodate the individual variation which is 'part and parcel of clinical practice, and largely what medical practice is all about'.[68]

8. Lessons for new tracers

Beyond improving staging, response evaluation and prognostication, imaging of targeted drugs might prove to be a valuable tool for selection of those patients that most likely will benefit from expensive treatments by measuring drug kinetics in the individual patient, for early selection of promising drugs in drug development and for radiation dosimetry of therapy with radiolabeled targeted drugs. These concepts are now being validated for small molecules as well as for monoclonal antibodies using PET isotopes with physical half-lives matching that of the biological ones of the compounds under study. For antibodies, generic chemistry has been developed so that any antibody can be brought into the clinic within weeks and compatible with GMP demands.[69] Labelling of targeted drugs with positron-emitting isotopes allows for quantitative PET imaging of their distribution and kinetics. The slow kinetics of e.g. monoclonal antibodies requires the use of isotopes with long radioactive half-lives, such as ^{89}Zr, ^{64}Cu or ^{124}I. Here, standardisation of quantification is as important as with [18F-]FDG-PET.

In conclusion, two decades of [18F-]FDG PET have introduced molecular imaging into daily clinical practice. *Visual* assessment of PET(-CT) suffices for many current clinical situations, and it has become a standard test for several indications in many institutes. Apart from NSCLC and lymphoma therapy evaluation, international heterogeneity about guidelines for PET prevails. In the next decades, we expect that *quantitative* PET (and especially its hybrid variants with CT and soon with MRI) will prove to be uniquely equipped for personalised medicine. Sofar, lack of 'quantitative' quality control and – assurance has slowed progress of PET into the field of prognostication and of qualification as a biomarker of response to therapy. Apparently, the switch from the qualitative into the quantitative domain requires a different mind-set of imaging specialists. Recent standardisation efforts have now introduced generally applicable procedures which will help to move the field into the quantitative domain that is essential for truly personalised medicine. An important lesson from the [18F-]FDG experience is that the introduction of new technology requires a structured and multidisciplinary approach of technical and biological validation, to avoid endless cycles of 'trial and error'.

Conflict of interest statement

None declared.

Acknowledgement

The authors thank Prof. Adriaan Lammertsma for help with the design of the biomarker development scheme.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.ejca.2010.05.018.

