GOOD SCIENCE
BETTER MEDICINE
BEST PRACTICE

ORIGINAL ARTICLE

Optimizing panel-based tumor mutational burden (TMB) measurement

J. Budczies[1,2]*, M. Allgäuer[1], K. Litchfield[3], E. Rempel[1], P. Christopoulos[4,5], D. Kazdal[1,5], V. Endris[1], M. Thomas[4,5], S. Fröhling[6,7], S. Peters[8], C. Swanton[3], P. Schirmacher[1,2] & A. Stenzinger[1,2]*

[1]Institute of Pathology, University Hospital Heidelberg, Heidelberg; [2]German Cancer Consortium (DKTK), Partner Site Heidelberg, Heidelberg, Germany; [3]Cancer Evolution and Genome Instability Translational Cancer Therapeutics Laboratory, Francis Crick Institute, London, UK; [4]Department of Thoracic Oncology, Thoraxklinik at Heidelberg University Hospital, Heidelberg; [5]German Center for Lung Research (DZL), Translational Lung Research Center Heidelberg (TLRC-H), Heidelberg; [6]Department of Translational Oncology, National Center for Tumor Diseases (NCT), Heidelberg; [7]German Cancer Research Center (DKFZ), Heidelberg, Germany; [8]Department of Oncology, Centre Hospitalier Universitaire Vaudois (CHUV), Lausanne University, Switzerland

Correspondence to: Dr Jan Budczies, Institute of Pathology, University Hospital Heidelberg, Im Neuenheimer Feld 224, 69120 Heidelberg, Germany. Tel: +49-6-22156-32757; E-mail: jan.budczies@med.uni-heidelberg.de

Prof. Albrecht Stenzinger, Institute of Pathology, University Hospital Heidelberg, Im Neuenheimer Feld 224, 69120 Heidelberg, Germany. Tel: +49-6-22156-34380; E-mail: albrecht.stenzinger@med.uni-heidelberg.de

Background: Panel sequencing based estimates of tumor mutational burden (psTMB) are increasingly replacing whole exome sequencing (WES) tumor mutational burden as predictive biomarker of immune checkpoint blockade (ICB).

Design: A mathematical law describing psTMB variability was derived using a random mutation model and complemented by the contributions of non-randomly mutated real-world cancer genomes and intratumoral heterogeneity through simulations in publicly available datasets.

Results: The coefficient of variation (CV) of psTMB decreased inversely proportional with the square root of the panel size and the square root of the TMB level. *In silico* simulations of all major commercially available panels in the TCGA pan-cancer cohort confirmed the validity of this mathematical law and demonstrated that the CV was 35% for TMB = 10 muts/Mbp for the largest panels of size 1.1–1.4 Mbp. Accordingly, misclassification rates (gold standard: WES) to separate 'TMBhigh' from 'TMBlow' using a cut-point of 199 mutations were 10%–12% in TCGA-LUAD and 17%–19% in TCGA-LUSC. A novel three-tier psTMB classification scheme which accounts for the likelihood of misclassification is proposed. Simulations in two WES datasets of immunotherapy treated patients revealed that small gene panels were poor predictors of ICB response. Moreover, we noted substantial intratumoral variance of psTMB scores in the TRACERx 100 cohort and identified indel burden as independent marker complementing missense mutation burden.

Conclusions: A universal mathematical law describes accuracy limitations inherent to psTMB, which result in substantial misclassification rates. This scenario can be controlled by two measures: (i) a panel design that is based on the mathematical law described in this article: halving the CV requires a fourfold increase in panel size, (ii) a novel three-tier TMB classification scheme. Moreover, inclusion of indel burden can complement TMB reports. This work has substantial implications for panel design, TMB testing, clinical trials and patient management.

Key words: tumor mutational burden, TMB, immune-oncology, panel sequencing, immune checkpoint blockade

Introduction

The immune system plays a central role in cancer recognition and control. It does not only fight virus-driven tumors and limits pro-tumorigenic states of inflammation, but also monitors cells for expression of neoantigens and cell stress induced proteins [1].

Upregulation of immune checkpoint proteins such as programed death ligand 1 (PD-L1) and cytotoxic T-lymphocyte associated protein 4 (CTLA-4) represents a mechanism for cancer cells to escape immune surveillance. Therapeutically, this can be overcome by immune checkpoint blockade (ICB); however, only a subset of

patients shows good clinical response and extended progression free survival under such therapies. Due to potentially severe adverse effects [2] and substantial costs it is paramount to prospectively identify patients that most likely benefit from such treatment strategies.

Currently, PD-L1 expression as assessed by immunohistochemistry is the most widely adopted and approved predictive biomarker for ICB [3] but its predictive power, especially negative predictive value, is limited [4]. Consequently, additional markers [5–7] including tumor mutational burden (TMB) are considered. The more mutations a tumor accumulates, the higher the likelihood of production and subsequent presentation of neoantigens on major histocompatibility complex (MHC) molecules resulting in a higher likelihood of tumor cell cytotoxicity after inhibition of checkpoint signals [8]. This paradigm is supported by accumulating evidence that tumors with higher TMB are more likely to respond to ICB in various settings including PD-(L)1 blockade in non-small-cell lung cancer (NSCLC) [9] and urothelial carcinoma [10], CTLA-4 blockade in malignant melanoma [11, 12] and combined PD(L)-1 and CTLA-4 blockade in NSCLC [13–15] and small-cell lung cancer (SCLC) [16]. Studies have shown that TMB is to a large extent independent of PD-L1 status and might thereby identify additional subgroups of patients who benefit from ICB [13, 15, 17–19].

The definition of cut-points to separate 'TMBhigh' from 'TMBlow' tumors is not consistent in recent NSCLC trials: For example, in the CheckMate (CM) trials CM012 [15], CM227 [13] and CM026 [17] cut-points of 158 mutations, 199 mutations (estimated from a panel-based cut-point of 10 mutations per Mbp [20]) and 243 somatic missense mutations were used, respectively.

While TMB was accurately measured by whole exome sequencing (WES) in several studies, this is currently not feasible in a routine clinical setting due to high costs, long turnaround times and limited availability of sufficient tissue samples. At the same time, panel-based sequencing of routinely available formalin-fixed and paraffin-embedded tissue samples has been implemented at many clinical centers. Therefore, stakeholders of academia and industry are working on implementing assays and workflows to reliably extrapolate TMB from panel sequencing data (psTMB). Many parameters influence psTMB measurement including pre-analytical factors, the assay itself and the bioinformatics analysis pipeline [21–25].

Here, we address key issues regarding the clinical implementation of psTMB measurement: First, using a random mutation model, we derive an algebraic formula for the coefficient of variation (CV) of psTMB as a function of panel size and number of detected mutations. Secondly, by the simulation of panel sequencing in WES data, we quantify psTMB variability, assess the degree of imprecision attributable to intratumoral heterogeneity, and analyze the capability of psTMB to predict response to ICB. Thirdly, we introduce a three-tier TMB classification scheme and shown how it can attenuate imprecision inherent to psTMB. Finally, we analyze how the inclusion of synonymous mutations, nonsense mutations and indels additionally to missense mutations can improve the precision of psTMB estimates.

Materials and methods

An algebraic formula for the CV of psTMB was derived assuming binomial distribution of the number of detected mutations. Panel sequencing

using five commercially available panels (Table 1) was simulated in WES data from the TCGA pan-cancer cohort, the TRACERx 100 cohort and two cohorts of ICB treated patients [13, 26–28]. Statistical computing and graphics generation was carried out with the programing language R. *P*-values <0.05 were considered significant. Statistical methods are described in detail in the supplementary Appendix S1 (available at *Annals of Oncology* online).

