Introduction

Lung cancer is the second most common cancer in both men and women, and the most common cause of cancer death in the world (1). Non-small cell lung cancer (NSCLC) comprises 80-85% of all lung cancer cases (2). Oncologic treatment options typically include surgery, radiation and chemotherapy, either alone or in combination. The treatment and prognosis currently primarily depend on the stage as defined by the Union for International Cancer Control (UICC)/American Joint Committee on Cancer (AJCC) staging system (3-6). Patients are classified into one of four stages determined by combining the T, N, and M components. Until now, staging and outcome prediction have largely been based on the resectability of the tumor. Within each stage, tumor and patient specific factors vary, creating a heterogenous population of patients, each with an individual prognosis that requires patient and tumor specific factors for best estimation. The goal of this paper is to review developments in evaluating the metabolic tumor burden and its role as a prognostic factor for lung cancer.

Quantitative measurements of tumor activity on FDG-PET imaging

\(^{18}\)F-fluorodeoxyglucose-positron emission tomography (\(^{18}\)F-FDG-PET) in cancer imaging takes advantage of the fact that most neoplasms are highly metabolically active, and can be detected on a background of relatively less active normal tissues. In addition to a qualitative evaluation by visual inspection, \(^{18}\)F-FDG-PET imaging can provide several semi-quantitative or quantitative measurements of radioactivity concentration, such as with the standardized uptake value (SUV) (7). The SUV is defined as the ratio of the FDG concentration in the region of interest (ROI) to the injected dose divided by the patient’s body weight (7,8); it is a frequently used and generally accepted semi-quantitative index for tumor glucose metabolism because of the relative ease of its calculation.

There are many factors affecting the SUV, such as a patient’s body habitus, body composition, blood glucose level, length of uptake period, the partial volume effect, definition of ROI, image reconstruction method, and...
resolution (9-11). Fortunately, most of these factors can be controlled in one institution by following standardized procedures for dose calibration, patient preparation, injection, acquisition, and processing. The patient's body habitus and composition can be controlled by normalizing to lean body weight or body surface area. SUV_{max} defined as maximal SUV in the ROI, has the advantage of being nearly operator-independent, meaning that regardless of how the ROI is drawn, the same value is generated. The most important concern with SUV_{max} is that a mildly active tumor may have a single 'hot' pixel that may arise from random error rather than an actual abnormal uptake in the body.

To overcome random variation from noise rather than an actual abnormal uptake in the body when using the SUV_{max}, SUV_{mean} can be utilized; this is calculated by averaging the SUV values generated from the entire tumor. The drawback of using SUV_{mean} is that differences in operator contouring will yield varying values. Furthermore, a standard approach has not yet been adopted by the nuclear medicine community to guarantee reproducible and accurate results. For small lesions, using the average counts within the ROI causes the SUV_{mean} to be susceptible to a partial volume effect as counts at the edge are averaged with normal surrounding tissue. Different outlines of the ROIs selected by different observers may therefore lead to significant variations in the SUV_{mean}.

More recently, metabolic tumor volume (MTV) and total lesion glycolysis (TLG) have been explored as measures of metabolic tumor burden. MTV indicates the volume of metabolically active tumor, typically assessed with semi-automatic positron emission tomography (PET) analysis software. TLG is the product of SUV_{mean} and MTV; it combines the volumetric and metabolic information of ^{18}F-FDG-PET (12). The metabolic tumor burden for the whole body can subsequently be determined by adding together either the MTV or TLG of the primary tumor, nodal metastases, and distant metastases. Therefore, the whole-body TLG and MTV serve as an index of the overall malignant process in the entire body.

MTV and TLG can be measured manually by nuclear-medicine physicians. However, measuring every tumor's SUV_{mean} and volume by hand is very time-consuming, highly operator-dependent, and impractical in routine clinical settings. This is especially true when a patient has many tumors, which is likely in extensive stage IV metastatic disease. New tumor contouring and segmentation methods, including maximal-intensity threshold and the gradient method, allow for more objective and efficient tumor detection and quantification (13,14). Based on a phantom study, the gradient method appears better than the maximal-intensity threshold method for tumor segmentation and quantification (15). The study showed that for tumor greater than 2 cm, the error in calculating diameter is small (<5%) regardless of the PET-computed tomography (PET-CT) scanner or segmentation method used. However, for smaller tumors (<2 cm), the gradient-based segmentation method yielded significantly lower mean absolute error (8.2%) than the 45%-maximal-intensity threshold method (49.2%). MTV estimates from the gradient method are also expected to be insensitive to the length of FDG uptake before PET imaging, because such volume estimates using the gradient method have been shown to be insensitive to the tumor to background ratio (15). With the semi-automatic gradient method, the radiologist identifies the tumor and selects its major and minor axes. Using spatial derivatives, the software automatically draws volume of interest, which is subsequently manually adjusted by the radiologist in order to include all tumor margins within the volume. Tumor volume is then calculated.

