Clinical Chemistry

Predictive Biomarkers for Immune Checkpoint Inhibitor Efficacy: Challenges, Innovations, and a Pathway to Precision Medicine in the Era of Cancer Immunotherapy

Article:
Matthew Lee, Jeffrey A SoRelle, Arun Everest-Dass, David E Gerber, Mitchell S von Itzstein.
Predictive Biomarkers for Immune Checkpoint Inhibitor Efficacy: Challenges, Innovations, and a Pathway to Precision Medicine in the Era of Cancer Immunotherapy.
Clin Chem 2026; 72(9): 926–47. https://doi.org/10.1093/clinchem/hvag062

Guest: Dr. Mitchell von Itzstein from the Division of Hematology and Oncology in the Harold C. Simmons Comprehensive Cancer Center at the University of Texas Southwestern Medical Center in Dallas, Texas.

Bob Barrett: This is a podcast from *Clinical Chemistry*, a production of the Association for Diagnostics & Laboratory Medicine. I'm Bob Barrett. Immune checkpoint inhibitors represent a dramatic new approach to cancer treatment, and their introduction has revolutionized oncology practice.

Despite the promise these drugs offer, there are substantial drawbacks. Very few patients experience long-term benefit, while a large number suffer from a variety of immune-related adverse events. Clearly, we need tools that can accurately predict which patients will respond favorably and which ones will not.

The FDA has approved three different tests to predict immune checkpoint inhibitor response, but all three have important limitations. A new review article in the September 2026 issue of *Clinical Chemistry* describes these existing tools and highlights where their shortcomings negatively impact patient care. It then shifts focus to the future, covering new tools under investigation that may enable us to use immune checkpoint inhibitors more intelligently and further improve care for people living with cancer.

Today, we'll speak with the article's senior author. Dr. Mitchell von Itzstein is an Assistant Professor in the Division of Hematology and Oncology in the Harold C. Simmons Comprehensive Cancer Center at the University of Texas Southwestern Medical Center in Dallas, Texas.

His research focuses on lung cancer clinical trials, predictive biomarkers for immune checkpoint inhibitors, and precision approaches to immunotherapy. So, Dr. Itzstein, could you start by outlining the three FDA-approved predictive biomarkers, PD-L1, tumor mutational burden, and MSI-H/dMMR, and explain why, despite their clinical adoption, they still remain only modest predictors of response?

Mitchell von Itzstein: Yes. The three approved biomarkers each capture a different slice of the immune system biology. So, PD-L1 is a cell

surface ligand measured by immunohistochemistry. It measures expression of the target for the PD-L1 immune checkpoint inhibitor drugs. So, tumors with high expression are thought to be more dependent on that pathway for immune evasion.

Tumor mutational burden, or TMB, counts somatic mutations per megabase of coding DNA, and it serves as a surrogate for neoantigen load. So, in other words, the more mutations, the greater the chance the tumor is producing novel peptides that the immune system can recognize as foreign and attack.

And then, MSI-H, or mismatch repair deficiency, identifies tumors with defective DNA mismatch repair, which are especially rich in mutations and therefore in neoantigens. And I call them modest predictors when I compare them to other biomarkers we use to guide cancer treatment.

So, for example, if I find an EGFR, an ALK, or a ROS1 alteration in a lung cancer patient, the response rate to the matched targeted treatment can be as high as 90%. If I find PD-L1 expression of greater than 50%, which is our best standard immune therapy biomarker in lung cancer, then less than a half of those patients respond to immune therapy. And also, a small but real proportion of PD-L1 negative patients still do get durable benefit from immunotherapy.

So, a negative test doesn't let me withhold the therapy either. And PD-L1 also doesn't reliably predict benefit in melanoma, kidney cancer, or liver cancer. So, we typically don't even test it there. And with TMB, it's currently approved as a one-size-fits-all cutoff that works in some tumors, but it doesn't work in others. And with MSI-H, it is a genuinely excellent biomarker when it's positive, but it's simply very rare. It's less than 1% of all lung cancers, although it is seen more frequently in some tumor types like colorectal, gastric, and uterine, where it does produce excellent responses when it's identified.