Results

psTMB measurement extrapolates the total number of mutations in the coding sequence (CDS) by analysis of a limited sequence part. The largest commercially available panels cover between 1.0 and 1.5 Mbp, <5% of the total CDS. A multi-step analysis was carried out to decompose different sources of imprecision inherent in psTMB estimates.

Random mutation model

Assuming that each base is mutated with the same probability, the number of mutations detected by panel sequencing is a random variable following a binomial distribution. This assumption reflects many current psTMB filtering approaches that disregard evolutionary selected genes (e.g. classic oncogenes and tumor suppressors) for TMB scores. Evaluating this binomial model, we derived a mathematical formula for the CV of psTMB: the CV inversely proportional to the square root of the panel size and inversely proportional to the square root of the TMB level. Figure 1A shows how the CV decreases when the panel size increases for tumors with TMB scores = 1, 3, 10, 30 and 100 muts/Mpb. Figure 1B shows how the CV decreases when the TMB level increases (exemplified for the major five commercially available panels investigated in the present study plus the F1 CDx panel by Foundation Medicine). For a tumor of 10 muts/Mbp, the CV of psTMB ranged between 69% and 27% for panel sizes ranging between 0.21 and 1.34 Mbp (Table 1).

Real-world cancer genomes

Next, we simulated psTMB measurement with the five major sequencing panels in the pan-cancer TCGA cohort of >10 000 tumors (Figure 2, supplementary Figures S1 and S2, available at *Annals of Oncology* online). Strong correlations were observed between the numbers of mutations detected by panel sequencing and the reference standard WES: Pearson correlation reached 0.95 for the OCAv3 panel, 0.97 for the TST170 panel and 0.99 for the three large panels (Figure 2A, D and G). Linear models were fitted to analyze the correlation of panel-approximated TMB and WES-measured TMB. As expected, the values for the slopes reflected the ratio of the panel size to the size of the sequence region covered by WES (supplementary Figure S2, available at *Annals of Oncology* online). Additionally, we detected intercepts ranging from 0.82 to 1.23 mutations, which were substantially different from the unbiased situation of an intercept of zero. This means that typically about one additional mutation was detected by panel-based sequencing in addition to what would be expected by multiplicative scaling, a bias that reflects the enrichment of typical panel designs for frequently mutated cancer genes. The variance of the residuals in the linear fit of psTMB versus WES-TMB showed a linear increase with the number of mutations
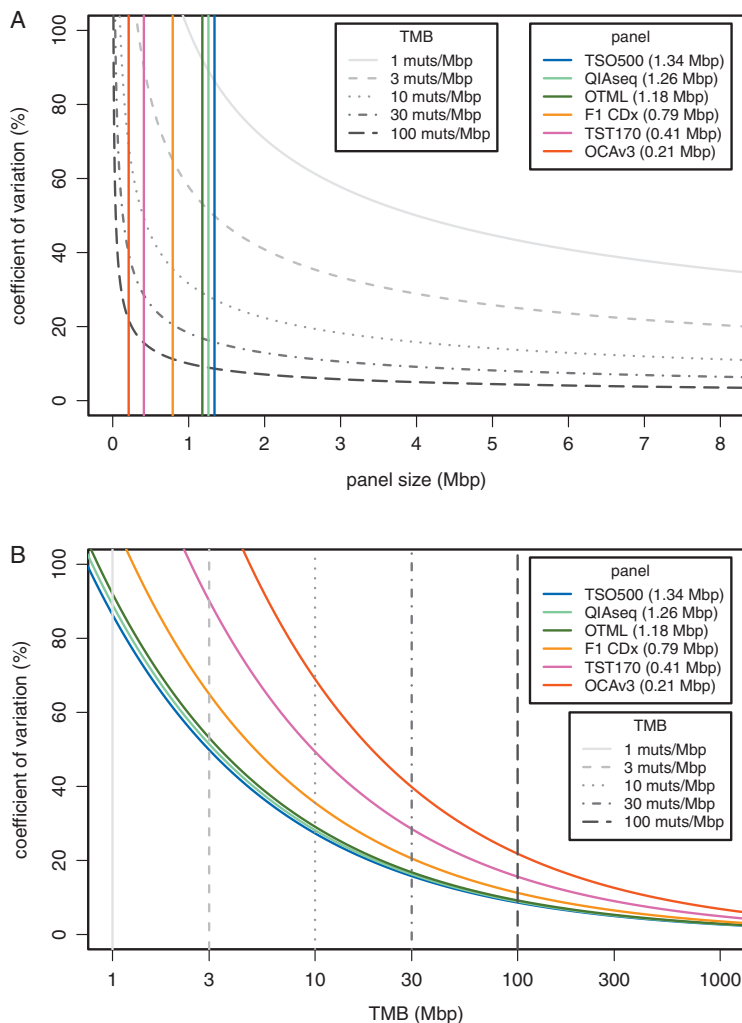
Figure 1. Analysis of the precision of tumor mutational burden (TMB) estimation by panel sequencing in a random mutation model. Visualization of the algebraic formula for the coefficient of variation (CV, in %). The CV decreases inversely proportional to the square root of the TMB level and inversely proportional to the square root of the panel size. (A) Decreasing CV with increasing panel size for tumors with TMB = 1, 3, 10, 30 and 100 muts/Mbp. (B) Decreasing CV with increasing TMB for five commercially available sequencing panels.

Table 1. Precision of TMB measurement of five commercially available sequencing panels

Panel acronym	Panel	Provider	Panel size (Mpb)	CDS covered by panel (Mbp)	CV (random mutation model) for TMB = 10 muts/Mbp (%)	Additional CV contribution from real-world cancer exomes (%)	CV (real-world cancer exomes) for TMB = 10 muts/Mbp (%)
TSO500	TruSight oncology 500	Illumina	1.95	1.34	27	28	35
QIAseq	Human tumor mutational burden panel	Qiagen	2.58	1.26	28	26	35
OTML	Oncomine tumor mutational load assay	Thermo Fisher Scientific	1.66	1.18	29	17	34
TST170	TruSight tumor 170	Illumina	0.53	0.41	49	21	60
OCAv3	Oncomine comprehensive assay v3	Thermo Fisher Scientific	0.35	0.21	69	28	88