REFERENCES

1. Vlayen J, Stordeur S, van den Bruel A, Mambourg F, Eyssen M. Positron emissive Tomografie: een update. Health Technology Assessment (HTA). Brussel: Federaal Kenniscentrum voor de Gezondheidszorg (KCE). <http://www.kce.fgov.be/index_nl.aspx?SGREF=12647&CREF=13553> [accessed on March 15th, 2010]; 2009 [Report No.: KCE reports 110A (D/2009/10.27/24)].
2. West J, Dollery C, Naimark A. Distribution of blood flow in isolated lung; relation to vascular and alveolar pressures. *J Appl Physiol* 1964;**19**:713–24.
3. Torizuka T, Zasadny KR, Recker B, Wahl RL. Untreated primary lung and breast cancers: correlation between F-18 FDG kinetic rate constants and findings of in vitro studies. *Radiology* 1998;**207**(3):767–74.
4. Vesselle H, Salskov A, Turcotte E, et al. Relationship between non-small cell lung cancer FDG uptake at PET, tumor histology, and Ki-67 proliferation index. *J Thorac Oncol* 2008;**3**(9):971–8.
5. Avril N, Menzel M, Dose J, et al. Glucose metabolism of breast cancer assessed by 18F-FDG PET: histologic and immunohistochemical tissue analysis. *J Nucl Med* 2001;**42**(1):9–16.
6. Bos R, van der Hoeven JJ, van der Wall E, et al. Biologic correlates of (18)fluorodeoxyglucose uptake in human breast cancer measured by positron emission tomography. *J Clin Oncol* 2002;**20**(2):379–87.
7. Kelloff GJ, Hoffman JM, Johnson B, et al. Progress and promise of FDG-PET imaging for cancer patient management and oncologic drug development. *Clin Cancer Res* 2005;**11**(8):2785–808.
8. Paesmans M, Berghmans T, Dusart M, et al. Primary tumor standardized uptake value measured on fluorodeoxyglucose positron emission tomography is of prognostic value for survival in non-small cell lung cancer: update of a systematic review and meta-analysis by the European Lung Cancer Working Party for the International Association for the Study of Lung Cancer Staging Project. *J Thorac Oncol* 2010;**5**(5):612–9.
9. Vesselle H, Freeman JD, Wiens L, et al. Fluorodeoxyglucose uptake of primary non-small cell lung cancer at positron emission tomography: new contrary data on prognostic role. *Clin Cancer Res* 2007;**13**(11):3255–63.
10. Dooms C, Verbeken E, Stroobants S, et al. Prognostic stratification of stage IIIA-N2 non-small-cell lung cancer after induction chemotherapy: a model based on the combination of morphometric-pathologic response in mediastinal nodes and primary tumor response on serial 18-fluoro-2-deoxy-glucose positron emission tomography. *J Clin Oncol* 2008;**26**(7):1128–34.
11. Cerfolio RJ, Bryant AS. Ratio of the maximum standardized uptake value on FDG-PET of the mediastinal (N2) lymph nodes to the primary tumor may be a universal predictor of nodal malignancy in patients with nonsmall-cell lung cancer. *Ann Thorac Surg* 2007 May;**83**(5):1826–9.
12. Greco M, Crippa F, Agresti R, et al. Axillary lymph node staging in breast cancer by 2-fluoro-2-deoxy-D-glucose-positron emission tomography: clinical evaluation and alternative management. *J Natl Cancer Inst* 2001;**93**(8):630–5.
13. Bossuyt PM, Reitsma JB, Bruns DE, et al. Towards complete and accurate reporting of studies of diagnostic accuracy: the STARD Initiative. *Ann Intern Med* 2003;**138**(1):40–4.
14. Pieterman RM, van Putten JW, Meuzelaar JJ, et al. Preoperative staging of non-small-cell lung cancer with positron-emission tomography [comment]. *N Engl J Med* 2000;**343**(4):254–61.
15. Zijlstra JM, Lindauer-van der Werf G, Hoekstra OS, et al. (18)F-fluoro-deoxyglucose positron emission tomography for post-treatment evaluation of malignant lymphoma: a systematic review. *Haematologica* 2006;**91**(4):522–9.
16. Fletcher JW, Djulbegovic B, Soares HP, et al. Recommendations on the use of 18F-FDG PET in oncology. *J Nucl Med* 2008 Mar;**49**(3):480–508.
17. Uyl-de Groot CA, Senft A, de Bree R, Leemans CR, Hoekstra OS. Chest CT and whole-body 18F-FDG PET are cost-effective in screening for distant metastases in head and neck cancer patients. *J Nucl Med* 2010;**51**(2):176–82.
18. Gould MK, Maclean CC, Kuschner WG, Rydzak CE, Owens DK. Accuracy of positron emission tomography for diagnosis of pulmonary nodules and mass lesions: a meta-analysis. *JAMA* 2001;**285**(7):914–24.
19. Gould MK, Ananth L, Barnett PG. A clinical model to estimate the pretest probability of lung cancer in patients with solitary pulmonary nodules. *Chest* 2007;**131**(2):383–8.
20. Herder GJ, van Tinteren H, Golding RP, et al. Clinical prediction model to characterize pulmonary nodules: Validation and added value of 18F-fluorodeoxyglucose positron emission tomography. *Chest* 2005;**128**(4):2490–6.
21. Swensen SJ, Silverstein MD, Ilstrup DM, Schleck CD, Edell ES. The probability of malignancy in solitary pulmonary nodules. Application to small radiologically indeterminate nodules. *Arch Intern Med* 1997;**157**(8):849–55.
22. Gould MK, Fletcher J, Iannettoni MD, et al. Evaluation of patients with pulmonary nodules: when is it lung cancer?: ACCP evidence-based clinical practice guidelines (2nd edition). *Chest* 2007;**132**(3 Suppl.):108S–30S.
23. Bruzzi JF. The words count – radiology and medical linguistics. *N Engl J Med* 2006;**354**(7):665–7.
24. Robinson PJ. Radiology's Achilles' heel: error and variation in the interpretation of the Rontgen image. *Br J Radiol* 1997;**70**(839):1085–98.
25. van Tinteren H, Hoekstra OS, Smit EF, et al. Effectiveness of positron emission tomography in the preoperative assessment of patients with suspected non-small-cell lung cancer: the PLUS multicentre randomised trial. *Lancet* 2002;**359**(9315):1388–93.
26. Maziak DE, Darling GE, Inculet RI, et al. Positron emission tomography in staging early lung cancer: a randomized trial. *Ann Intern Med* 2009;**151**(4):221–48.
27. Herder GJ, Verboom P, Smit EF, et al. Practice, efficacy and cost of staging suspected non-small cell lung cancer: a retrospective study in two Dutch hospitals. *Thorax* 2002;**57**(1):11–4.
28. Verboom P, Herder GJ, Hoekstra OS, et al. Staging of non-small-cell lung cancer and application of FDG-PET. A cost modeling approach. *Int J Technol Assess Health Care* 2002;**18**(3):576–85.
29. Fischer B, Lassen U, Mortensen J, et al. Preoperative staging of lung cancer with combined PET–CT. *N Engl J Med* 2009;**361**(1):32–9.
30. Viney RC, Boyer MJ, King MT, et al. Randomized controlled trial of the role of positron emission tomography in the management of stage I and II non-small-cell lung cancer. *J Clin Oncol* 2004;**22**(12):2357–62.
31. Verboom P, van Tinteren H, Hoekstra OS, et al. Cost-effectiveness of FDG-PET in staging non-small cell lung cancer: the PLUS study. *Eur J Nucl Med Mol Imaging* 2003;**30**(11):1444–9.

