EUROPEAN JOURNAL OF CANCER 44 (2008) 39–45

available at www.sciencedirect.com

ScienceDirect

journal homepage: www.ejconline.com

Review

How to assess anti-tumour efficacy by imaging techniques

Stephen J. Gwyther[a,*], *Lawrence H. Schwartz*[b]

[a]*Department of Medical Imaging, East Surrey Hospital, Canada Avenue, Redhill, Surrey RH1 5RH, UK*
[b]*Memorial Sloan-Kettering Cancer Center, New York, USA*

ARTICLE INFO

Article history:
Received 8 October 2007
Accepted 10 October 2007

Keywords:
Response evaluation
RECIST
Cytocidal agents
Cytostatic agents
PET–CT
Dynamic contrast enhanced MRI and CT
Volumetric studies

ABSTRACT

Response evaluation in the assessment of potential new anti-cancer therapies is undergoing intense investigation and change. Current imaging techniques most commonly used in early phase clinical trials are limited to providing reliable and reproducible anatomical data demonstrating a change in size and reduction in tumour volume thereby inferring patient benefit. Current imaging modalities such as computed tomography (CT) and magnetic resonance imaging (MRI) by their nature require computer programs and software. This is a constantly evolving field and upgraded technology enables faster acquisition times for scans, greater anatomical detail and accurate volumetric data to be acquired. Dynamic studies allow contrast agents to be visualised in any given structure over time, so blood flow, blood volume and permeability can be assessed thereby demonstrating function. The advent of many new anti-cancer agents with novel modes of action such as anti-angiogenesis agents act by preventing the development of a suitable blood supply to sustain tumour growth. Such agents do not actively destroy tumour cells so do not exhibit a 'cytocidal' effect as traditional anti-cancer agents do but prevent tumour growth, so can be regarded as 'cytostatic' agents. Therefore, traditional response evaluation criteria may not be appropriate to assess drug efficacy or 'activity' in achieving patient benefit. New techniques have also been developed so the 'function' or metabolism can be demonstrated and tumour serum markers and other factors also require consideration rather than relying on a single modality alone.

This article reviews the current accepted response criteria and highlights some newer techniques which will almost certainly play a major role in the assessment of new anti-cancer therapy, particularly in the development of cytostatic agents which are playing an ever increasing role.

1. Introduction

The use of imaging for the assessment of response to cytotoxic and cytostatic therapies is undergoing intense investigation and change. This article will review the current role of imaging and response end-points for cytotoxic therapies, possible imaging novel techniques and paradigms for the assessment of cytostatic agents.

There is a need to develop new anti-cancer therapies, and a large potential pipeline is currently under development. A major goal is to develop these therapies in the shortest time span while ensuring they are safe, reliable and effective.

* *Corresponding author:* Tel.: +44 01737 782 852; fax: +44 01737 780 395.
E-mail address: GwytherSJ@aol.com (S.J. Gwyther).
0959-8049/$ - see front matter
doi:10.1016/j.ejca.2007.10.010

Ultimately, these drugs should prolong patient survival and improve patient quality of life. However, it is rarely feasible to use survival as an end-point from the outset because of the large numbers of patients required for a meaningful statistical analysis and the potential for confounding effects of other therapies and ethical issues related to exposure of patients to non-optimal therapies. Biomarkers for drug activity are therefore used to predict potentially active agents in early phases of drug discovery and surrogate markers may be used in later phase clinical trials to avoid some of the pitfalls on other therapeutic end-points.

Various imaging modalities, techniques and post-processing algorithms may be used as biomarkers and surrogates for response assessment in oncological drug discovery. The use of these imaging tools has undergone dramatic change in the past decade with the development of novel therapeutic agents.

Until 10 years ago, the development of chemotherapeutic agents was predominantly cytotoxic in nature producing irreversible lethal effects within all cells and particularly within rapidly dividing malignant cells. The clinical trials created to analyse these agents were generally divided into phases, where phase 1 studies evaluated primarily dose escalation to determine the maximum tolerated dose (MTD) of the drug and to determine the appropriate dose for further studies. Toxicity was the primary end-point and imaging played little if any role in phase 1 studies.