Tumor volume measurements based on PET have been shown to be more reliable than those based on CT or magnetic resonance imaging (MRI). A study in head and neck cancer compared tumor volume measured on CT, MRI, and ^{18}F-FDG-PET scans and found that the PET-derived MTV was the closest to the reference tumor volume from surgical specimens (16). In NSCLC, ^{18}F-FDG-PET/CT fusion images were found to more faithfully reflect pathological tumor size measurement than PET alone or CT alone (17). In addition, numerous other studies have shown that PET/CT is more accurate than CT or MRI for lesion characterization in various cancers, including NSCLC (18-24), which is important because it is often difficult to tell tumor from post-therapeutic change, infection, or inflammatory change on CT and MR images. Finally, oncologic PET/CT scans are routinely performed from “eyes to thighs” (25), thus facilitating whole-body metabolic tumor burden measurements for staging.

**Prognostic value of metabolic tumor burden**

^{18}F-FDG-PET imaging has become the standard of care for the purpose of initial staging of NSCLC, restaging recurrence (26), and monitoring the response to therapy (27). When analyzing the prognostic capability of PET, the most common measure analyzed was the SUV_{max}
of the primary lung tumor (28). Studies have shown that the degree of FDG uptake by the tumor, as assessed with SUV\textsubscript{max}, is a significant prognostic factor in NSCLC (29-32). However, predicting patient prognosis is still predominantly determined with the TNM clinical stage; whether SUV\textsubscript{max} provides prognostic information in addition to that provided by tumor stage in the TNM staging system remains a subject of debate (29,32,33).

Metabolic tumor burden has been shown to have significant and independent prognostic value in patients with lung cancer. Since disease stage is in part based on presence or absence of nodal and distant metastases, it is not surprising that the total number of tumors and nodal metastases are prognostic markers for NSCLC (34). More advanced methods of tumor burden assessment, including MTV and TLG, have also been evaluated in multiple recent studies.

Lee \textit{et al.} performed the first study to show that baseline whole-body MTV measured semi-automatically was a statistically significant prognosticator in 19 patients with lung cancer, and was better than SUV\textsubscript{max} and SUV\textsubscript{mean} (35). A subsequent study by the same group but with a larger group of 61 patients with NSCLC confirmed the inverse association of total body MTV with overall survival and progression-free survival (36). However, MTV was only an independent prognostic factor in the subgroup of patients treated with definitive intent; analysis of the entire cohort revealed that MTV did not have a statistically significant association with survival, likely due to inclusion of those with advanced disease treated only palliatively.

A study of 270 consecutive patients with NSCLC by Dehing-Oberijie \textit{et al.} demonstrated that MTV of the primary tumor and nodal metastases, in combination with the number of positive lymph node stations, is a more important prognostic factor than TNM stage for survival of inoperable NSCLC patients treated with chemoradiation (37).

Zhang \textit{et al.} studied 104 patients with NSCLC who ranged in stage from I to IV but were surgical candidates. MTV\textsubscript{WB} and TLG\textsubscript{WB} were both significantly associated with overall survival and independent of tumor TNM stage and other prognostic factors including patient’s age, gender, chemoradiation therapy and type of surgical procedure received. There was high inter-observer agreement for both measurements with their methods. Performing MTV and TLG was not excessively time-consuming, taking only 3.6 min per case in surgical patients (38). The majority of patients had TNM stage I disease, with less tumor lesions in the body; the time needed for whole-body tumor burden measurement may be dependent on the total number of tumoral lesions.

In a larger study by the same author including 328 surgical and non-surgical patients, baseline MTV\textsubscript{WB} and TLG\textsubscript{WB} both were prognostic markers independent of stage, treatment intent, patient age and gender, as well as tumor histology. Both MTV\textsubscript{WB} and TLG\textsubscript{WB} were significantly better than whole body SUV\textsubscript{max} and SUV\textsubscript{mean} (39).