32. van Tinteren H, Hoekstra OS, Smit EF, Boers M. The implementation of PET in non-small-cell lung cancer in the Netherlands. *Clin Oncol (R Coll Radiol)* 2006;**18**(2):156–7.
33. Herder GJ, Kramer H, Hoekstra OS, et al. Traditional versus up-front [18F] fluorodeoxyglucose-positron emission tomography staging of non-small-cell lung cancer: A Dutch Cooperative Randomized Study. *J Clin Oncol* 2006;**24**(12):1800–6.
34. Ruers TJ, Wiering B, van der Sijp JR, et al. Improved selection of patients for hepatic surgery of colorectal liver metastases with (18)F-FDG PET: a randomized study. *J Nucl Med* 2009;**50**(7):1036–41.
35. van Westreenen HL, Westerterp M, Sloof GW, et al. Limited additional value of positron emission tomography in staging oesophageal cancer. *Br J Surg* 2007;**94**(12):1515–20.
36. Senft A, de Bree R, Hoekstra OS, et al. Screening for distant metastases in head and neck cancer patients by chest CT or whole body FDG-PET: a prospective multicenter trial. *Radiother Oncol* 2008;**87**(2):221–9.
37. Bastiaannet E, Wobbes T, Hoekstra OS, et al. Prospective comparison of [18F]fluorodeoxyglucose positron emission tomography and computed tomography in patients with melanoma with palpable lymph node metastases: diagnostic accuracy and impact on treatment. *J Clin Oncol* 2009;**27**(28):4774–80.
38. Herder GJ, van Tinteren H, Comans EF, et al. Prospective use of serial questionnaires to evaluate the therapeutic efficacy of 18F-fluorodeoxyglucose (FDG) positron emission tomography (PET) in suspected lung cancer. *Thorax* 2003;**58**(1):47–51.
39. Kalff V, Hicks RJ, MacManus MP, et al. Clinical impact of (18)F fluorodeoxyglucose positron emission tomography in patients with non-small-cell lung cancer: a prospective study. *J Clin Oncol* 2001;**19**(1):111–8.
40. Hillner BE, Tunuguntla R, Fratkin M. Clinical decisions associated with positron emission tomography in a prospective cohort of patients with suspected or known cancer at one United States center. *J Clin Oncol* 2004;**22**(20):4147–56.
41. Hillner BE, Siegel BA, Liu D, et al. Impact of positron emission tomography/computed tomography and positron emission tomography (PET) alone on expected management of patients with cancer: initial results from the National Oncologic PET Registry. *J Clin Oncol* 2008;**26**(13):2155–61.
42. Comans EF, Hoekstra OS, van Lingen A, Lammertsma AA, Teule GJ. Diagnostic value and therapeutic efficacy of PET FDG in clinical oncology as measured by systematic clinical value assessment. *J Nucl Med* 1999;**40**(5):407.
43. CMS. Medicare national coverage determinations manual; Chapter 1, Part 4 (Sections 200–310.1). Coverage Determinations; 2009.
44. Facey K, Bradbury I, Laking G, Payne E. Overview of the clinical effectiveness of positron emission tomography imaging in selected cancers. *Health Technol Assess* 2007;**11**(44).
45. Fineberg HV. Evaluation of computed tomography: achievement and challenge. *AJR Am J Roentgenol* 1978;**131**(1):1–4.
46. Steinberg EP. Magnetic resonance coronary angiography – assessing an emerging technology. *N Engl J Med* 1993;**328**(12):879–80.
47. Nair VS, Barnett PG, Ananth L, Gould MK. Positron emission tomography (PET) 18F-fluorodeoxyglucose (FDG) uptake and prognosis in patients with resected, clinical stage Ia non-small cell lung cancer. *Chest* 2009;**137**(5):1150–6.
48. Juweid ME, Cheson BD. Positron-emission tomography and assessment of cancer therapy. *N Engl J Med* 2006;**354**(5):496–507.