The CV (coefficient of variation) of the panel TMB (tumor mutational burden) score was calculated in a random mutation model and estimated in simulations in the real-world cancer exomes (TCGA pan-cancer data). The numbers of mutations in the CDS (coding sequence) regions covered by the panels were compared with the total number of mutations in the CDS.
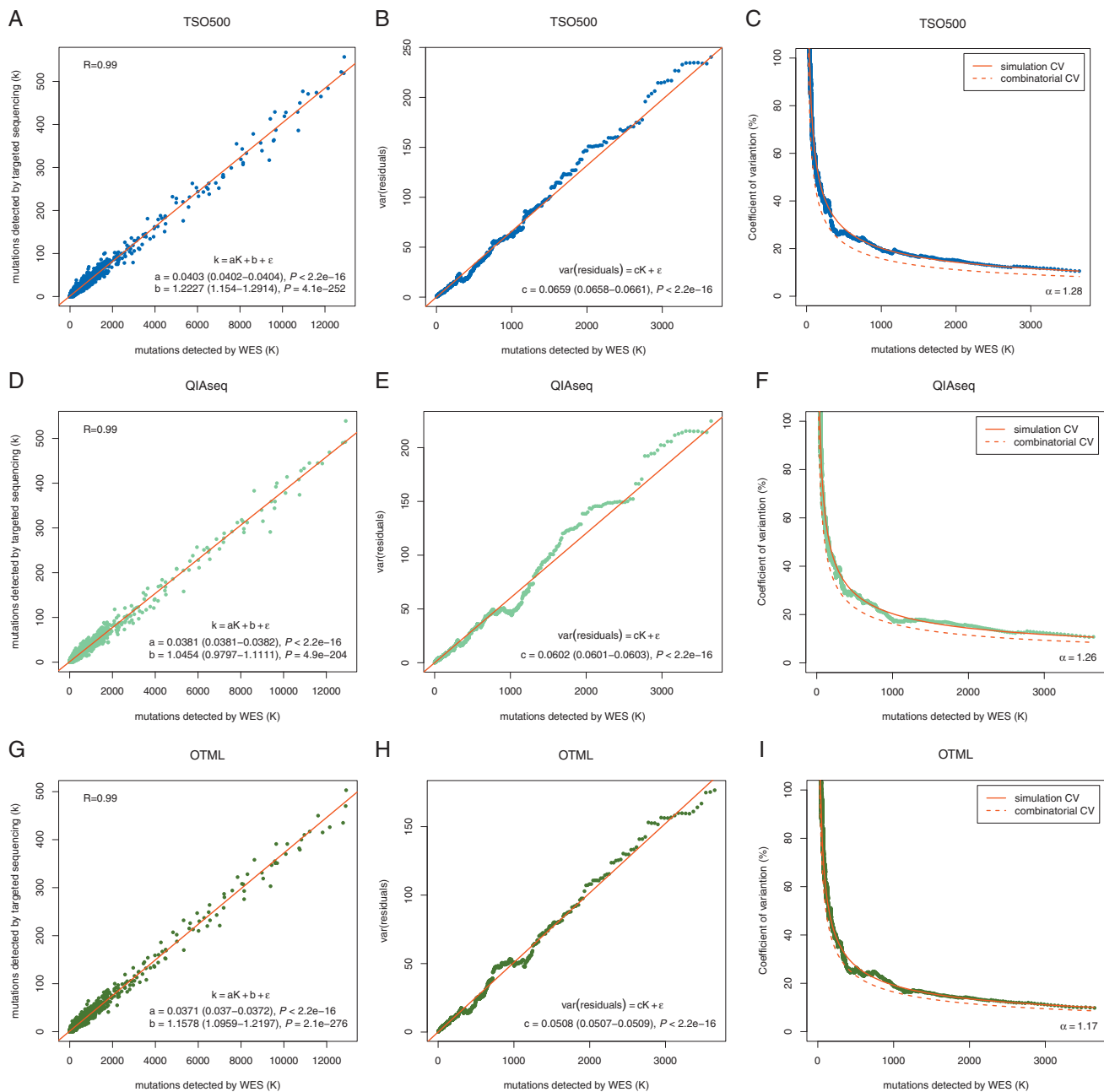
Figure 2. Analysis of the precision of tumor mutational burden (TMB) estimation by panel sequencing in real-world cancer exomes. Simulation of sequencing panels in the TCGA WES data. Only missense mutations were included in the calculation of TMB. Linear fit of the number of mutations in the panel region against the total number of mutations (A, D and G). Linear fit of the variance of the residuals against the total number of mutations (B, E and H). The variance was estimated within sliding windows of 200 tumors with similar TMB. Coefficient of variation (CV, in %) describing the real world (simulation in the TCGA data, solid line) and combinatorial (random mutation model, dashed line) imprecision in TMB estimation by targeted sequencing (C, F and I).

(Figure 2B, E and H). These data demonstrate that the mathematical law derived in the random mutation model—a decrease of the CV of the panel TMB estimate with the square root of the TMB level—remained valid for real-world cancer exomes. Of note, the CV in the real-world situation was 17%–28% higher than in the random mutation model (Figure 2C, F and I; Table 1).

To analyze the influence of selected genes on the variance of psTMB we simulated three different kinds of panels: (i) panels

exclusively composed of oncogenes and tumor suppressor genes, (ii) panels of randomly drawn genes, (iii) panels of randomly drawn genes excluding oncogenes and tumor suppressor genes. Again, the mathematical law before could be confirmed (supplementary Figure S3, available at *Annals of Oncology* online). The additional variance adding to the one predicted by the random mutation model was approximately twice as high (15%) for panels composed of oncogenes and tumor suppressor genes compared with the other panels (6%).

Intra-tumor inhomogeneity of psTMB

We compared the TMB in different regions of 100 tumors of the TRACERx lung cancer cohort [27]. TMB and simulated psTMB were calculated including the missense mutations with variant allele frequency (VAF) greater or equal to 5% (Figure 3A). Analyzing a TMB cut-point of 199 mutations, 11% of the 100 tumors were classified inconsistently by different regions. Classification with psTMB resulted in higher inconsistency rates of 18%, 16%, 17%, 15% and 24% for the TSO500, QIAseq, OTML, TST170 and OCAv3 panels, respectively; however, these differences reached significance only for one of the smaller panels: OCAv3 ($P = 0.026$). Next, we quantified the intra-tumor variation of TMB by calculating the CV between the regions of each of the 100 tumors (Figure 3B–D, supplementary Figure S4, available at *Annals of Oncology* online). The average CV of the cohort was substantially higher for all panel estimations of TMB (23.1%, 24.1%, 24%, 27.6% and 28.2% for TSO500, QIAseq, OTML, TST170 and OCAv3, respectively) compared with the WES calculation of TMB (15.3%).

Prediction of response to IO therapy

We analyzed two clinically annotated cohorts of ICB treated patients with tumors characterized by WES [13, 28].

In the Miao et al. dataset [28], we compared the capability of psTMB to separate responders (CR/PR, $n = 70$) from patients with progressive disease (PD, $n = 123$). The analysis was carried out in the subcohort of lung cancer patients ($n = 36$), melanoma patients ($n = 125$) as well as in the mixed cohort types (five cancer types, $n = 193$) for which WES data were available (Figure 4). In lung cancer, high TMB was strongly predictive of response to ICB: areas under the curve (AUC) were between 0.78 and 0.94. WES performed significantly better than the OCAv3, TST170 and QIAseq panels ($P = 0.011$, $P = 0.01$ and $P = 0.048$), but not significantly better than the OTML and the TSO500 panel ($P = 0.063$ and $P = 0.11$). Analyzing the cut-point of 199 mutations, misclassification rates compared with WES (16.7%) were considerable higher for the small panels OCAv3 (33.3%, $P = 0.087$) and TST170 (36.1%, $P = 0.054$). Odds ratios of objective response rates (OR; TMB above 199 mutations versus TMB below 199 mutations) were considerably lower for the OCAv3 and the TST170 panels (OR = 4.0 and OR = 3.1) compared with WES (OR = 27.6, 95% CI 5.2–233.1). Using a very low cut-point of 100 mutations, misclassification rates compared with WES (16.7%) were considerably higher for all five panels (OCAv3: 36.1%, $P = 0.054$; TST170: 30.6%, $P = 0.13$; OTML: 38.9%, $P = 0.033$; QIAseq: 36.1%, $P = 0.054$ and TSO500: 33.3%, $P = 0.087$). Results of the analysis in melanoma cases and the overall cohort are described in the supplementary Appendix (SR.1, available at *Annals of Oncology* online).

Analysis of the Hellmann et al. dataset of NSCLC patients treated with PD-1 plus CTLA-4 blockade [13] showed similar results (supplementary Figure S5, available at *Annals of Oncology* online).

