Introduction

In cancer immunotherapy, the patient’s own adaptive immunity is harnessed to detect and eradicate tumor cells. Current strategies involve the utilization of adoptive T-cell transfer, oncolytic viruses, cancer vaccines and immune checkpoint inhibitors. These latter are monoclonal antibodies directed against immune checkpoint proteins that regulate immune homeostasis and are essential to self-tolerance. Neoplastic cells are hypothesized to evade immune surveillance, notably by expressing at their surface a variety of immune checkpoint proteins that negatively regulate T-cell activity [reviewed in (1)]. Several receptor-ligand immune checkpoints have been identified, such as the receptor programmed cell death protein-1 (PD-1) interacting with its ligands programmed death-ligand 1,2 (PD-L1,2). Current immunomodulatory anticancer drugs for non-small cell lung cancer (NSCLC) target the PD-1/PD-L1,2 axis and the receptor cytotoxic T-lymphocyte-associated antigen 4 (CTLA4). Nivolumab and pembrolizumab are highly specific antibodies that bind PD-1, while atezolizumab, avelumab and durvalumab target

Methods of measurement for tumor mutational burden in tumor tissue

Bárbara Meléndez*, Claude Van Campenhout*, Sandrine Rorive, Myriam Remmelink, Isabelle Salmon, Nicky D’Haene

Department of Pathology, Erasme Hospital, Université Libre de Bruxelles, Brussels, Belgium

Contributions: (I) Conception and design: B Meléndez, C Van Campenhout, N D’Haene; (II) Administrative support: B Meléndez, C Van Campenhout, N D’Haene; (III) Provision of study materials or patients: None; (IV) Collection and assembly of data: B Meléndez, C Van Campenhout, N D’Haene; (V) Data analysis and interpretation: B Meléndez, C Van Campenhout, N D’Haene; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

*These authors contributed equally to this work.

Correspondence to: Dr. Nicky D’Haene. Department of Pathology, Erasme Hospital, 808 Route de Lennik - B-1070, Brussels, Belgium. Email: nicky.d.haene@erasme.ulb.ac.be.

Abstract: Immunotherapies based on immune checkpoint inhibitors are emerging as an innovative treatment for different types of advanced cancers. While the utility of immune checkpoint inhibitors has been clearly demonstrated, the response rate is highly variable across individuals. Due to the cost and toxicity of these immunotherapies, a critical challenge in this field is the identification of predictive biomarkers to discriminate which patients may respond to immunotherapy. Recently, a high tumor mutational burden (TMB) has been identified as a genetic signature that is associated with a favorable outcome for immune checkpoint inhibitor therapy. The TMB is defined as the total number of nonsynonymous mutations per coding area of a tumor genome. Initially, it was determined using whole exome sequencing, but due to the high costs and long turnaround time of this method, targeted panel sequencing is currently being explored to measure TMB. In the near future, TMB evaluation may play an important role in immuno-oncology, but its implementation in a routine setting involves robust analytical and clinical validation. Standardization is also needed in order to make informed decisions about patients. This review presents the methodologies employed for determining TMB and discusses the factors that may have an impact on its measurement.

Keywords: Tumor mutational burden (TMB); whole-exome sequencing (WES); gene panels; methods

View this article at: http://dx.doi.org/10.21037/tlcr.2018.08.02
PD-L1 and ipilimumab targets CTLA-4. Treating solid tumors with immune checkpoint inhibitors can be highly effective and can lead to durable responses in approximately one in five patients with NSCLC, melanoma or renal-cell cancer (2). Further studies confirmed the utility of anti-PD-1 antibodies in the treatment of cases of advanced NSCLC, with an objective response rate (ORR) of 19.4% and a median duration of response of 12.5 months (3). However, despite this major oncologic advance, immune checkpoint inhibitors are not universally successful across different patients and tumor types and the response rate is highly variable, ranging from complete to no response. Only a minority of patients achieves long-term survival and there is thus a great need to find predictive biomarkers for patient stratification and selection.