But really, the deeper reason why they all underperform is that cancer immunotherapy response is governed by multiple interacting systems. So, the tumor's genomics, the tumor microenvironment, antigen presentation machinery, the patient's own systemic immunity, and even the gut microbiome. So, no single assay captures more than one of these dimensions.

Bob Barrett: Now, PD-L1 is the most widely used biomarker guiding immunotherapy, but your review describes four different FDA-approved assays, three different scoring systems, and meaningful discordance between them all. Is PD-L1 a good biomarker being used badly or a weak biomarker that we've just become accustomed to?

Mitchell von Itzstein: Well, I do think it's a bit of both. And let me give PD-L1 its credit first. It is the most widely used immunotherapy biomarker and one real strength is the tumor-specific scoring systems. So, we use tumor proportion score in lung cancer, we use combined positive score in headed neck and gastric cancers, and immune cell score in bladder and breast.

And each of these are validated prospectively in their own registrational trials. So, the level of evidence is strong, but the test has practical problems. There are four FDA-approved assays, each with different antibody clone, each locked to its own manufacturer's staining platform, and each tied to specific immunotherapy drugs. And when we looked at the blueprint studies, it showed that three of them, 22C3, 28-8, and SP263, are reasonably concordant, but SP142 systematically stained fewer tumor cells. So, in other words, the same tumor can be positive on one assay and negative on the other.

There's also the biological issues of spatial and temporal heterogeneity. So, for example, a small biopsy from one site of a tumor at one time point may simply not represent the overall cancer biology in that patient. PD-L1 expression is also a biologically continuous variable that we forced into cutoffs for practical clinical use.

A lung cancer with a PD-L1 above 90% responds better than one at 50%, but both get reported as PD-L1 high. So, for the reporting laboratories, the practical advice is to use the companion diagnostic validated for that intended cancer and drug and state the assay, the antibody clone, and the scoring system explicitly, as well as the actual PD-L1 percentage on the report.

Bob Barrett: So, tumor mutational burden received a tissue-agnostic approval at a threshold of 10 mutations per megabase. Your review argues that a single universal cutoff might not be the right approach. Why is that? Explain the problem.

Mitchell von Itzstein: Well, the cutoff and tumor-agnostic approval came from a basket study across 27 different cancer types, where collectively patients with tumors above 10 mutations per megabase, measured by NGS from a tumor biopsy tissue, had a response rate of 29% versus 6% for those below 10.

The problem is that mutational burden has different distribution for every cancer type. So, 10 mutations per megabase is pretty unremarkable for a smoking-related lung cancer, but it's extremely high for a prostate cancer. And when other groups have done pan-cancer analyses of more than 70,000 patients, the TMB performed best when the thresholds were set as percentiles for within each cancer

type, with the most predictive cutoff sitting near the top of each cancer's own distribution.

There are also technical measurement issues. So, the specific panel size, the gene content, variant inclusion rules, and germline filtering chosen for each assay all affect the number reported. When we do germline filtering across reference databases, it can produce falsely elevated TMB in patients from ancestries that are underrepresented in those databases, and this has been documented in published real-world studies. The TMB can also be artificially inflated by tumor subclonal mutations that don't actually generate effective neoantigens, and low tumor content in the biopsy tissue can deflate the TMB.

So, while the greater than 10 mutations per megabase threshold opened important tumor agnostic indication, it oversimplifies the biology. And tumor-specific or percentile-based cutoffs and better assay harmonization and integration with other markers, such as neoantigen clonality, would make TMB a more precise biomarker.

Bob Barrett: Doctor, your review surveys a broad landscape of investigational biomarkers. Which of these areas do you consider closest to near-term clinical implementation, and what evidence supports that view?