Three-tier versus two-tier TMB classification

In the TCGA data, we analyzed the feasibility to classify tumors by panel sequencing compared with the gold standard of WES. The four 'WES thresholds' 158 muts, 199 muts (equivalent to 10 muts/Mpb [20]), 243 muts and 'median TMB' were converted to corresponding 'panel thresholds' using a linear transformation (supplementary Table S1, available at *Annals of Oncology* online). The median TMB was analyzed additionally to the clinical validated thresholds, because for many cancer types very few tumors had TMB above these thresholds (supplementary Table S2, available at *Annals of Oncology* online). For most of the cancer types, misclassification rates were substantially lower when using the three larger panels (with panel sizes >1 Mbp) compared with the two smaller panels (Figure 5A and supplementary Table S2, available at *Annals of Oncology* online); however, even with the three large panels, misclassification rates (exemplified for a threshold of 199 mutations) were substantial: 17%–19% for lung squamous cell carcinoma (LUSC), 10%–12% for lung adenocarcinoma (LUAD), 7%–11% for cutaneous melanoma (SKCM) and 5%–6% in the pan-cancer cohort. To deal with the substantial fraction of misclassified tumors, we studied a refined classification approach of replacing the cut-point by a three-tier classifier that is based on the likelihood of misclassifications (Figure 5B). The interval width was determined in such a way, that the percentage of strong misclassifications (tumor classified as 'TMBhigh' by panel sequencing, when harboring low TMB and vice versa) was <5%. The interval width was substantially smaller for the three larger panels compared with the two smaller panels and smaller in the pan-cancer cohort compared with the lung cancer and melanoma subcohorts (Figure 5C–F).

In an exploratory analysis of the 36 ICB treated lung cancer patients [28], we compared the performance of three-tier and two-tier classification using a cut-point of 199 mutations. Performance of the OCAv3, TST170, OTML, QIAseq and TSO500 panels improved from misclassifying 33.3%, 36.1%, 19.4%, 19.4% and 19.4% to misclassifying 21.1%, 19%, 14.3%, 16.7% and 14.7% of the patients (supplementary Table S3, available at *Annals of Oncology* online).

Tumor classification by psTMB and tumor indel ratio

In addition to missense mutations, we analyzed synonymous mutations, nonsense mutations and indels and calculated the ratio of the number of the latter mutations to the number of the former mutations for each of the tumors in the TCGA pan-cancer cohort (Figure 6). Indel burden in conjunction with either high or low TMB identified specific tumor types and genetic subgroups including MSI-H tumors which are known to respond well to ICB (for details see supplementary Appendix SR.2 and Figure S6, available at *Annals of Oncology* online).

Discussion

The predictive power of TMB as biomarker for response to ICB is currently investigated in many clinical trials [22, 23] across various cancer types. At first WES was widely used to determine TMB, but now there are a growing number of clinical studies which interrogate subsets of the genome by gene panels to approximate TMB (psTMB) [22, 23]. While in the clinical trial context these analyses are mainly carried out by commercial providers, many clinical laboratories depending on the regulatory
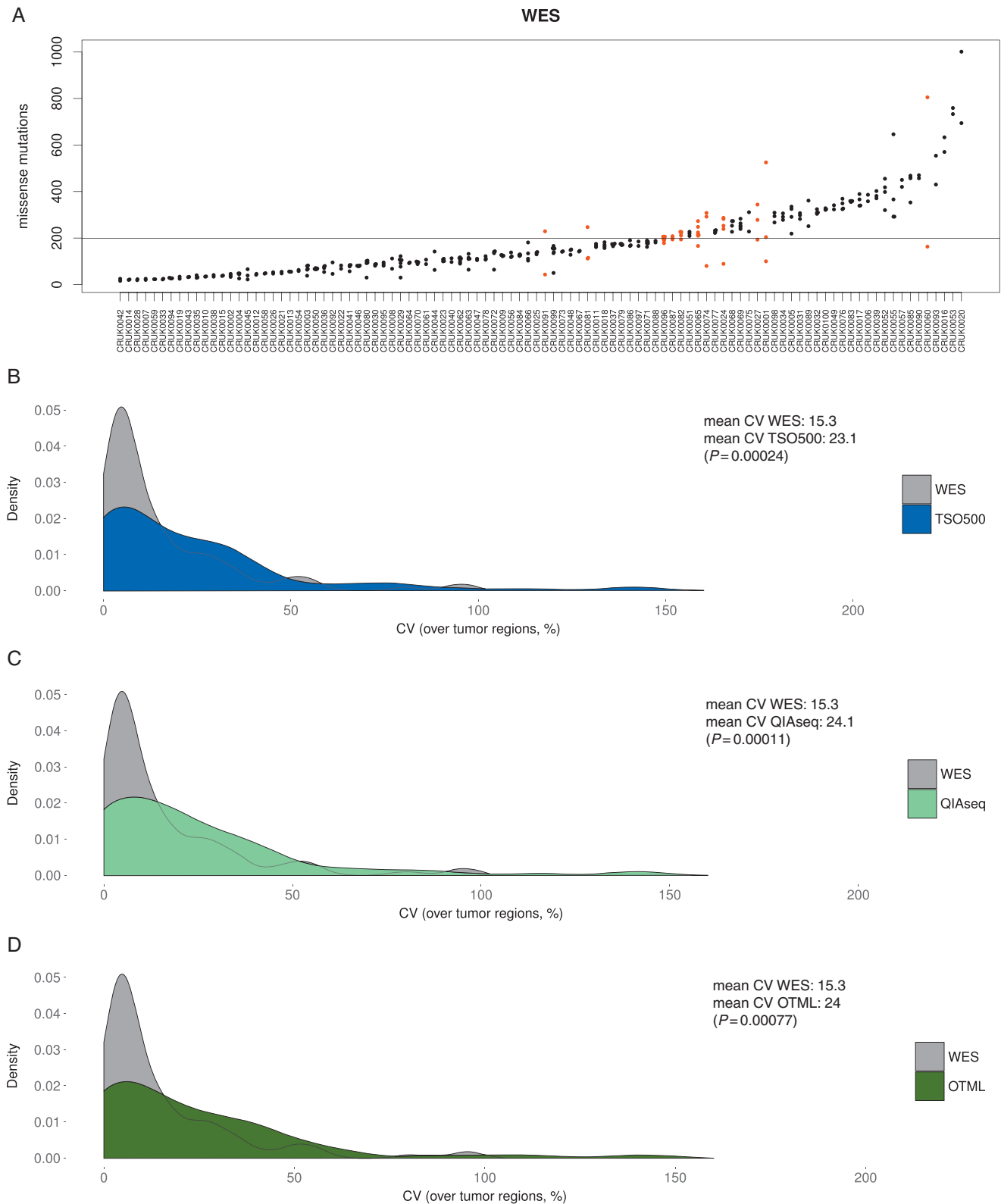
Figure 3. Inhomogeneity of tumor mutational burden (TMB) across different regions of lung tumors TRACERx 100 data [27]. (A) TMB of 323 regions of 100 lung carcinoma measured by WES and classification using a cut-point of 199 mutations. Regions of the same tumor were classified inconsistently for 11 tumors (red dots). (B–D) The coefficient of variation (CV) of TMB across regions was calculated for each of the tumors. The CV was substantially higher in simulated panel sequencing data (TSO500, QIAseq and OTML panels) compared with the WES data.
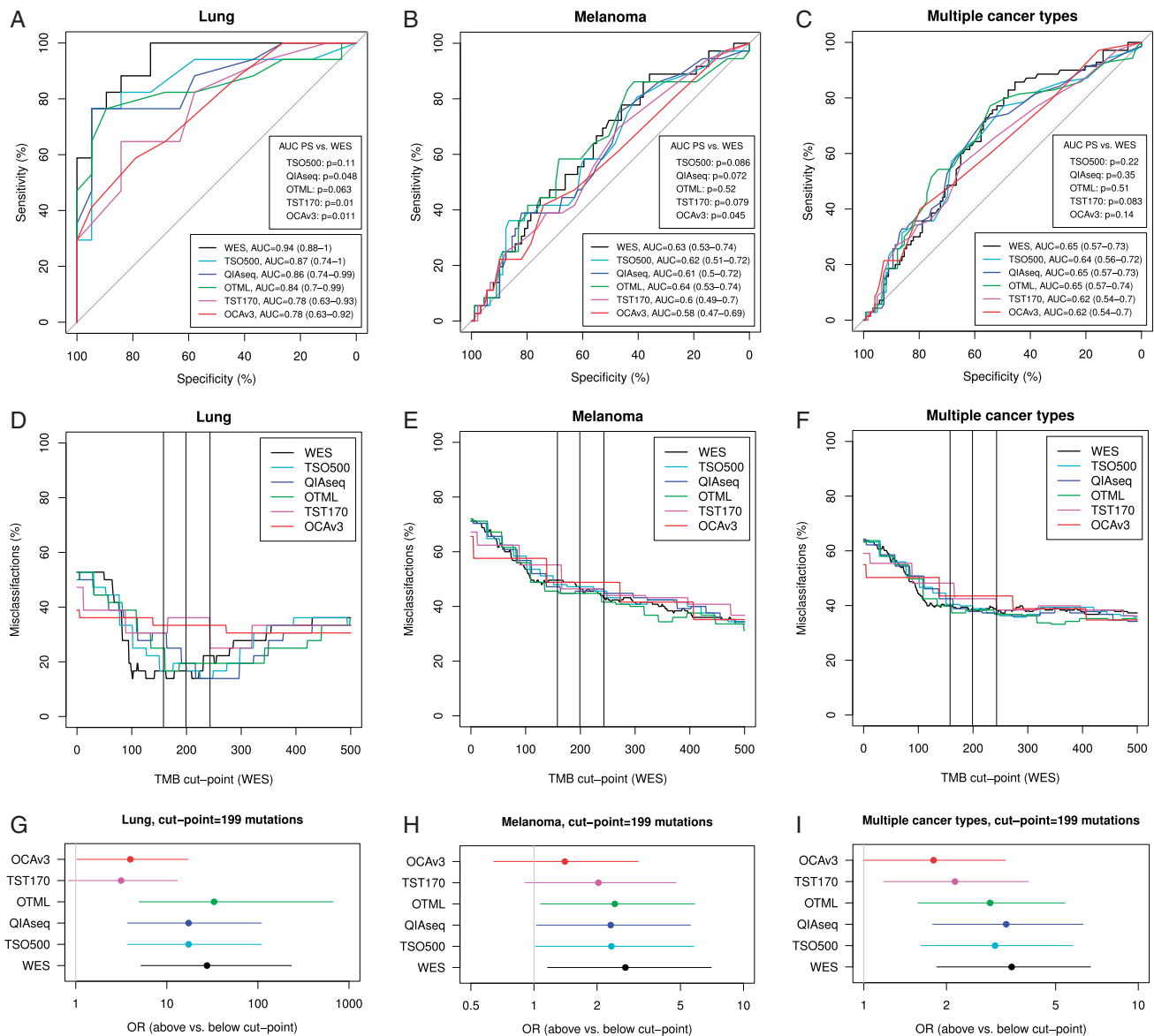
Figure 4. Prediction of response to immune checkpoint blockade (ICB) (RECIST CR/PR versus PD) by tumor mutational burden (TMB) (Miao et al. [28] data). Performance comparison of TMB measured by panels (simulations) with TMB determined by WES. (A–C) Receiver operating characteristic (ROC) curves and areas under the curve (AUC). (D–F) Misclassification rates for different TMB cut-points. (G–I) Odds ratios of response rates (OR; TMB above 199 mutations versus TMB below 199 mutations; with 95% confidence intervals).