At present, PD-L1 expression assay is the most used biomarker to predict a favorable response to immune checkpoint inhibitors. Current guidelines for NSCLC recommend testing for actionable molecular alterations of EGFR, ALK, ROS1, BRAF and KRAS genes and immunohistochemistry testing for PD-L1 expression. In the absence of an identified mutated or rearranged driver oncogene and when PD-L1 expression is ≥50% (representing ~30% of the NSCLC patients), pembrolizumab is indicated as the initial treatment (National Comprehensive Cancer Network. NSCLC, version 3. 2018; https://www.nccn.org/professionals/physician_gls/default.aspx). However, PD-L1 expression is considered to be an “enrichment factor”, meaning that the higher the percentage of PD-L1 positive tumor cells, the higher the likelihood of response to the immune checkpoint inhibitors targeting the PD-1/PDL-1 axis, but some patients that are negative do in fact can respond. Moreover, PD-L1 immunohistochemistry could be challenging for several reasons: (I) it is a continuous biomarker with a range of expression levels that complicates its evaluation; (II) it is a heterogeneous marker, with intratumoral and temporal variation of expression, and (III) different assays have been developed with different scoring criteria and different positivity thresholds.

Among new predictive biomarkers for response to immunotherapy, the mutational load or tumor mutational burden (TMB) has been shown to correlate with the response to immunotherapy treatment (4-6). TMB is a quantitative measure of the total number of somatic nonsynonymous mutations per coding area of a tumor genome. It has been hypothesized that tumors with a higher mutation burden are more likely to express neoantigens and to induce a more robust immune response in the presence of immune checkpoint inhibitors (7). Recently, the CheckMate 227 trial (NCT02477826) demonstrated a longer progression-free survival (PFS) with nivolumab plus ipilimumab versus chemotherapy in first-line treatment of NSCLC patients whose tumors had a high (≥10 mut/Mb) TMB regardless of PD-L1 expression (8).

It is worth noting that the implementation of a new biomarker such as TMB involves robust analytical and clinical validation. Robustness of sample processing and methodology are important factors that influence the reproducibility of TMB measurement. Moreover, in the case of this new genome-based biomarker, bioinformatic analysis is also a critical determinant of successful clinical implementation, because TMB measurement is based on next-generation sequencing (NGS) techniques. Further, across studies, different methods have been used to calculate and report TMB. Here we review the methodologies employed for determining the levels of TMB and discuss the factors that may have an impact on its measurement.

### TMB from whole-exome sequencing (WES)

Initial TMB measurements were performed using WES of tumor tissue and matched non-tumor tissue using NGS technology. In humans, the exome spans 30–50 megabases of coding sequences that correspond to less than 2% of the genome (9). WES provides a clear landscape of most coding mutations that can contribute to tumor progression, making it a cost-effective alternative to whole-genome sequencing (WGS) (6,10-12). Most WES protocols require a minimum of 150–200 ng of genomic DNA, an amount that can prevent its use in small biopsies. Moreover, sequencing of a tumor sample and matched normal samples is highly recommended for identifying tumor-specific variants, which increases the cost of the test.

The ability to detect somatic variants with NGS is primarily dependent upon the variant frequency within the tumor as well as on the sequencing depth. Theoretically, at 50x coverage, 95% of single nucleotide variants (SNVs) and short insertions or deletions (INDELs) with a variant allele frequency of 15% or higher can be consistently detected (13). However, contamination of non-tumor cells, tumor heterogeneity and aneuploidy are additional complexity factors. Therefore, deeper sequencing may be advisable to maintain high sensitivity in samples with lower purity or higher heterogeneity.