49. Johnson KR, Ringland C, Stokes BJ, et al. Response rate or time to progression as predictors of survival in trials of metastatic colorectal cancer or non-small-cell lung cancer: a meta-analysis. *Lancet Oncol* 2006;**7**(9):741–6.
50. Weber WA, Petersen V, Schmidt B, et al. Positron emission tomography in non-small-cell lung cancer: prediction of response to chemotherapy by quantitative assessment of glucose use. *J Clin Oncol* 2003;**21**(14):2651–7.
51. Hoekstra CJ, Stroobants SG, Smit EF, et al. Prognostic relevance of response evaluation using [18F]-2-fluoro-2-deoxy-D-glucose positron emission tomography in patients with locally advanced non-small-cell lung cancer. *J Clin Oncol* 2005;**23**(33):8362–70.
52. Young H, Baum R, Cremerius U, et al. Measurement of clinical and subclinical tumour response using [18F]-fluorodeoxyglucose and positron emission tomography: review and 1999 EORTC recommendations. European Organization for Research and Treatment of Cancer (EORTC) PET Study Group. *Eur J Cancer* 1999;**35**(13):1773–82.
53. EORTC, GELA. Phase III randomized study of early Fludeoxyglucose F18 positron emission tomography scan-guided treatment adaptation vs. standard combined modality treatment in patients with previously untreated supradiaphragmatic stage I or II Hodgkin's Lymphoma. http://www.cancer.gov/clinicaltrials/EORTC-20051; 2010 March 3.
54. Schoder H, Fury M, Lee N, Kraus D. PET monitoring of therapy response in head and neck squamous cell carcinoma. *J Nucl Med* 2009;**50**(Suppl. 1):74S–88.
55. Eisenhauer EA, Therasse P, Bogaerts J, et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). *Eur J Cancer* 2009;**45**(2):228–47.
56. Juweid ME, Stroobants S, Hoekstra OS, et al. Use of positron emission tomography for response assessment of lymphoma: consensus of the Imaging Subcommittee of International Harmonization Project in Lymphoma. *J Clin Oncol* 2007;**25**(5):571–8.
57. Cheson BD, Pfistner B, Juweid ME, et al. Revised response criteria for malignant lymphoma. *J Clin Oncol* 2007;**25**(5):579–86.
58. Stroobants S, Goeminne J, Seegers M, et al. 18FDG-Positron emission tomography for the early prediction of response in advanced soft tissue sarcoma treated with imatinib mesylate (Glivec). *Eur J Cancer* 2003;**39**(14):2012–20.
59. Boellaard R, O'Doherty MJ, Weber WA, et al. FDG PET and PET/CT: EANM procedure guidelines for tumour PET imaging: version 1.0. *Eur J Nucl Med Mol Imaging* 2010;**37**(1):181–200.
60. Shankar LK, Hoffman JM, Bacharach S, et al. Consensus recommendations for the use of 18F-FDG PET as an indicator of therapeutic response in patients in National Cancer Institute Trials. *J Nucl Med* 2006;**47**(6):1059–66.
61. Cerfolio RJ, Bryant AS, Winokur TS, Ohja B, Bartolucci AA. Repeat FDG-PET after neoadjuvant therapy is a predictor of pathologic response in patients with non-small cell lung cancer. *Ann Thorac Surg* 2004;**78**(6):1903–9.
62. Prenen H, Deroose C, Vermaelen P, et al. Establishment of a mouse gastrointestinal stromal tumour model and evaluation of response to imatinib by small animal positron emission tomography. *Anticancer Res* 2006;**26**(2A):1247–52.
63. Su H, Seimbille Y, Ferl GZ, et al. Evaluation of [(18)F]gefitinib as a molecular imaging probe for the assessment of the epidermal growth factor receptor status in malignant tumors. *Eur J Nucl Med Mol Imaging* 2008;**35**(6):1089–99.
64. Moskowitz CH, Schoder H, Teruya-Feldstein J, et al. Risk-Adapted Dose-dense immunochemotherapy determined by interim FDG-PET in advanced-stage diffuse large B-cell lymphoma. *J Clin Oncol* 2010;**28**(11):1896–903.