approval context will eventually use pre- or self-designed gene panels to determine TMB scores.

TMB is a continuous variable and cut-points that discriminate between likely ICB responders and likely non-responders have to be investigated in clinical trials separately for each cancer type; however, many parameters influence TMB measurement and thus the TMB score for an individual tumor. In this work, we describe key influencing factors that are inherent to the nature of the analytical design of psTMB testing and need to be carefully considered when implementing any TMB assay in the clinic. Our analysis lays the foundation for future studies that will investigate additional parameters influencing wet-lab performance (e.g. pre-analytical factors, enrichment and sequencing technologies [25]), which is beyond the scope of this work. Several consortial approaches are under way to address these issues [29]; however,

while careful control of wet-lab parameters may minimize their influence and potential bias, the methodological limitations described in this article will remain unaffected by any of these efforts and apply to any panel.

Analyzing a stochastic mutation model, we observed that the variability of TMB counts (CV) is an algebraic function of panel size and TMB. Specifically, the CV is inversely proportional to the square root of the panel size and inversely proportional to the square root of the TMB of the tumor. Practically speaking, a scenario with low cut-points and TMB scores determined by small gene panels will result in a high imprecision of TMB measurement and will not reliably identify patients who benefit from ICB. For a tumor with a TMB close to the cut-point of 10 muts/Mb the CV turned out to be 22%, 26%, 32%, 45% and 63% for sequencing with panel sizes of 4, 2, 1, 0.5 and 0.25 Mbp. Translating this
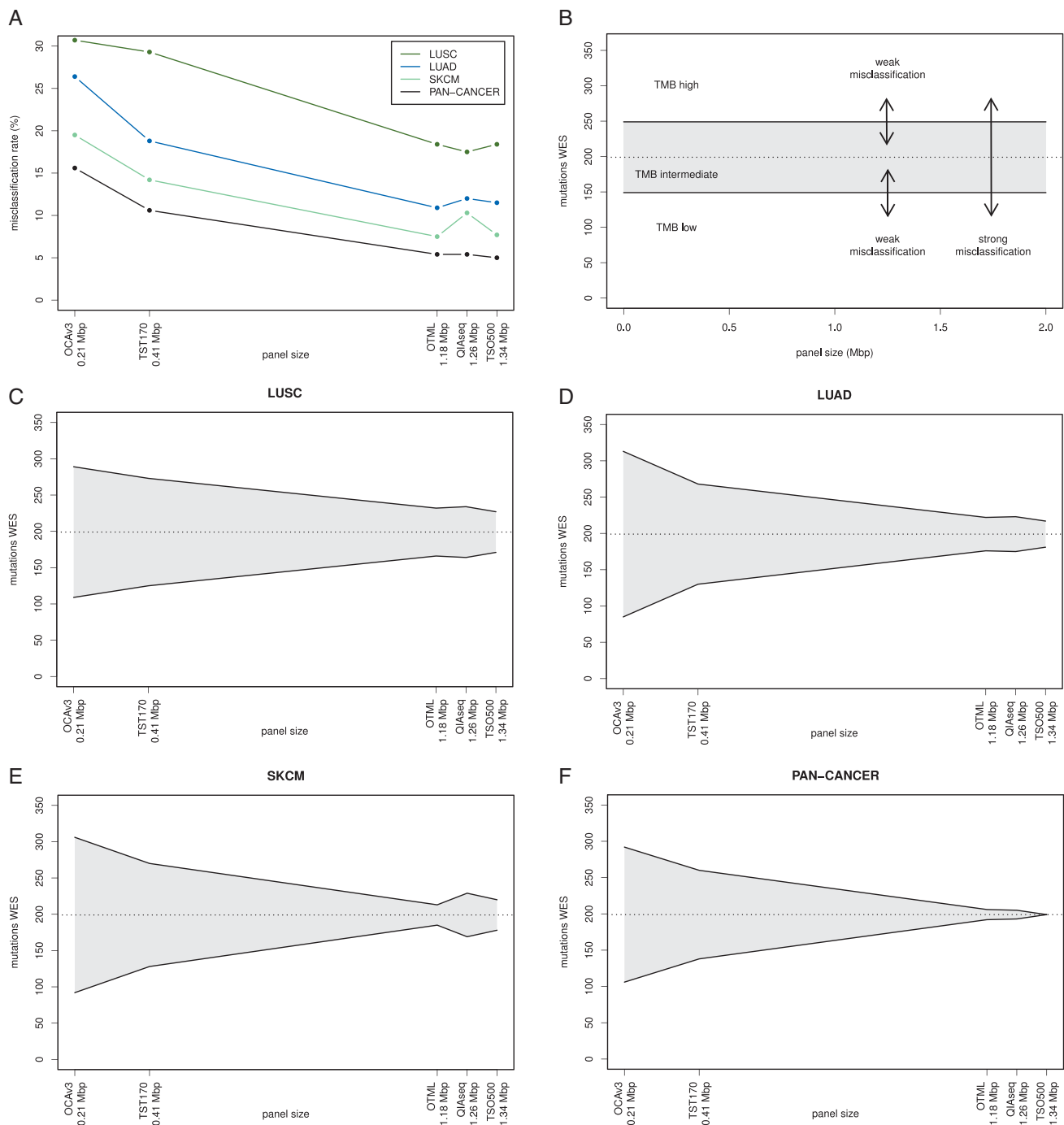
Figure 5. Analysis of the cut-point of 199 missense mutations to separate 'TMBhigh' from 'TMBlow' tumors. Simulations in the TCGA lung squamous cell carcinoma (LUSC), lung adenocarcinoma (LUAD), cutaneous melanoma (SKCM) and pan-cancer cohorts. (A) Numbers of misclassified tumors (in %) using tumor mutational burden (TMB) measurements of five commercial sequencing panels. (B) Introduction of a three-tier scheme to keep the rate of strongly misclassified tumors (classified as TMBhigh instead of as TMBlow and vice versa) under control. (C–F) The interval width was determined in such a way that strong misclassifications (classification above the gray by panel sequencing, while the tumor is classified below the gray area by WES and vice versa) occurred for <5% of the tumors.