Two studies by Liao \textit{et al.} demonstrated significant association between overall survival and MTV/TLG in non-surgical patients with NSCLC (40,41). MTV and TLG were prognostic indices independent of tumor TNM stage, not only considering the tumor burden at the whole body level, but also at the level of the primary tumor, nodal disease, and distant metastases. Both TLG and MTV had similar prognostic value, which were better compared to SUV\textsubscript{max} and SUV\textsubscript{mean} of the tumor. There was low inter-observer variability in assessing MTV and TLG measurements. Similarly, Kim \textit{et al.} found MTV and TLG were both correlated with progression-free survival and overall survival in a study involving 91 patients (42).

A study by Chung \textit{et al.} specifically focused on patients with adenocarcinoma. In univariate and multivariate analysis, high MTV and TLG values (≥90 mL and ≥600 g, respectively) were independent predictors of poor overall and progression-free survival in patients with advanced, stage III and IV disease. However, in patients with only stage I and II disease, neither MTV nor TLG were significant prognostic predictors (43).

Hyun \textit{et al.} evaluated preoperative MTV and TLG in 529 patients with early-stage (stage I and II) NSCLC treated surgically. In multivariate analyses, MTV and TLG were significantly associated with an increased risk of recurrence and death, independent of histology, tumor stage, and type of surgery. SUV\textsubscript{max} was not a significant prognostic factor (44).

The same author also considered the value of MTV and TLG in surgical and non-surgical stage III NSCLC patients. Using the Cox proportional hazards models, MTV and TLG were significantly associated with overall survival, independent of histological cell type as well as T stage, N stage, and treatment variables in the surgical group. Again, SUV\textsubscript{max} was not a significant prognostic factor (45).

Chen \textit{et al.} studied the prognostic value of PET measures in 105 patients with NSCLC. Whole-body TLG was significantly associated with overall survival, independent of patient’s TNM stage, tumor histology, age, gender, and performance status, and treatment type. A whole-body
SUV cutoff value of 655 g was determined from ROC analysis to yield specificity of 95%; this value was then used to differentiate patients with high versus low TLG. Using multivariate analysis, only TLG and treatment method (surgery vs. other treatment) proved to have significant prognostic value for overall and progression-free survival. The overall 1-year survival for patients with low TLG was 89% while for those with high TLG it was 42%. The fact that only TLG remained significant in multivariate analysis suggests that it may be a better predictor compared to MTV, perhaps due to a more complete assessment of both tumor volume and metabolic activity. Additionally, their results suggested that whole-body TLG may be a better predictor than TNM stage (46).

Similarly to the above study, Melloni et al. determined that TLG was the only independent prognostic factor of local recurrence in patients with stage I NSCLC based on multivariate analysis (47). Other measures evaluated were MTV and lesion SUV\textsubscript{index} (ratio of lesion SUV\textsubscript{max} to liver SUV\textsubscript{mean}).

MTV and TLG have also been shown to be valuable metrics in small cell lung carcinoma and mesothelioma. Oh et al. reviewed 106 patients with small cell lung cancer who underwent pre-treatment PET. Whole body MTV was an independent predictor of progression and overall survival in both univariate and multivariate analysis. Based on their results, incorporation of whole body MTV into staging could more accurately classify patients into subgroups compared to the traditional staging system, leading to more accurate prognosis and ideally better treatment decisions (48). In a study by Zhu et al., MTV and TLG were both significant prognostic factors of progression and survival in patients with both limited and extensive small cell lung cancer. SUV\textsubscript{max} of the primary tumor did not correlate with survival (49). Several studies have shown prognostic value of MTV in malignant mesothelioma (50-53). Klabatsa et al. showed a significant association of MTV and TLG with overall survival in 60 patients with mesothelioma (53). Nowak et al. demonstrated that total glycolytic tumor volume had better prediction of survival than TNM staging, but only in patients with non-sarcomatoid histology; sarcomatoid histology remained the strongest predictive factor (51).

