What you see is (not) what you get: tools for a non-radiologist to evaluate image quality in lung cancer

Evelyn E.C. de Jong¹,¹, Lizza E.L. Hendriks⁵,¹, Wouter van Elmpt⁶, Hester A. Gietema⁷, Paul A.M. Hofman⁸, Dirk K.M. De Ruysscher⁹, Anne-Marie C. Dingemans⁵,⁎

¹ The D-Lab: Decision Support for Precision Medicine, GROW-School for Oncology and Developmental Biology, Maastricht University Medical Center+, Universiteitssingel 40, 6229 ER, Maastricht, The Netherlands
² Department of Pulmonary Diseases, GROW-School for Oncology and Developmental Biology, Maastricht University Medical Center+, P. Debyelaan 25, 6229 HX, Maastricht, The Netherlands
³ Department of Radiation Oncology (MAASTRO), GROW-School for Oncology and Developmental Biology, Maastricht University Medical Center+, Doctor Tanslaan 12, 6229 ET, Maastricht, The Netherlands
⁴ Department of Radiology, Maastricht University Medical Center+, P. Debyelaan 25, 6229 HX, Maastricht, The Netherlands

ARTICLE INFO

Keywords:
Non-small cell lung cancer
Contrast-enhanced computed tomography
Gadolinium-MRI
¹⁸F-FDG-PET
Image guidelines

ABSTRACT

Medical images are an integral part of oncological patient records and they are reviewed by many different specialists. Therefore, it is important that besides imaging experts, other clinicians are also aware that the diagnostic value of a scan is influenced by the applied imaging protocol.

Based on two clinical lung cancer trials, we experienced that, even within a study protocol, there is a large variability in imaging parameters, which has direct impact on the interpretation of the image. These two trials were: 1) the NTR3628 in which the added value of gadolinium magnetic resonance imaging (Gd-MRI) to dedicated contrast enhanced computed tomography (CE-CT) for detecting asymptomatic brain metastases in stage III non-small cell lung cancer (NSCLC) was investigated and 2) a sub-study of the NVALT 12 trial (NCT01171170) in which repeated ¹⁸F-fluorodeoxyglucose positron emission tomography (¹⁸F-FDG-PET) imaging for early response assessment was investigated.

Based on the problems encountered in the two trials, we provide recommendations for non-radiology clinicians, which can be used in daily interpretation of imaging. Variations in image parameters cannot only influence trial results, but sub-optimal imaging can also influence treatment decisions in daily lung cancer care, when a physician is not aware of the scanning details.

1. Introduction

Medical imaging is an essential component of the diagnostic procedures performed in lung cancer. Next to that, it is also used for response assessment. The imaging modalities used in oncology have evolved from simple X-rays to computed tomography (CT) and magnetic resonance imaging (MRI) scans. Nuclear imaging has innovated by the introduction of positron-emission tomography (PET) with several tracers being ¹⁸F-fluorodeoxyglucose (¹⁸F-FDG) most frequently used. An ¹⁸F-FDG-PET-scan using an integrated PET-CT scanner combined with a contrast enhanced CT (CE-CT) is nowadays a standard staging technique in thoracic oncology. Due to ongoing technological innovations, the sensitivity and specificity of these modalities have significantly improved. However, many factors, as patient preparation, image acquisition and reconstruction parameters affect the quality and accuracy of all these exams [1].

Images are nowadays an integral part of electronic patient records and can be reviewed directly by many different health care specialists. Consequently, it is important that besides the imaging experts (i.e. radiologists and nuclear medicine physicians), other reviewing clinicians are aware that the diagnostic value of a scan is influenced by the applied imaging protocol and can recognize common artifacts (e.g. breathing artifacts). More knowledge on this topic will provide clinicians tools to communicate with their imaging colleagues to prevent...
image misinterpretation and to select the appropriate image acquisition protocol.

In two recently performed clinical trials (NTR3628 and NCT01171170) studying patients with non-small cell lung cancer (NSCLC) we experienced that, even though specific imaging guidelines were mandated by the trial protocol, there was a large variability in imaging parameters. This influenced not only the outcome of the trial but could also have impact on treatment by their clinical physician (medical oncologist/pulmonologist). In this manuscript, we describe the imaging problems encountered in these two trials. Furthermore, we will provide recommendations to guide clinicians in the interpretation of medical imaging based on our experience. Our goal is that this will result in improved clinical care as well as imaging standardization, not only in future multicenter studies, but also in daily clinical care.

2. Methods

2.1. NVALT12 18F-FDG PET/CT imaging sub-study

In all lung cancer patients eligible for therapy with curative intent, not only a CE- chest-CT, but also a whole body 18F-FDG-PET is recommended [2,3]. The 18F-FDG-PET, performed with a non-diagnostic low dose CT (LD-CT) for attenuation correction can be extended by an additional diagnostic CE-CT of the chest (with or without the upper abdomen and brain).