data to a concrete real-world case, let us assume a trial enrollment scenario where a patient's tumor has a true TMB score of 15 and 10 muts/Mb is used as the cut-point separating TMBhigh from TMBlow patients. When TMB is estimated by a 2 Mbp panel the CV value mentioned above corresponds to a CI of 10.1–21.4 muts/Mbp meaning that there is a strong likelihood that this

patient will be enrolled into the TMBhigh group. Now let us consider two slightly different scenarios where a 1 or a 0.5 Mbp panel was used to estimate TMB for the very same patient: the corresponding CIs would be 8.4–24.7 muts/Mbp and 6.3–30.2 muts/Mbp, respectively. In the latter scenarios one cannot exclude that this patient although belonging to the TMBhigh group would be
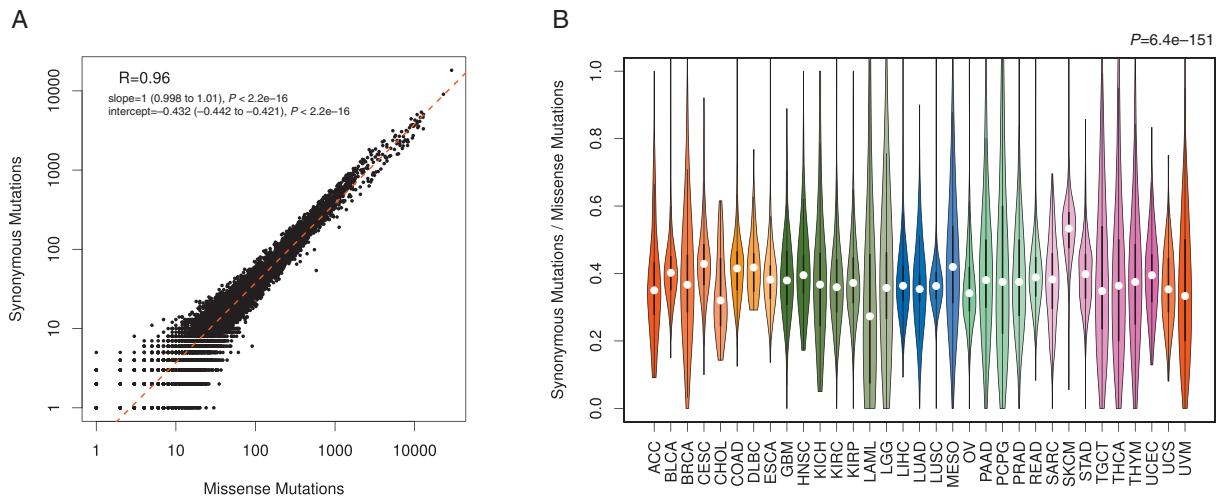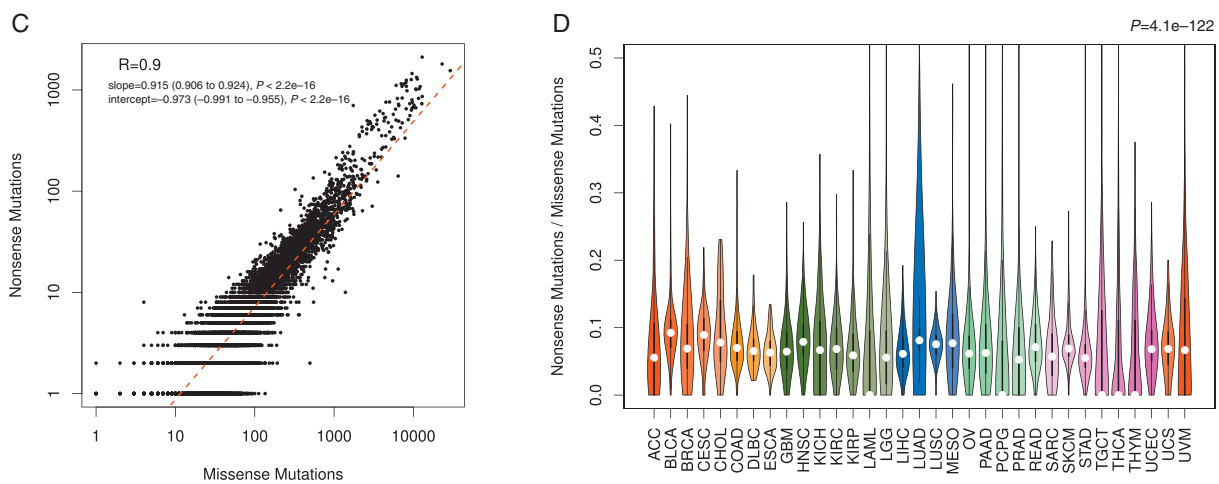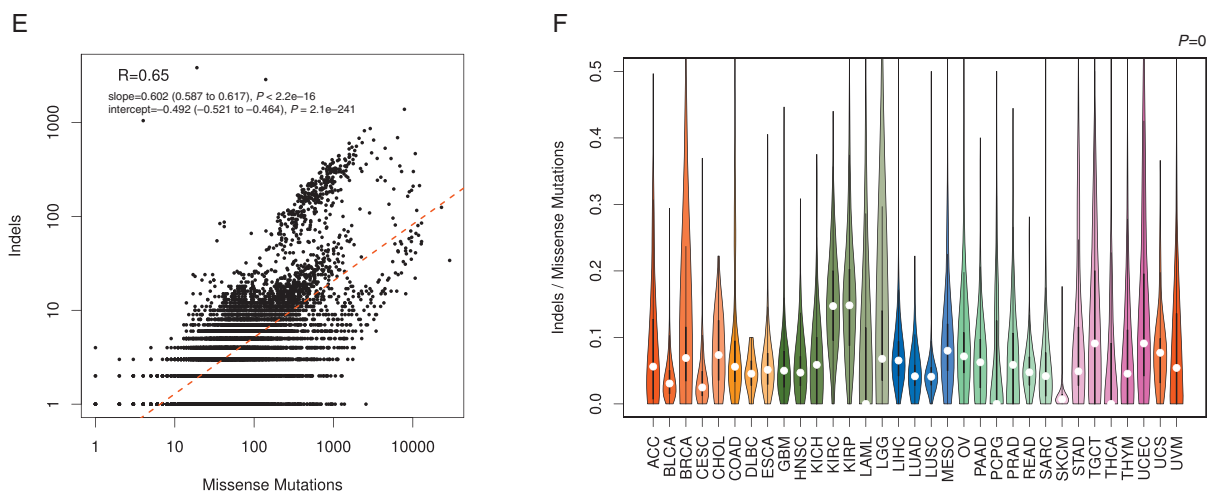
Figure 6. Analysis of other mutation types in relation to missense mutations. (A, C, E) Scatterplots showing strong correlations of synonymous mutations and nonsense mutations with missense mutations, but moderate correlation of indels with missense mutations. (B, D, F) Violin plots showing the ratio of numbers of a specific mutation type to missense mutations.