The primary goal of early phase 2 studies was to demonstrate activity and efficacy of a single agent in a specific predefined tumour type. If the agent showed evidence of response, primarily defined by imaging and an acceptable toxicity profile then larger phase 2 studies may investigate the efficacy of single agent or multimodality combination treatment and randomised studies may be undertaken to compare the agent or combination with standard therapy. Imaging, in particular, played a pivotal role in these studies because of the primary end-point of 'response'. The response rate as determined in most phase 2 studies was based on imaging. A decrease in tumour sizes during treatment by a predefined percentage was used to define patients 'responding' to the therapy and forms the basis of response rates, which produced a surrogate for efficacy and patient benefit. Finally, phase 3 clinical trials were used as confirmatory trials to establish the patient risk–benefit profile of the agent under investigation and constitute large, multi-centre, randomised studies.[1] Imaging played an important but sometimes, secondary role in these trials.

These paradigms have changed dramatically with the development of cytostatic anti-cancer therapies which, for example, may inhibit cell division without causing cell death. Examples include anti-angiogenesis agents and immune modulators. The classic phase 1 and 2 clinical trial design using toxicity and image based response rates may not be entirely relevant for the development of these cytostatic therapies. Toxicity as defined in phase 1 studies may not help to determine the optimal dose of the therapy, and response rate as defined by a fixed percentage of tumour size reduction may not be the appropriate end-point for surrogate definition. This is not to limit the role of imaging, in fact, quite the contrary. Imaging is likely to play an ever increasing role in drug discovery, however, the role will change and the imaging techniques utilised will need to be redefined with these changing therapies.

2. Imaging in the development of cytocidal agents

Response rates as determined with imaging are the primary end-point of phase 2 studies and determines whether suitable activity appears to be present (with acceptable toxicity, which is generally not assessed by imaging).

The World Health Organisation (WHO) devised the first standardised response criteria which were developed and published in 1981 by Millar.[2] Each well-defined lesion is measured in its largest diameter and multiplied by its greatest perpendicular diameter to give a bi-dimensional product. Each lesion measured is summed before treatment commences and the same lesions are measured during treatment, typically after every two courses of therapy. The degree of response is assigned depending on the predetermined percentage reduction in tumour product compared to the pre-treatment product. Four categories are recognised, complete response (CR), partial response (PR), stable disease (SD) and progressive disease (PD), see Table 1. At least one well-defined, measurable lesion is necessary for measurement at the pre-treatment imaging study. If all disease disappears, the patient is assigned a category of CR. During treatment, if the measurable disease decreases by the pre-determined percentage or resolves but there is evidence of non-measurable disease, then PR is assigned. Any evidence of new disease elsewhere during the study is assigned PD, even if the measurable lesions are responding. When the criteria were devised in 1981 almost all lesions were assessed by the standard chest radiograph (CXR) or by non-imaging physical examination. This was a simple process because with the radiograph on a light box lesions could be measured using a ruler.

Great advances in imaging technology have seen the widespread introduction of new radiological modalities including ultrasound (US), computerised tomography (CT) and mag-

Table 1 – Categories of response by WHO criteria

Complete response (CR): complete disappearance of all disease, measurable or otherwise during treatment
Partial response (PR): decrease in sum of tumour products by 50–99% compared with the pre-treatment examination and/or evidence of residual non-measurable disease
Stable disease (SD): reduction in disease product by less than 50% compared to the pre-treatment examination and an increase in product of any given lesion by less than 25% compared to its smallest diameter during treatment
Progressive disease (PD): increase in product of any lesion by 25% or more compared to its smallest product during treatment or the appearance of incontrovertible new disease

netic resonance imaging (MRI). During the 1980s, these modalities became more widely available and were incorporated into clinical trials as response assessment tools. While this was beneficial for lesion detection and measurement, there was confusion regarding the technical parameters used with these modalities. There were no guidelines regarding number or minimum size at the baseline examination and no imaging stipulations such as the slice thickness in CT with regard to the minimum lesion size, the use of oral and intravenous contrast agents was not standardised, even during individual clinical trials. Finally, the claimed responders were not initially verified by an independent third party and in some early studies where independent verification was undertaken, up to 30% of claimed responders were rejected [3,4] Although this reduction seems high, this was the maximum and the revised response rates were more robust and were found to be more uniform from one trial to another.