In the multicenter randomized phase II NVALT12 trial (NCT01171170) chemo-naïve patients with stage IV non-squamous NSCLC were treated with paclitaxel-carboplatin-bevacizumab with or without nitroglycerin patches. For patients in whom an 18F-FDG-PET at baseline was performed as part of standard work-up, a second study 18F-FDG-PET was performed within three weeks after start of treatment. The two 18F-FDG-PETs were used for response assessment, and results have been reported [4]. It was assumed that image acquisition was per the European Association of Nuclear Medicine (EANM) guidelines for tumor imaging version 1 as introduced in 2009, therefore no PET guidelines were added to the protocol [1].

2.2. Study NTR3628: brain imaging

Although patients with brain metastases often have neurological complaints, 3–21% of neurologically asymptomatic patients with otherwise stage I–III lung cancer are diagnosed with brain metastases on MRI [5]. This diagnosis is especially important in patients that are potentially eligible for therapy with curative intent. A post-gadolinium-MRI (post-Gd-MRI) is the imaging modality of choice, but when MRI is contra-indicated or too difficult to arrange within a reasonable time scale, a diagnostic CE-CT is an acceptable alternative [2,3]. Except for the recommendation to include Gd-contrast series, no recommendations are given in the ESMO and NCCN lung cancer guidelines regarding the minimal requirements for this brain MRI (e.g. applied MRI sequences (e.g. T1, T2 FLAIR, diffusion weighted imaging) and minimum contrast amount) [2,3]. For brain CTs, intravenous administration of iodine-containing contrast is advised but otherwise no recommendations are made (e.g. minimum number of mAs and minimum contrast dose) [2,3].

In the multicenter NTR3628 study, the additional value of a post-contrast brain MRI was evaluated in stage III (based on 18F-FDG-PET/CE-CT) NSCLC patients. All patients underwent a dedicated brain CE-CT as part of the staging whole body 18F-FDG-PET as standard of care [6]. Imaging requirements were: a standard 18F-FDG-PET/CE-CT protocol that included a diagnostic CE-CT brain, and a 1.5T Gd-MRI brain (1 mm slices, 0.1 mmol/kg gadolinium), with a magnetization transfer contrast (MTC) pre-pulse to increase sensitivity and an additional post-contrast FLAIR sequence. MRI parameters were as recommended by an experienced neuro-radiologist (PH) and followed the American College of Radiology Appropriateness Criteria (ACR AC) [7]. After inclusion of all patients, CE-CTs and MRIs were per protocol centrally reviewed by PH for protocol adherence and presence of brain metastases.

3. Results

3.1. NVALT12 18F-FDG PET/CT imaging sub-study

In the imaging analysis sub-study of the NVALT12, 167 baseline scans and 118 follow-up PET scans for response evaluation were present. Only 97 (34%) of the 285 18F-FDG-PETs performed in this study had an 18F-FDG uptake time as recommended by the EANM. Fifty-four (19%) scans had both uptake times in agreement with the uniform protocols for imaging in clinical trials (UPICT) guidelines. Supplementary material S1 shows the uptake times of the baseline- and the response scans ranked in ascending order for the baseline scan, only the uptake times between the red lines can be used for response assessment. The other investigated parameters in this imaging sub-study, correction factors (attenuation, randoms, scatter) and reconstruction
settings (method, convolution kernel, matrix size) were assessed in a quality score.

In this study, 60% (171/285) of the \(^{18}\text{F}-\text{FDG-PETs} had a SUV\text{mean} of the liver within reference values. Fig. 1 shows an example of how a difference in SUV\text{mean} of the liver between the baseline and the response scan can lead to a visual underestimation of tumor response. Although we expected a non-compliant SUV\text{mean} of the liver for scans with a low quality score, our results did not show such a relation between the quality score and SUV\text{mean} of the liver [8].

### 3.2. Study NTR3628: brain imaging

With respect to dedicated brain imaging, imaging related problems were encountered both during set-up and after central imaging review. Trial set-up was in 2012 and it was planned to have four hospitals involved, but one did not have a MRI scanner that could meet the required technical standards and was excluded from participation. Two other hospitals used MRI protocols that did not meet the required standards (5 mm instead of 1 mm slices, no MTC technique), but they agreed to update their protocols to be able to participate in the study. Two-hundred-fifteen neurologically asymptomatic, extracranial stage III (based on the \(^{18}\text{F}-\text{FDG-PET-CE-CT} \)-NSCLC patients were included between December 2012 and July 2016. After central review, it turned out that in 56 patients (26%) the CE-CT of the brain did not meet the criteria for a dedicated brain imaging protocol. Remarkably, for these patients CE-CT brain results were reported in the radiology report without noting the limitations of the study. Problems with these brain CE-CTs were the field of view (FOV), adjusted to the thorax, but not to the brain and the arm position (above the head instead of at thorax level, see Fig. 1).