enrolled into the TMBlow group based on test results. This can evidently influence trial results but can also affect individual patient management outside clinical trials. In summary, the inherent imprecision of psTMB estimates drastically increases for panel sizes <1 Mbp. Simulation of panels in the TCGA dataset showed that this algebraic law remained valid in real-world cancer genomes with CVs (and corresponding CIs) that are even larger than in the random mutation model.

For only a few cancer types, including colorectal cancer, stomach cancer and uterine corpus endometrial carcinoma, TMB shows a bi- or multimodal distribution [30]. In these cancers, the TMB distribution is shaped by the occurrence of hypermutation in MMR deficient and/or POLE/POLD1 mutated tumors and permits a clean dichotomization; however, for most of the other cancer types including lung adenocarcinoma, lung squamous cell carcinoma and cutaneous melanoma, TMB is unimodally distributed with a dense point cloud of TMB scores scattering around the cut-point. For accurate classification of these tumors, TMB scores need to be determined by gene panels of a considerable size to obtain reliable results, i.e. with an acceptable CV. While we acknowledge that the CV will never be zero, it is evident that uncontrolled high CV values derived from insufficiently powered gene panels can misclassify individual patients. Moreover, it can lead to false assignment to TMB groups in clinical trials and thus impact clinical trial datasets as well as subsequent treatment guidelines.

Employing two datasets [13, 28], we showed that the AUC of psTMB substantially differs from WES-TMB particularly when using small gene panels [31] leading to substantial clinical misclassifications. This scenario can be controlled by establishing a three-tier classification that depends on the likelihood of misclassification instead of a single cut-point. Setting the error margin for cases which were designated TMBhigh whereas they were TMBlow in reality (and vice versa) at 5%, a subset of patients is identified where the panel itself operates with such substantial variability that clinically meaningful assignment of an individual patient to either one of the groups (low/high) is not possible—even though a cut-point has demonstrated clinical utility when analyzing a cohort of patients in a clinical trial. Introducing this concept of a 'gray zone' of TMB values would provide a safety margin in which clinicians could weigh in additional factors to their decision making (i.e. co-morbidities, other treatment options). This concept is transparent for clinicians as it conveys which TMB values to take for certain and which might be more error prone, a concept proposed for quantitative diagnostic tests before [32].

psTMB scores are also influenced by regional sampling of the tumor further contributing to the imprecision of psTMB when considering an individual case detailed above. This is a substantial finding as psTMB will be determined in biopsy material from stage IV lung cancer patients reflecting a proportion but not the entirety of the tumor, and is in line with seminal work on intratumoral genetic heterogeneity [33–35].

Both synonymous mutation burden and nonsense mutation burden turned out to be proportional to tumor (missense) mutation burden for the majority of tumors; however, in line with seminal work by Turajlic et al. [36], our work highlights indel burden as an independent parameter in the mutational spectrum of TMB. We found that the proportion of indels (in relation to missense mutations) in conjunction with either high or low TMB values

identifies different tumor types and genetic subgroups including MSI-H cases which are well known to respond to ICB [37].

In conclusion, our work shows that a universal mathematical law describes an imprecision inherent to psTMB, which is independent of pre-analytics and specific analysis parameters (e.g. coverage) and can influence patient management and clinical trial results. This scenario is further aggravated by spatial inhomogeneity of psTMB scores. Panel designs that consider the mathematical law described in this article as well as a novel three-tier classification system can control for the variable precision of psTMB. Moreover, our analysis suggests that indel burden can complement TMB results. These findings have implications for panel designs, psTMB testing and clinical decision making.

Disclosure

KL has a patent on indel burden and checkpoint inhibitor response pending, and a patent on targeting of frameshift neoantigens for personalized immunotherapy pending.

MT received advisory board honoraria from Novartis, Lilly, BMS, MSD; Roche, Celgene, Takeda, AbbVie and Boehringer, speaker's honoraria from Lilly, MSD, Takeda, research funding from AstraZeneca, BMS, Celgene, Roche and travel grants from BMS, MSD, Novartis and Boehringer.

SP, receipt of honoraria or consultation fees: Abbvie, Amgen, AstraZeneca, Bayer, Biocartis, Boehringer Ingelheim, Bristol-Myers Squibb, Clovis, Daiichi Sankyo, Debiopharm, Eli Lilly, F. Hoffmann-La Roche, Foundation Medicine, Illumina, Janssen, Merck Sharp and Dohme, Merck Serono, Merrimack, Novartis, Pharma Mar, Pfizer, Regeneron, Sanofi, Seattle Genetics and Takeda; Speakers engagements: AstraZeneca, Boehringer Ingelheim, Bristol-Myers Squibb, Eli Lilly, F. Hoffmann-La Roche, Merck Sharp and Dohme, Novartis, Pfizer, Takeda; Receipt of grants/research supports: (Sub)investigator in trials (institutional financial support for clinical trials) sponsored by Amgen, AstraZeneca, Boehringer Ingelheim, Bristol-Myers Squibb, Clovis, F. Hoffmann-La Roche, Illumina, Merck Sharp and Dohme, Merck Serono, Novartis and Pfizer.

CS, receipt of grants/research support: Pfizer, AstraZeneca, BMS, Ventana. Receipt of honoraria or consultation fees: Boehringer Ingelheim, Eli Lily, Servier, Novartis, GlaxoSmithKline, MSD, BMS, Celgene, AstraZeneca, Illumina, Sarah Canon Research Institute. Stock shareholder: Apogen Biotechnologies, Epic Bioscience, GRAIL. Co-Founder & stock options: Achilles Therapeutics.

PS received advisory board honoraria from Pfizer, Roche, Novartis and AstraZeneca as well as speaker's honoraria and research funding from Roche, AstraZeneca and Novartis.