To try and add more certainty to the studies, different co-operative groups developed their own response criteria and modified existing criteria. In fact, this confused the issue further because direct comparisons between one study and another became even more difficult when different response criteria were used.

In the late 1990s, an international working group was formed to look into this matter and from this the response evaluation criteria in solid tumours (RECIST) arose in 2000.[5] The RECIST criteria sought to unify the various criteria by introducing imaging criteria which it was hoped would clarify and ease the situation. Most examinations employ CT, though MRI and CXR are acceptable. US should be avoided with the exception of imaging very superficial lesions as it is largely a subjective examination. The key imaging points were to define a minimum-sized lesion at the baseline examination, particularly with regard to slice thickness and the minimum lesion size at baseline was set at 1.0 cm, suggest the use of intravenous contrast agents unless medically contra-indicated, ensure all relevant soft tissue, lung and other 'windows' were imaged, measure all lesions up to a maximum of 10 (and a maximum of 5 per organ) and most controversially introduce uni-dimensional measurements, the longest diameter only per lesion and sum the lengths. CR remains unaltered but a PR is achieved if there is a reduction in the sum of the uni-dimensional measurements of 30% compared to the baseline figure. This is the equivalent of a bi-dimensional reduction of 50%, so response has the same meaning, and there is some continuity from previous studies. PD is defined as a 20% or more increase in the sum of the lesions compared to the examination showing the greatest reduction from baseline. This represents the equivalent of a 44% rise in bi-dimensionally measurable disease, so differs from the WHO criteria, which only requires a 25% increase. The reason is that a 25% increase in bi-dimensionally measurable disease is only a 12% increase in uni-dimensionally measurable lesions and with small lesions a small error in measurement leads to a large percentage error. It is noteworthy that PD is based on the sum of the lesions attaining a 20% or more increase, not any given lesion attaining this figure. The appearance of any new disease does, of course, constitute PD, Table 2. The relationship between tumour volume, product and length is shown in Table 3. The use of uni-dimensional measurements has been validated by the RECIST group and others.[5–7]

Table 3 – Relationship between change in tumour diameter (length), product and volume

Diameter (%)	Product (%)	Volume (%)
30	50	65 (Decrease)
12	25	40 (Increase)
20	44	73 (Increase)

The RECIST criteria were formulated over several years including consultation with various co-operative groups, opinion leaders, regulatory authorities and pharmaceutical companies, all of whom have valid though different perspectives on the overall equation, though all with a similar overall objective. During this time, there were inevitable changes in technology most notably in CT scanning, due to the advent of multi-slice CT (MSCT). This single change in technology has allowed very fast scanning through body parts, allowing IV contrast enhanced scans to be performed in various phases of enhancement, namely arterial, equilibrium, portal venous and delayed phases. The subtleties emanating from this technology were not foreseen by RECIST and have not, to date, been addressed by it. At the time of publication in 2000, it was felt that clinical trials should reflect the imaging technology throughout the world where trials are undertaken, but virtually all CT scans throughout the world are now performed using MSCT.

In retrospect, there are several flaws in the RECIST criteria, bearing in mind the technological advances since its introduction. The timing of the scan with regard to the administration of the IV contrast agent is important. Some visceral lesions can appear a different size depending on the timing of the scan. Some tumour types, such as metastatic carcinoid to the liver show maximal uptake in the arterial phase, whereas most solid tumours show maximal uptake in the portal venous phase. The timing of the scan must, therefore,

Table 2 – Categories of response by RECIST Criteria

CR: Complete disappearance of all disease, measurable or otherwise during treatment
PR: Decrease in sum of tumour lengths by 30–99% compared with the baseline examination and/or evidence of residual non-measurable disease
SD: Reduction in sum of tumour length by less than 30% and an increase in the sum of the tumour lengths by less than 20% during treatment
PD: Increase in the sum of the tumour lengths by 20% or more compared to the smallest total length during treatment or the appearance of incontrovertible new disease