65. Duhrsen U, Huttmann A, Jockel KH, Muller S. Positron emission tomography guided therapy of aggressive non-Hodgkin lymphomas–the PETAL trial. *Leuk Lymphoma* 2009;**50**(11):1757–60.
66. Hoekstra CJ, Paglianiti I, Hoekstra OS, et al. Monitoring response to therapy in cancer using [18F]-2-fluoro-2-deoxy-D-glucose and positron emission tomography: an overview of different analytical methods. *Eur J Nucl Med* 2000;**27**(6):731–43.
67. Eccles M, Steen N, Grimshaw J, et al. Effect of audit and feedback, and reminder messages on primary-care radiology referrals: a randomised trial. *Lancet* 2001;**357**(9266):1406–9.
68. Rosenbaum S, Frankford DM, Moore B, Borzi P. Who should determine when health care is medically necessary? *N Engl J Med* 1999;**340**(3):229–32.
69. van Dongen GA, Visser GW, Lub-de Hooge MN, de Vries EG, Perk LR. Immuno-PET: a navigator in monoclonal antibody development and applications. *Oncologist* 2007;**12**(12):1379–89.
70. Bree Rd, Deurloo EE, Snow GB, Leemans CR. Screening for distant metastases in patients with head and neck cancer. *Laryngoscope* 2000;**110**(3 Pt. 1):397–401.