AS is a consultant/advisory board member of AstraZeneca, Bristol-Myers Squibb, Novartis, Thermo Fisher Scientific and

References

1. Swann JB, Smyth MJ. Immune surveillance of tumors. J Clin Invest 2007; 117(5): 1137–1146.
2. Brahmer JR, Tykodi SS, Chow LQ et al. Safety and activity of anti-PD-L1 antibody in patients with advanced cancer. N Engl J Med 2012; 366(26): 2455–2465.
3. Herbst RS, Soria JC, Kowanetz M et al. Predictive correlates of response to the anti-PD-L1 antibody MPDL3280A in cancer patients. Nature 2014; 515(7528): 563–567.
4. Kerr KM, Tsao MS, Nicholson AG et al. Programmed death-ligand 1 immunohistochemistry in lung cancer: in what state is this art? J Thorac Oncol 2015; 10(7): 985–989.
5. Linnemann C, Mezzadra R, Schumacher TN. TCR repertoires of intratumoral T-cell subsets. Immunol Rev 2014; 257(1): 72–82.
6. Chowell D, Morris LGT, Grigg CM et al. Patient HLA class I genotype influences cancer response to checkpoint blockade immunotherapy. Science 2018; 359(6375): 582–587.
7. Cristescu R, Mogg R, Ayers M et al. Pan-tumor genomic biomarkers for PD-1 checkpoint blockade-based immunotherapy. Science 2018; 362(6411): 197.
8. Gubin MM, Artyomov MN, Mardis ER, Schreiber RD. Tumor neoantigens: building a framework for personalized cancer immunotherapy. J Clin Invest 2015; 125(9): 3413–3421.
9. Rizvi NA, Hellmann MD, Snyder A et al. Cancer immunology. Mutational landscape determines sensitivity to PD-1 blockade in non-small cell lung cancer. Science 2015; 348(6230): 124–128.
10. Rosenberg JE, Hoffman-Censits J, Powles T et al. Atezolizumab in patients with locally advanced and metastatic urothelial carcinoma who have progressed following treatment with platinum-based chemotherapy: a single-arm, multicentre, phase 2 trial. Lancet 2016; 387(10031): 1909–1920.
11. Snyder A, Makarov V, Merghoub T et al. Genetic basis for clinical response to CTLA-4 blockade in melanoma. N Engl J Med 2014; 371(23): 2189–2199.
12. Van Allen EM, Miao D, Schilling B et al. Genomic correlates of response to CTLA-4 blockade in metastatic melanoma. Science 2015; 350(6257): 207–211.
13. Hellmann MD, Ciuleanu TE, Pluzanski A et al. Nivolumab plus ipilimumab in lung cancer with a high tumor mutational burden. N Engl J Med 2018; 378(22): 2093–2104.
14. Ramalingam S, Hellmann M, Awad M et al. Tumor mutational burden (TMB) as a biomarker for clinical benefit from dual immune checkpoint blockade with nivolumab + ipilimumab in first-line non-small cell lung cancer: identification of TMB cutoff from CheckMate 568. In American Association for Cancer Research (AACR) Annual Meeting. Chicago, IL 2018.
15. Hellmann MD, Nathanson T, Rizvi H et al. Genomic features of response to combination immunotherapy in patients with advanced non-small-cell lung cancer. Cancer Cell 2018; 33(5): 843–852.
16. Hellmann MD, Callahan MK, Awad MM et al. Tumor mutational burden and efficacy of nivolumab monotherapy and in combination with ipilimumab in small-cell lung cancer. Cancer Cell 2018; 33(5): 853–861.
17. Carbone DP, Reck M, Paz-Ares L et al. First-line nivolumab in stage IV or recurrent non-small-cell lung cancer. N Engl J Med 2017; 376(25): 2415–2426.
18. Rizvi H, Sanchez-Vega F, La K et al. Molecular determinants of response to anti-programmed cell death (PD)-1 and anti-programmed death-

ligand 1 (PD-L1) blockade in patients with non-small-cell lung cancer profiled with targeted next-generation sequencing. J Clin Oncol 2018; 36(7): 633–641.
19. Goodman AM, Kato S, Bazhenova L et al. Tumor mutational burden as an independent predictor of response to immunotherapy in diverse cancers. Mol Cancer Ther 2017; 16(11): 2598–2608.
20. Szustakowski J, Green G, Geese W et al. Evaluation of tumor mutation burden as a biomarker for immune checkpoint inhibitor efficacy: a calibration study of whole exome sequencing with FoundationOne®. In American Association for Cancer Research (AACR) Annual Meeting. Chicago, IL 2018.
21. Garofalo A, Sholl L, Reardon B et al. The impact of tumor profiling approaches and genomic data strategies for cancer precision medicine. Genome Med 2016; 8(1): 79–89.
22. Allgauer M, Budczies J, Christopoulos P et al. Implementing tumor mutational burden (TMB) analysis in routine diagnostics—a primer for molecular pathologists and clinicians. Transl Lung Cancer Res 2018; 7: 703–715.
23. Chan TA, Yarchoan M, Jaffee E et al. Development of tumor mutation burden as an immunotherapy biomarker: utility for the oncology clinic. Ann Oncol 2018; 30(1): 44–56.
24. Buchhalter I, Rempel E, Endris V et al. Size matters: dissecting key parameters for panel-based tumor mutational burden analysis. Int J Cancer 2019; 144(4): 848–858.
25. Endris V, Buchhalter I, Allgauer M et al. Measurement of tumor mutational burden (TMB) in routine molecular diagnostics: in silico and real-life analysis of three larger gene panels. Int J Cancer 2019; 144: 2303–2312.
26. Cancer Genome Atlas Research Network, Weinstein JN, Collisson EA et al. The cancer genome atlas pan-cancer analysis project. Nat Genet 2013; 45: 1113–1120.
27. Jamal-Hanjani M, Wilson GA, McGranahan N et al. Tracking the evolution of non-small-cell lung cancer. N Engl J Med 2017; 376(22): 2109–2121.
28. Miao D, Margolis CA, Vokes NI et al. Genomic correlates of response to immune checkpoint blockade in microsatellite-stable solid tumors. Nat Genet 2018; 50(9): 1271–1281.
29. Stenzinger A, Allen JD, Maas J et al. Tumor mutational burden (TMB) standardization initiatives: recommendations for consistent TMB assessment in clinical samples to guide immunotherapy treatment decisions. Genes Chromosomes Cancer 2019; 58(8): 578–588.
30. Budczies J, Seidel A, Christopoulos P et al. Integrated analysis of the immunological and genetic status in and across cancer types: impact of mutational signatures beyond tumor mutational burden. Oncoimmunology 2018; 7(12): e1526613.
31. Lih CJ, Takebe N. Considerations of developing an NGS assay for clinical applications in precision oncology: the NCI-MATCH NGS assay experience. Curr Probl Cancer 2017; 41(3): 201–211.
32. Coste J, Pouchot J. A grey zone for quantitative diagnostic and screening tests. Int J Epidemiol 2003; 32(2): 304–313.
33. McGranahan N, Swanton C. Clonal heterogeneity and tumor evolution: past, present, and the future. Cell 2017; 168(4): 613–628.
34. Jamal-Hanjani M, Quezada SA, Larkin J, Swanton C. Translational implications of tumor heterogeneity. Clin Cancer Res 2015; 21(6): 1258–1266.
35. McGranahan N, Swanton C. Biological and therapeutic impact of intratumor heterogeneity in cancer evolution. Cancer Cell 2015; 27(1): 15–26.
36. Turajlic S, Litchfield K, Xu H et al. Insertion-and-deletion-derived tumour-specific neoantigens and the immunogenic phenotype: a pan-cancer analysis. Lancet Oncol 2017; 18(8): 1009–1021.
37. Mandal R, Samstein RM, Lee KW et al. Genetic diversity of tumors with mismatch repair deficiency influences anti-PD-1 immunotherapy response. Science 2019; 364(6439): 485–491.