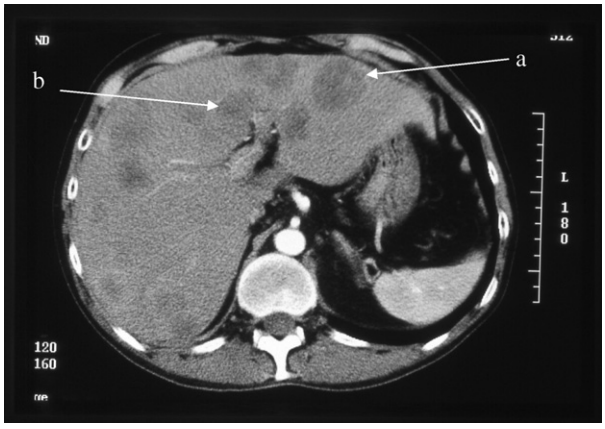

Fig. 1 – IV contrast-enhanced scan through the liver during the arterial phase, about 30 s after commencing the injection. Multiple metastases from colonic carcinoma are seen. Note the lesions labelled a and b.

be approximately similar, within 10 s, from one examination to another (Figs. 1 and 2). RECIST assumes all lesions are spherical and will decrease in size uniformly with treatment. This is not the case, so lesions, where there is an elliptical decrease in tumour product and obvious significant reduction in tumour bulk, may well be assigned SD during treatment as the greatest diameter is measured. This leads to an increase in the number of patients with SD, rather than PR. Does this matter? It probably doesn't matter because those patients who have progressed on conventional treatment, but when on the experimental regime, showed some stabilisation of disease and are experiencing some benefit from treatment, assuming that this is sustained and that there is an acceptable quality of life for the patient. This is even more pertinent with the introduction of cytostatic agents (see Figs. 3 and 4).

There is evidence to suggest that with therapy the change in the short axis measurement of lymph nodes is of greater prognostic significance than the measurement of the greatest diameter.[8–10] One study using oesophageal cancer as a model

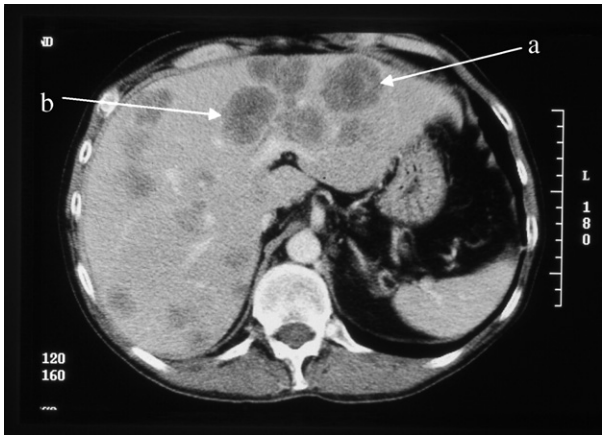

Fig. 2 – IV contrast enhanced scan through the corresponding level as Fig. 1. during the portal venous phase, about 80 s after commencing the injection. Note lesion a is similar in dimensions to Fig. 1, but lesion b appears much larger.

compared the response rate using the WHO and RECIST criteria for measuring the lesions. There was a discordance rate of 26% between the criteria.[10] This was most pronounced when lesions regressed in an elliptical manner, though for lymph nodes the discordance was less when the short axis was measured rather than the greatest diameter. This leads to potential problems in determining whether lesions represent solid tumour metastases or nodes, which in the case of mesenteric masses could be difficult. If solid tumours were to be measured uni-dimensionally in their long axis and nodes in their short axis great confusion will ensue and the system will be unworkable. If PR is to remain a major end-point, then perhaps bi-dimensional measurements of the lesions would circumvent this problem. Why not measure the tumour volume? The recent advent of reliable software combined with fast multi-slice CT and MRI scanners allows the accurate reconstruction of tumours in three dimensions and small changes in volume during therapy can be detected which may predict response outcomes more reliably than traditional uni or bi-dimensional measurements. This is the subject of considerable interest using both CT and MRI.

Some tumour types are not easily measured because of their mode of growth. Malignant pleural mesothelioma grows as a 'rind' of tissue arising from the pleura rather than by discrete nodules, so measuring the long axis is not appropriate. Instead modified criteria are employed, where the tumour thickness is measured perpendicular to the chest wall or mediastinum in two positions at three different levels on the thoracic CT scan. The uni-dimensional measurements can be summed in the usual way for RECIST and the response is defined as >30% reduction in tumour measurements compared to the baseline examination. This has been validated and correlates with survival and changes in lung function tests.[11] With the development of cytostatic agents, where lesion regression may not occur to the same extent or magnitude, then perhaps SD should be included in the 'preferred' categories of response. The definition of a PR is somewhat arbitrary and does not have a biological correlate. Potentially with cytostatic agents, where lesion regression may not occur to the same magnitude, SD may be included in determining 'response'.