The Cancer Genome Atlas (TCGA) project and several
other studies have used WES to measure TMB across cancer types and found a wide distribution of TMB (more than 1,000-fold variation) across approximately 20–30 cancer types. A very low TMB, about 0.1 mutations per Mb, is observed in certain pediatric cancers. At the opposite end of the spectrum, more than 400 mutations per Mb were identified in some cancers related to extensive exposure to carcinogens such as malignant melanoma (exposure to ultraviolet light) or lung cancer (exposure to tobacco smoke) (14-16). TMB variability is particularly broad in NSCLCs. As expected, the median of somatic mutations per Mb is higher in NSCLC patients that are former/current smokers (~10.5 mutations/Mb) than in never smokers (0.6 mutations/Mb) (17). Nevertheless, smoke carcinogens alone do not account for the wide range of the mutational burden in NSCLCs. Additional factors, such as mutations in genes important for DNA repair and replication, are known to contribute to the accumulation of somatic mutations.

In 2015, Rizvi and colleagues showed that for advanced NSCLC patients treated with an anti-PD-1, a high TMB strongly associates with clinical efficacy. The authors performed WES on two independent cohorts of NSCLC patients treated with pembrolizumab (discovery, n=16; validation, n=18). For both cohorts, tumor DNA sequencing generated a mean target coverage of 164× and a mean of 94% of the target sequence was covered to a depth of at least 10×. A higher nonsynonymous mutation burden in tumors was associated with an improved objective response, a more durable clinical benefit (DCB) and a longer progression-free survival (PFS). The rate of DCB in patients harboring ≥178 mutations was 75% compared to 17% in those with <178 (5). In the CheckMate 026 trial (NCT02041533) of nivolumab vs. chemotherapy in first-line NSCLC, the TMB scores were determined by WES for tumor and matched whole blood DNA in 312 patients. In patients with TMB ≥243 mutations, PFS was longer (median: 9.7 vs. 5.8 months) and the ORR was higher with nivolumab than with chemotherapy (47% vs. 28%) (18). In sum, while a high TMB score is clearly associated with a more durable clinical benefit (DCB) and a longer progression-free survival (PFS), the rate of DCB is significantly higher in patients with a high TMB compared to those with a low TMB. TMB variability is particularly broad in NSCLCs.

In an exploratory analysis of 240 patients with advanced NSCLC treated with either PD-1 or PD-L1 inhibitors at Memorial Sloan Kettering Cancer Center (MSKCC), the MSK-IMPACT targeted NGS assay (341, 410 or 468 genes, targeting ~1.1 Mb of coding genome, 315 genes) accurately assessed the TMB, compared with sequencing the whole exome (21). From both the WES and targeted cancer gene panel samples, the numbers of somatic, coding and base substitutions and INDEL mutations per megabase of interrogated genome (with the exclusion of germline polymorphisms) were calculated. They found that the TMBs calculated by these two assays were highly correlated. It should be noted that the tumor samples must contain at least 20% tumor cells and that at least 50 ng of FFPE DNA is needed.

**TMB from gene-targeted sequencing**

Measuring mutation load using WES can be difficult due to its high cost and extensive analysis and data management requirements. Implementing a test in a clinical setting is challenging because (I) the test must be performed on routine samples with a sometimes limited amount of DNA, (II) the test results must be delivered within an adequate time, (III) the test results must be accurate and facilitate clinical decision-making and (IV) the test must be affordable. In these senses, gene panels may be preferable because of lower sequencing costs, lower DNA input requirements and shorter turnaround time (TAT). Furthermore, gene panels can generate deeper sequencing compared to WES and could potentially improve mutation detection sensitivity when using a small amount (a few nanograms) of DNA or even fragmented DNA. This is an important aspect to consider, as in daily practice, the use of formaldehyde fixation, the small sample size and the highly variable tumour cell content could be limiting factors.