**Prognostic value of MTV/TLG compared to SUV\textsubscript{max} and SUV\textsubscript{mean}**

TLG and MTV are 3D measures that incorporate both tumor volume and metabolic activity. Therefore, MTV and TLG reflect changes throughout the entire tumor mass and, in theory, should be more accurate methods of detecting global changes than a single-pixel value measure like SUV\textsubscript{max}. Indeed, multiple studies evaluating prognostic value of whole-body metabolic tumor burden including MTV, TLG, and the number of metastatic tumors, have shown that these measures are either more accurate than either SUV\textsubscript{max} and SUV\textsubscript{mean}, or the sole prognostic marker of outcome in NSCLC (34-36,38,39-42,46,54). Although studies have demonstrated the prognostic value of SUV measurement (55-59), it may be less accurate if metastatic disease is present, and does not give more prognostic information compared to tumor size and stage (33,60). Although the prognostic value of SUV\textsubscript{max} may be inferior to MTV or TLG, SUV\textsubscript{max} should still be considered in monitoring disease response.

**Pitfalls of MTV/TLG measurements**

Most lung cancers are hypermetabolic and therefore result in increased tumor to background ratio for MTV/TLG measurements. One pitfall could be encountered with mucinous adenocarcinoma, also known as bronchoalveolar carcinoma; this tumor type is known to often have no increased FDG activity, which would falsely underestimate extent of disease, and, therefore, the patient prognosis. False-positive PET results may be due to active infection or inflammation, including granulomatous diseases such as sarcoidosis, histoplasmosis or tuberculosis, as well as metabolically active brown fat. Therefore, PET images (attenuation-corrected and non-attenuation-corrected), as well as the diagnostic CT images should be evaluated together before contouring tumor and measuring MTV and TLG values.

**Clinical implications of tumor burden measurements**

As discussed above, multiple studies have shown that whole body tumor burden, best assessed with whole-body MTV and TLG, has significant prognostic value in patients with lung cancer, which is independent of TNM stage. Therefore, the addition of tumor burden measurements can help further stratify patients within each stage and optimize treatment method.

The components of the TNM staging system attempt to roughly approximate the whole-body tumor volume by describing primary tumor size and local involvement...
(T), extent of nodal disease (N), and presence of metastatic disease (M), with the primary goal of assessing resectability. However, with more advanced disease, the accuracy of the TNM system as a surrogate for overall tumor burden breaks down due to a wide spectrum of disease severity represented by only a few different stages. For example, a patient with only a single extrathoracic metastatic lesion would have the same stage as a patient with diffusely metastatic disease. While both patients would not be surgical candidates based on their stage, their prognosis and treatment may not necessarily be the same.

PET based measures such as whole-body MTV and TLG can provide a more complete estimation of the true volume and biological aggressiveness, and can contribute a quantitative prognostic measure to the TNM staging system. The more accurate risk stratification may aid clinician and patient decision making for optimal treatment choices and better outcome prediction. Furthermore, a quantifiable prognostic factor may better define patient grouping for clinical trials. Incorporating metabolic tumor burden in trials and staging could help subselect patients groups which would most benefit from adjuvant or neoadjuvant chemotherapeutic therapies.

It is incompletely clear which measure, MTV or TLG, is more superior in prognostication. While some studies have shown these to be equivalent (39), others have argued the superiority of TLG given that it reflects both tumor volume and degree of metabolic activity (46,47). However, with either MTV or TLG, the value based on the whole-body assessment, tumor which includes the primary tumor, nodal and distant metastases, should be used rather than the value based solely on the primary tumor.

Future research

There is no consensus on how exactly metabolic tumor burden measurements should be used in clinical practice. It is also unclear how sensitive MTV and TLG are to FDG uptake time and what effect different PET/CT scanners and reconstruction methods have on their values. These questions need to be addressed with additional research before wider application in patient management can take place, including prospective clinical trials utilizing metabolic tumor burden measures. An additional future goal is the development of more reliable computer assisted diagnostic (CAD) systems which will enable automated accurate and reproducible values.

Conclusions

Whole-body tumor burden assessment with MTV and TLG has significant prognostic value in patients with lung cancer. These markers are independent of stage and other clinical prognostic factors, and of better prognostic value than either SUV_{max} or SUV_{mean}.

Acknowledgements

Disclosure: The authors declare no conflict of interest.

References

12. Larson SM, Erdi Y, Akhurst T, et al. Tumor treatment response based on visual and quantitative changes in global


37. Dehing-Oberije C, De Ruysscher D, van der Weide H, et al. Tumor volume combined with number of positive lymph node stations is a more important prognostic factor than TNM stage for survival of non-small-cell lung cancer patients treated with (Chemo)radiotherapy. Int J Radiat Oncol Biol Phys 2008;70:1039-44.


55. Pelosi E, Billè A, Skanjeti A, et al. Prognostic role of...


Lung Cancer 2011;71:49-54.