Duration of response is important. A total (or significant) reduction in tumour bulk is meaningless if the duration is short lived and the disease recurs. For a response to be confirmed, the response must be demonstrated on two successive examinations not less than 28 days apart. If this does not occur then SD is assigned, unless PD ensues. However, the range and mean periods of response should be recorded for all responders.

The number of lesions measured is a matter of consternation. Both inter-observer and intra-observer errors occur for a number of reasons such as poorly defined lesions, irregular lesions and when imaging occurs at different phases of intravenous contrast enhancement. It is not feasible to measure all lesions when the tumour burden is high. RECIST requires that all measurable lesions, up to a maximum of 10 should be measured and followed, no more than 5 per organ. This is a time-consuming process, not so much measuring the lesions, but determining which lesions are which on subsequent examinations. Schwartz measured a number of

lesions at the baseline examination and at the first post-treatment examination in one and two dimensions then ran a computerised modelling analysis to determine the response assignation for every possible number of lesions. The variance in response assignment was calculated and 6 lesions were required to be measured to reduce the variance by 90% when bi-dimensional measurements were made and 4 lesions when uni-dimensional measurements were made.[12] A further dilemma is which lesions to measure. The largest may not necessarily be the best because they may be partially necrotic and thus may not necessarily decrease in size to the same extent as smaller lesions containing viable tissue. Similarly, when tumours respond and decrease in size to less than 5 mm a small error in measurement will lead to a large percentage error and may falsely suggest PD. One possible solution is for any visible lesion which has decreased in size to less than 5 mm but persists should be assigned a nominal measurement of 5 mm.

3. Imaging in the development of cytostatic agents

Most cytostatic agents or class of agents have a potentially different mode of action, so it is difficult to produce response assessment criteria without extensive validation. Since these agents inhibit new cell growth without necessarily causing cell death, the traditional imaging criteria may not necessarily be valid. Some agents do show cytocidal activity, but others do not, so a uniform means of assessment of therapies may be difficult. This has led to different agents being assessed with varying imaging modalities and techniques. The emergence of functional imaging enables physiological processes within cells to be monitored rather than simply assessing the physical size changes. This takes several forms, for instance, dynamic contrast enhanced (DCE) CT and MRI studies reflect changes in blood flow, blood volume, perfusion and permeability in angiogenesis. An IV contrast agent is given and the change in concentration with time within a tumour can be visualised, either graphically or colour coded. This is well demonstrated in breast cancer where the type 3 curve is indicative of cancer.[13] Changes in blood flow, blood volume and perfusion with therapy can, therefore, be recorded using a non-invasive technique and act as a surrogate for anti-angiogenesis effects, with tumour vessel stabilisation and thus patient benefit. A recent study assessing a tyrosine kinase inhibitor targeting vascular endothelial growth factor (VEGF) receptors in metastatic renal cell cancer used DCE-MRI measuring both the volume transfer constant of the contrast agent (k-trans) and the area under the curve at 90 s (IAUC90) showed correlation between progression free survival and patients with a high baseline k-trans versus a low baseline k-trans, though little correlation with the IAUC90.[14] DCE-MRI and DCE-CT will inevitably play a significant role in response evaluation particularly for anti-angiogenesis agents and cytostatic agents once standardised protocols are developed and instituted. DCE-CT may be more successful in the assessment of thoracic lesions where motion artefact may degrade DCE-MRI examinations to a greater extent than the liver and central nervous system.