Although TMB has been traditionally assessed by WES, initial “in silico” studies using WES data concluded that TMB can be estimated using large cancer gene-targeted sequencing panels (19,20). In a cohort of 29 tumors, Chalmers and collaborators showed that targeted comprehensive genomic profiling with FoundationOne assay (targeting ~1.1 Mb of coding genome, 315 genes) accurately assessed the TMB, compared with sequencing the whole exome (21). From both the WES and targeted cancer gene panel samples, the numbers of somatic, coding and base substitutions and INDEL mutations per megabase of interrogated genome (with the exclusion of germline polymorphisms) were calculated. They found that the TMBs calculated by these two assays were highly correlated. It should be noted that the tumor samples must contain at least 20% tumor cells and that at least 50 ng of FFPE DNA is needed.
Recently, a commercially available gene panel was released, i.e., the Oncomine Tumor Mutation Load Assay (Life Technologies), targeting 1.7 Mb of the genome in 409 key cancer genes. The performance of this targeted panel for TMB was assessed using a set of 24 FFPE tumor samples, including two samples from NSCLC tumors. The TMB was determined from 20 nanograms of DNA by counting the somatic SNVs per Mb (custom filtering) that (I) were exonic; (II) had a mutant allele frequency >10% (to avoid counting the variants caused by FFPE artefacts); (III) were not found in either the dbSNP (to exclude germline mutations) or in the COSMIC database (to eliminate the potential bias of studying cancer-related genes in the panels). The results showed a good correlation between TMB values derived from WES and those from the targeted panel. However, the authors highlighted the difficulties in the analysis of some FFPE samples with many artefacts, suggesting the use of duplicates (reducing the depth coverage by half in order to maintain costs) and underlining the importance of the filtering pipeline (23,24).

Therefore, to date, the TMB results obtained from large cancer gene-targeted panels demonstrate the feasibility of using TMB as a biomarker that can extrapolate the genome-wide mutation load. Various commercial companies propose TMB evaluation in combination with genomic profiling for targeted therapies using panels varying between 315 and >600 genes at a cost varying between 4,800 and 6,500 USD (25). However, not only analytical, but also clinical validations for the calculation of TMB are required for all of these available panels and thus the results obtained by these methods should be viewed with caution if the platforms have not been clinically validated.

In summary, a high TMB has been defined at various thresholds and with different NGS targeted panels. Further, the molecular testing platforms currently in use also vary widely in the composition of the targeted genes: hence, the harmonization of TMB measurement methodology across platforms and the definition of the thresholds are necessary steps to implement in the clinic as part of an approach to personalized medicine (26).

Factors with an impact on TMB calculation

Tumor cell content and sequencing coverage

Targeted panels enable deeper sequencing compared to WES. That means that higher sensitivity is achieved when using these targeted platforms even when the tumor cell content of the sample is low (<10%) and especially when the depth of sequencing (coverage) increases (500–1,000×). The initial pilot studies that inferred whether targeted-gene panels would be suitable for calculating TMB did not consider this factor, as these models used data from WES (19,20). Recent studies have achieved much higher coverage in targeted panels than in WES. For example, the recent study by Rizvi and collaborators reached a mean sequencing coverage across all tumor samples of 744x when using the MSK-IMPACT panel, while when using WES the mean target coverage was 232x in tumor sequences (22). This may imply that, by using gene-targeted panels, data on TMB may be obtained from samples with a low tumor cell content and that lower frequency variants may be detected. Nevertheless, the clinical impact of this higher sensitivity obtained by increasing the sequencing coverage has not yet been studied in depth.

Preanalytic factors

Fixation of tissues in buffered formaldehyde is the standard procedure because it preserves tissue and cellular morphology. Nevertheless, formaldehyde induces various crosslinks (protein–protein, protein–DNA, interstrand DNA crosslinks, etc.) and is a well-known source of sequencing artifacts (apparent sequence changes that are not present in the original sample). The most prevalent alterations in FFPE samples are DNA fragmentation, denaturation and deamination of cytosine bases. Cytosine deamination is the major cause of baseline noise in NGS, leading to the most prevalent sequence artifacts in FFPE (C:G > T:A). The existence of these artefactual alterations can complicate the analysis of NGS results, particularly when analyzing low-frequency mutations. The impact of sequence artifacts in FFPE DNA can be minimized through: pre-analytical assessment to identify tumor-rich areas for macrodissection, decross-linking by heat treatment prior to DNA extraction (use a FFPE dedicated DNA extraction kit), short amplicon design, HIFI polymerase use, sequencing of both strands of DNA, capture-based sequencing approaches and frequency and variant filtering [reviewed in (27)]. This issue, however, could be highly challenging for clinical laboratories.