### 4. Discussion

Imaging techniques have significantly evolved over the last years and have become more sensitive for the detection of subtle abnormalities. In (lung) cancer guidelines, recommendations regarding imaging techniques are made and even separate detailed \(^{18}\text{F}-\text{FDG-PET-CT} \) recommendations exist \([2,3]\). However, no recommendations are made on the minimal technical requirements, resulting in a large variation in accuracy. Therefore, there is a large variability in imaging parameters and scanners, which can influence not only the clinical trial outcome but also the clinical interpretation by the treating physician. In this manuscript, we report on two trials in which problems were encountered due to variation in applied imaging protocols, which people did not realize beforehand. In NTR3628, after central imaging review, it turned out that the CE-CTs brain of several included patients suffered from unreported serious flaws. In this specific trial, the non-diagnostic CE-CT had no clinical consequences, as all patients underwent a post-Gd-MRI. However, when limitations of an examination are not recognized nor reported, this can have serious consequences (see Table 1). Depending on geographical location and institution, in 48–62% of patients a CE-CT is performed instead of a post-Gd-MRI to screen for brain metastases \([9]\). Although not reported in CE-CT imaging guidelines, is the importance of the patient position: it should be reproducible and standardized, and it matters what the FOV and patient position is. The FOV should be adjusted to the size of the head and arms should be at thorax level to reduce artifacts and acquire diagnostic images (see Fig. 1). Furthermore, just the application of iodine contrast administration does not make the CT-brain diagnostic: timing of the scan relative to contrast injection and the amount of contrast material injected are important. For tumor response assessment using \(^{18}\text{F}-\text{FDG-PET} \), it is important that scans are performed with the same acquisition and reconstruction settings as baseline. Because varying uptake times lead to variations in measured SUV, a larger uptake time will result in a higher SUV because tumor SUV increases over time. In response assessment studies, changes in SUV over time are the parameter of interest and therefore differences in uptake times between repeated scans can lead to scan misinterpretation. Variability in uptake times of the baseline as well as the response scan made that only 46 scans could be used for early response assessment in the NVALT12 sub-study. A study of Kuhnert et al. \([10]\) showed that for settings like the injected activity and the uptake time the SUV\text{mean} of the liver could be used for quality assessment in compliance with the EANM guideline. This means that in the NVALT12 sub-study the large variability in SUV\text{mean} of the liver is most probably caused by the differences in uptake time (although a treatment effect cannot be excluded). In the NVALT12 sub-study no relation between deviation in image parameters, reflected in the quality score, and the SUV\text{mean} of the liver was reported. The SUV\text{mean} of the liver can therefore not exclude all quality issues, although, it could be used as a quick quality control with reference values between 1.3 and 2.5. Our recommendations for clinicians regarding repeated \(^{18}\text{F}-\text{FDG-PET imaging for response assessment are to try to use the same
scanning parameters as for the first scan, or check if the same scan parameters are used as in the previous scan. Scans with an uptake time less than 55 min or more than 75 min cannot reliably be used and checking the SUV\textsubscript{mean} of liver can be used as quality assurance. For data analysis it is important that all scan parameters are noted in the DICOM header, but at least injection time, scan time, patient’s body weight, radionuclide dose and the SUV conversion tag should be present. An example of a DICOM header with the minimum required tags is shown in the supplementary material S2. For brain CE-CT combined with the 18F-FDG-PET scan, we recommend that the clinician checks the head FOV, the patient arm position and the contrast dose and timing.

For MRI, at least a T2-weighted sequence, a T2-FLAIR and a T1-weighted sequence pre- and post-contrast should be performed. See for an overview of these recommendations Table 1. If any of the recommendations cannot be met it is advised to consult an imaging expert.

In addition, radiologists need to be informed about the indication for the requested scan and therefore clinicians must provide all clinical information that is important to determine the optimal imaging protocol to apply. Next to that, for non-imaging clinicians it is important to evaluate whether the applied imaging protocol is indeed adequate for the requested purpose. Our quality recommendations will potentially lead to a more uniform clinical interpretation and could be helpful during multidisciplinary meetings. These recommendations should preferably be tested prospectively.

**Funding**

This work was supported by the Dutch Cancer Society (grant UM 2010-4883).

**Disclosure**

Author AD is advisory board member of Roche, Eli Lilly, Astra Zeneca, MSD and Pfizer. Author LH is advisory board member of Boehringer-Ingelheim and MSD. All remaining authors have declared no conflicts of interest.

**Appendix A. Supplementary data**

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.lungcan.2018.07.014.

**References**