One of the most exciting changes to have occurred is positron emission tomography (PET), where a positron emitting isotope can be incorporated into normal physiological substances. The most widely used agent is 18F fluorodeoxy-D-glucose (FDG), a glucose analogue which is incorporated into actively metabolising cells, but unlike glucose is not metabolised and so becomes concentrated and can be visualised. Tumour activity can thus be recorded in normal-sized lymph nodes and unlike anatomical imaging is independent of the enlargement of the nodes. The spatial resolution is relatively poor but this has been overcome by combining a PET and a CT scanner together enabling accurate anatomical localisation of suspicious lesions. In clinical practice, treatment effects can be monitored using both cytostatic and cytocidal agents by recording the decrease in glucose metabolism with treatment. Many tumours do show activity such as non-small cell lung cancer, colonic cancer, melanoma and high grade lymphomas, but others such as mucinous gastrointestinal tumours, some low grade lymphomas and slowly growing tumours are not

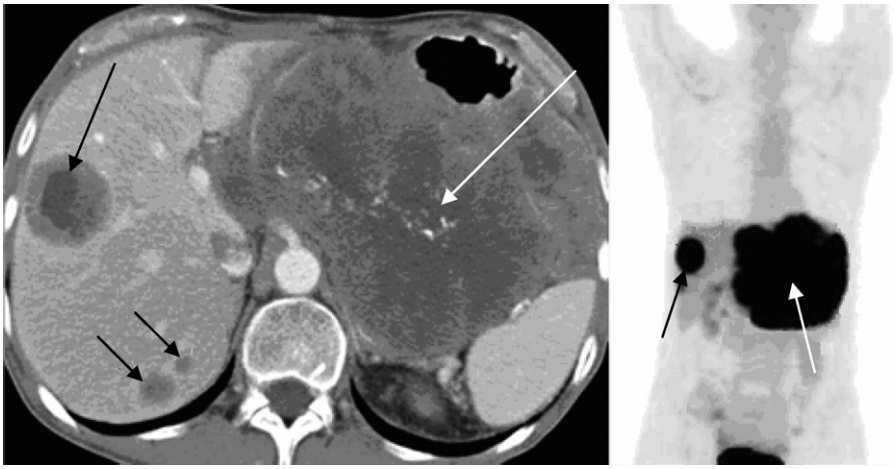

Fig. 3 – Left image, CT scan through the upper abdomen in a patient with a large newly diagnosed gastro intestinal stromal tumour (GIST) lying between the stomach and liver (white arrow) and liver metastases (black arrows). Right image, the FDG PET images show avid FDG uptake in the primary tumour (white arrow) and in one of the hepatic metastases (black arrow).

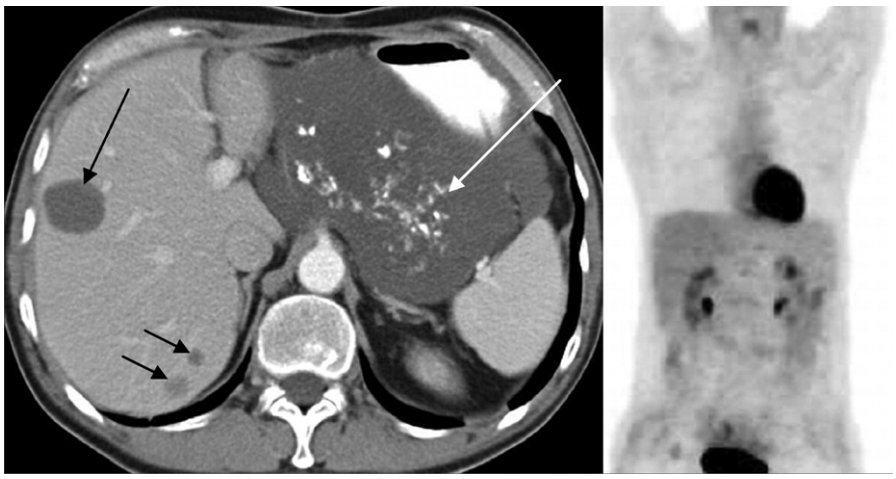

Fig. 4 – Same patient as Fig. 3. Left image, CT scan through the upper abdomen 3 months later after treatment with imatinib. The primary mass is slightly smaller (white arrow) and the liver metastases are smaller (black arrows), but using anatomical criteria, this does not attain a PR. Right image, the FDG PET scan shows no evidence of FDG activity in either the primary tumour or the hepatic metastases indicating a CR.