Size of the gene-targeted panel and the genes included

Several sizes of panels have been tested to determine TMB. The most widely used panels across the studies include the MSK-IMPACT panel, which in the latest version targets
468 genes (1.22 Mb of the genome), and the Foundation Medicine Panel, which targets 315 genes (1.2 Mb). Two other commercially available panels have also been developed recently, the Oncomine Tumor Mutation Load Assay (Life Technologies; 409 genes, 1.7 Mb) and the TruSight Tumor 170 (Illumina; 170 genes, 0.524 Mb), although these commercial panels have not been widely used to date.

It has been shown that as panels become smaller in size, the zone of uncertainty associated with TMB estimation increases rapidly. Additionally, the coefficient of variance increases rapidly when the size of the targeted panels is less than 1 Mb (21). Therefore, the minimum size of the panel for determining the TMB of more than 300 genes or 1 Mb has been calculated (20,21). Regarding the selected genes included in these panels, only around 50% of the genes included in the MSK-IMPACT platform are also present in the FoundationOne panel, with ~190 genes also common to the Oncomine Tumor Mutation Load Assay. The impact of the mutational study of different genes on TMB calculation has not been analyzed yet.

**Bioinformatic pipeline**

Although the higher occurrence of synonymous variants may indicate a mutational process that also results in nonsynonymous changes, in general, synonymous and germline variants are discarded in the calculation of TMB, as it is assumed that these variants are not likely to be directly involved in creating neoantigens. In order to filter out germline variants, the ideal situation would be to sequence a matched non-tumor sample from each patient. However, in a clinical practice setting, the availability of this matched sample may vary across different institutions and diverse organizational factors may hamper the feasibility of maintaining short TATs and the affordability of the test.

In the setting of tumor-only sequencing, germline false positive variants may be filtered out by using large, publicly available germline variant data sets. The use of these germline databases is a critical step in this process, as it is necessary to use germline databases with a sufficiently broad representation of all populations. Otherwise, patients with ethnic backgrounds who are underrepresented in these databases may have elevated rates of germline false positive mutations. This fact was emphasized in the work of Garofalo and coworkers, which found an increased rate of false positive germline variants in non-white ancestries when using the dbSNP and 1,000 Genomes databases compared to the ExAC public database (20). Furthermore, these authors suggested that manual interpretation of the variants by a molecular pathologist was highly effective in reducing these false positive errors. This manual filtering, however, is time-consuming for implementation in a routine clinical setting.

There is a certain degree of variability in the TMB calculation between the different studies, because some of them consider all somatic alterations, including nonsynonymous mutations and copy number alterations (22,28), while others considered only sequence mutations (29,30). Similarly, most of the studies removed the variants included in COSMIC in order to avoid bias toward genes with functional mutations in cancer or in order to just exclude alterations likely or known to be bona fide oncogenic drivers and germline polymorphisms.

Qiu and colleagues compared WES data and mutations reported by multiple commercial sources based on the same set of three matched tumors and normal samples (31). Two of the samples were obtained from FFPE tumor blocks (one was expected to have a high TMB) and the third one was a mix of a tumor and a normal cell line with a known genomic profile that mimicked 30% tumor content. A large discrepancy was observed among the number of mutations reported by different vendors, underlining the need for standardization. For instance, vendors reported a wide range of numbers of mutations, varying up to 2.3-fold (after removing overlaps between replicates). In a bioinformatic pipeline, the estimation of mutation rates is affected by the aligner that is used to align reads to a reference genome and, more importantly, by the variant caller or parameters that are used to identify single-nucleotide variants. Given the major potential effect of the latter, standards should be established. Interestingly, Qiu and colleagues showed that the analysis of raw data (FASTQ files) from the different WES commercial sources with the same data analysis pipeline [Genome Analysis Toolkit (GATK)/Mutect-based pipeline] gave more consistent results.