FDG avid, in which case FDG-PET is of no use. FDG-PET has been used in the development and registration of imatinib, a tyrosine kinase inhibitor, with particular effect on the c-KIT gene which is over expressed in gastrointestinal stromal tumours (GISTs). These tumours tend to be large and when treated with imatinib become necrotic, so although the lesion size remains large and by anatomical criteria does not fulfill a PR, the PET–CT demonstrates little or no FDG uptake indicating a PR, Figs. 3 and 4. A corresponding clinical improvement has been demonstrated. At present, set protocols have yet to be developed and quantitative measurements are dependent on many factors, not least the dose injected and the length of time after the injection to the acquisition of the images. Combining volumetric studies with functional imaging may be of particular benefit for patients treated with cytostatic agents who would traditionally be defined as having stable disease and may yield further information regarding tumour function rather than tumour size.

Response assessment in haematological malignancies employs slightly different end-points to those in solid tumours. The main difference is that treatment of haematological malignancies aims to eradicate all disease, failure to do so inevitably results in recurrence, so imaging and other techniques are designed to determine if a CR has been achieved. Anything less requires further treatment. Cheson described criteria for the assessment of lymphomas in 1999, primarily using CT, but also included biochemistry and clinical assessment to help build up a more accurate assessment of events.[15] Juweid added FDG-PET to the existing criteria in a group of patients with aggressive non-Hodgkin's lymphoma and found a subset of patients who had PR by conventional imaging but CR by FDG-PET and this group had a better overall prognosis so suggested modifying the criteria accordingly.[16] This is not surprising, because he selected the patients with no evidence of increased metabolism and thus no evidence of active disease, rather than assuming an enlarged mass contained tumour. An international imaging committee made some recommendations regarding PET.[17] These included ensuring imaging after chemotherapy or chemo-immunotherapy that occurred at least 3 weeks after the completion of therapy and preferably 6–8 weeks and 8–12 weeks after the completion of radiotherapy or chemo-radiotherapy. Visual assessment alone is sufficient to determine response at the end of therapy, rather than measure standardised uptake values (SUV). For residual masses greater than 2 cm diameter, the mediastinal blood pool should be used as a reference, activity greater than this in the mass indicates active disease and for smaller lesions, activity greater than the surrounding background is regarded as positive. Specific criteria are suggested in the assessment of the liver, spleen, lungs and bone marrow, attenuation-corrected PET is recommended and PET treatment monitoring should only be undertaken in the context of a clinical trial or prospective registry. The 1999 criteria were subsequently revised by Cheson in 2007 and other additions were also made such as the role of immune-histochemistry and flow cytometry.[18]

It is interesting that the revised lymphoma imaging criteria take into account the various information available, including biochemistry and physical examination, not just the PET and CT data, though this has a high weighting. This allows an overall 'picture' to be built up. It is likely that new agents under development for the treatment of solid tumours will require similar investigation. This is of benefit in some tumour types such as epithelial ovarian cancer, where extensive disease often takes the form of widespread peritoneal studding with small tumours which are not easily measurable with conventional anatomical imaging and ascites. Adding functional imaging such as PET–CT and recording the serum CA-125 levels during treatment will give a more accurate overall assessment of response to treatment than by relying on any one given modality. It may be that some modalities will have greater weighting than others and at the moment this is likely to be PET–CT because the change in tumour metabolism can be appreciated in a much shorter time frame, such as after one course of therapy, rather than waiting to see a macroscopic change in tumour size which inevitably occurs much later. A

change in tumour function would also be expected when cytostatic agents are used as they prevent new cell growth, so although lesions may not regress, any change in metabolism from pre-treatment to during therapy would suggest patient benefit and these changes are typically seen early during treatment, even after one course of treatment. Conversely, ineffective therapy could also be detected early and withdrawn. Not all solid tumours are FDG avid, so only those which are avid should be investigated using FDG-PET. PET–CT should also only be undertaken in well-designed clinical trials or in a prospective registry where the data is recorded and captured for meaningful scrutiny at a later stage. Paradoxically, despite the advent of more sophisticated imaging modalities the data they yield needs to be taken in context with other data such as tumour markers, immune-histochemistry and clinical examination. The aim is to help develop active new anti-cancer therapies which will prolong survival, provide a good quality of life and which are safe and effective. The most likely means of achieving this is to use all the data at our disposal rather than rely on any particular modality, though different weightings may be given to different modalities.

Conflict of interest statement

None declared.

REFERENCES

1. CPMP/EWP/205/95/Rev.3/Corr.2 Guideline on the evaluation of anticancer medicinal products in man. June 2006. <http://www.emea.eu.int/pdfs/human/ewp/020595en.pdf> [accessed 26.02.2007].
2. Millar AB, Hoogstraten B, Staquet M, Winkler A. Reporting results of cancer treatment. *Cancer (Phila)* 1981;**47**:207–14.
3. Gwyther SJ, Aapro MS, Hatty SR, Postmus PE, Smith IE. Independent review of gemcitabine in patients with non-small cell lung cancer (NSCLC). *Eur Respirat J* 1997;**10**(25):196s–7s.
4. Gwyther SJ, Bolis G, ten Bokkel Huinink W, et al. Experience with independent radiological review during a topotecan trial in ovarian cancer. *Annal Oncol* 1997;**8**:463–7.
5. Therasse P, Arbuck SG, Eisenhauer EA, et al. New guidelines to evaluate the response to treatment in solid tumors. *J Natl Cancer Inst (Bethesda)* 2000;**92**:205–16.
6. James K, Eisenhauer EA, Christian MC, et al. Measuring response in solid tumors: uni-dimensional versus bi-dimensional measurement. *J Natl Cancer Inst (Bethesda)* 1999;**91**:523–8.
7. Dachman AH, MacEneaney PM, Adedipe A, Carlin M, Schumm LP. Tumor size on computed tomography scans: is one measurement enough? *Cancer (Phila)* 2001;**91**:555–60.
8. Kamiyoshihara M, Kawashima O, Ishikawa S, Morishita Y. Mediastinal lymph node evaluation by computed tomographic scan in lung cancer. *J Cardiovasc Surg (Torino)* 2001;**42**:119–24.
9. Konishi J, Yamazaki K, Tsukamoto E, et al. Mediastinal lymph node staging by FDG-PET in patients with non-small cell lung cancer: analysis of false-positive FDG-PET findings. *Respiration* 2003;**70**:500–6.
10. Schwartz LH, Colville JAC, Ginsberg MS, et al. Measuring tumor response and shape change on CT: esophageal cancer as a paradigm. *Annal Oncol* 2006;**17**:1018–23.
11. Byrne MJ, Nowak AK. Modified RECIST criteria for assessment of response with malignant pleural mesothelioma. *Annal Oncol* 2004;**15**:257–60.
12. Schwartz LH, Mazumdar M, Brown W, Smith A, Panicek DM. Variability in response assessment in solid tumors: Effect of number of lesions chosen for measurement. *Clin Cancer Res* 2003;**9**:4318–23.
13. Kuhl CK, Mielcareck P, Klaschik S, et al. Dynamic breast MR imaging: are signal intensity time course data useful for differential diagnosis of enhancing lesions. *Radiology* 1999;**211**:101–10.
14. Hahn OM, Yeng C, Medved M, et al. Dynamic contrast enhanced MRI (DCE-MRI) pharmacodynamic (PD) study of sorafenib in metastatic renal cell carcinoma (RCC): Results of a randomized phase II trial. In: JCO 2007; ASCO annual meeting proceedings (post-meeting edition); 25: No.18S (June 20 Suppl.); 2007. p. 3545.
15. Cheson BD, Horning SJ, Coiffier B, et al. Report of an international workshop to standardize response criteria for Non-Hodgkin's lymphomas. *J Clin Oncol* 1999;**17**:1244–53.
16. Juweid ME, Wiseman GA, Vose JM, et al. Response assessment of aggressive Non-Hodgkin's lymphoma by integrated international workshop criteria and fluorine-18-fluorodeoxyglucose positron emission tomography. *J Clin Oncol* 2005;**23**:4652–61.
17. Juweid ME, Stroobants S, Hoekstra OS, et al. Use of positron emission tomography for response assessment in lymphoma: consensus of the imaging subcommittee of international harmonization project in lymphoma. *J Clin Oncol* 2007;**25**:571–8.
18. Cheson BD, Pfiestner B, Juweid ME, et al. Revised response criteria for malignant lymphoma. *J Clin Oncol* 2007;**25**: 579–86.