**Cut-offs**

Several studies reported a wide range of TMBs across many cancer types, with melanoma and lung tumors representing two of the cancer types with the highest TMB (15,16,21). In the lung, the median TMB across more than 18,000 lung cancer cases was 7.2 mutations/Mb, with approximately 12% of the patients showing more than 20 mutations/ Mb (21). Nevertheless, cut-offs may not be coincident for
all tumor types and clinical cut-offs should be accurately established for each tumor type in order to translate the use of TMB biomarker into clinical practice.

The initial WES study from Rizvi and collaborators in advanced NSCLC patients defined 178 nonsynonymous mutations [approximately 5 mutations/Mb if we use 38 Mb as the estimate of the exome size (21)] as the cut-point that combined maximal sensitivity (100%) with maximal specificity (67%); this cut-point was associated with a likelihood ratio of a DCB of 3.0 (5). Note that the study of Campesato and collaborators using the same WES data yielded median numbers of nonsynonymous somatic mutations per Mb (calculated by analyzing only the genes contained in the Foundation Medicine panel) of 9 and 5 for tumors from patients with a DCB and with a non-durable clinical benefit (NDCB), respectively (19). Later on, by using data from targeted NGS (MSK-IMPACT panel), Rizvi and collaborators found that the TMB was greater in advanced NSCLC patients with a DCB than in those with an NDCB (median: 8.5 vs. 6.6 mutations/Mb). TMB was also greater in patients with a complete response or partial response versus those with stable disease and those with progressive disease (median of 8.5, 6.6 and 6.6 mutations/Mb, respectively) (22). In the phase 3 CheckMate 026 study, exploratory analyses were performed to determine the value of TMB as a predictive biomarker in 312 NSCLC tumor samples using WES. A TMB in the upper tertile (≥243 mutations, approximately 6 mutations/Mb, a high burden) was associated with an increased ORR to nivolumab and a longer median PFS, compared to a TMB in the lowest 2 tertiles (<100 mutations, approximately 3 mutations/Mb, a low burden; and 100–242 mutations, a medium burden) and compared to a TMB of patients treated with chemotherapy in the upper tertile (18).

Finally, in the CheckMate 227 phase 3 trial, the TMB was evaluated using the FoundationOne CDx panel. Here, a cut-off of 10 mutations per Mb of DNA was associated with an ORR of approximately 44%. Increasing this cut-off to 15 mutations/Mb did not identify patients with a greater likelihood of response, and a lower cut-off was associated with significantly lower response rates.

Therefore, in patients with advanced NSCLC whose first-line treatment is the combination immunotherapy (nivolumab-ipilimumab), a tumor mutational burden greater than or equal to 10 mutations/Mb was identified as the optimal cut-off for benefits due to combination immunotherapy, showing a longer PFS than chemotherapy treatment (8).

Overall, from these different studies, we observe that the optimal cut-offs to establish a tumor with a high TMB have yet to be defined. Three categories of TMB seem to be emerging, one clearly high (>10 mutations/Mb), one clearly low (<5) and the last “undetermined”, for which more studies are needed in order to evaluate the response to immune checkpoint inhibitors.

**Conclusions**

As already underlined by others authors, evaluation of the TMB may play an important role in immuno-oncology (26). However, implementation of this test in a routine setting is still challenging. Sample size and DNA amount, varying testing platforms, varying bioinformatic pipelines, cut-off definition and costs are factors that potentially limit the implementation of the TMB evaluation. Moreover, standardizing the method of TMB measurement is needed in order to ensure reliability, reproducibility and clinical utility.

**Acknowledgements**

This work was supported by funding from the “Fonds Yvonne Boël” (Brussels, Belgium).

**Footnote**

Conflicts of Interest: The authors have no conflicts of interest to declare.

**References**


Cite this article as: Meléndez B, Van Campenhout C, Rorive S, Remmelink M, Salmon I, D’Haene N. Methods of measurement for tumor mutational burden in tumor tissue. Transl Lung Cancer Res 2018;7(6):661-667. doi: 10.21037/tlcr.2018.08.02

© Translational lung cancer research. All rights reserved.