Mitchell von Itzstein: Well, there's two that stand out to me. The first is actually blood-based TMB and circulating tumor DNA more broadly. The practical case for this is pretty compelling because for lung cancer, up to 30% of biopsies are inadequate for sequencing, so a blood test may be the only option for a patient.

And since it's only a blood test, it's readily repeatable, which is not practical for a tumor tissue and repeat biopsies. The evidence base is also the most mature of any investigational marker, with multiple prospective trials showing immunotherapy benefit in high-blood TMB patients. Longitudinal ctDNA can also capture early treatment effect more dynamically than a one-time tissue biopsy.

For example, greater reductions in post-treatment ctDNA have shown prognostic benefit. But commercial assays are not interchangeable, and the specific TMB cutoff thresholds for benefit vary substantially between vendors. And then there are clonal hematopoiesis issues that can contaminate the signal, but these should be solvable problems with the appropriate bioinformatics.

And then secondly, *POLE* and *POLD1* mutations. These are proofreading polymerase mutations, and they produce ultra-high mutational burdens in cancers and are very strongly

associated with immunotherapy benefit across different tumor types. Very importantly for these two mutations, they're already being sequenced on standard commercial NGS panels, and so most cancer patients already have the result of these mutations. No new assay is required. Just recognition of the mutation when it does appear, and appropriate reporting on the molecular pathology reports.

Bob Barrett: Now, you conclude that composite integrative models, potentially artificial intelligence-driven, are the most promising path forward. Some listeners will hear that as trading a flawed but interpretable test for a black box. How would you answer that concern?

Mitchell von Itzstein: Well, let me first say I don't think that the current standard testing is going away anytime soon. I anticipate continuing to utilize PD-L1, TMB, and MSI-H for my patients for the foreseeable future. But composites are the right direction because it follows directly from why single markers underperform. If immunotherapy responses depend on all of these different factors like tumor genomics, antigen presentation, the microenvironment, systemic immunity, the microbiome, simultaneously, then no single assay can ever capture all of this complexity.

And we already have the proof of concept that combining multiple tests works. So when we've studied HLA heterozygosity plus TMB, it predicts better than TMB alone. When we studied the tumor microenvironment classification plus TMB, it beats TMB alone. So these examples are not black box per se because they are transparent combinations of different types of tests integrated together. And where the artificial intelligence comes into play is that it can act as the integrator for all of these different axes together.

Because at some point, when we're measuring all of these different variables, it can exceed what is possible through standard statistical regression where a human being can interpret. So using composite tests with AI to assist in the analysis can be helpful. Now, of course, with anything like this that's being newly developed, prospective validation and transparent reporting of how the model arrives at its prediction will still be essential.

And we should aim for models that are both more accurate, but still clinically interpretable, not a black box system.

Bob Barrett: Well, finally, Dr. von Itzstein, let's look ahead for a laboratory director listening today. What should they be preparing for over the next three to five years?

Mitchell von Itzstein: Well, first, I think we should expect for increasing volumes of non-tissue specimens. So particularly plasma for blood TMB

and ctDNA and then eventually stool for microbiome analyses. So laboratories will need robust processes for handling these types of specimens and clear logistics for send-out testing.

And then second, most of the advanced investigational assays, for example, the multiplex spatial profiling, advanced proteomics from plasma, they will remain specialized. And the smart approach for the average laboratory is to establish formal partnerships with either academic or commercial reference labs that have already validated and scaled pipelines for these types of specialized tests. And then the local lab is responsible for attaining the specimen processing and handling at the local site, similar to what is already done for the send-out NGS tests.

And for optimal patient care, we need to aim for a clinically actionable result within a few weeks from when the testing is ordered. And then third, continue participation in the harmonization efforts for the assays that are already being run. So PD-L1, IHC, and TMB for example, there are nuances and complexities that we've discussed in how these assays are run and reported, but we need to continue the efforts across the country and even across the world to make sure there's harmonization and interoperability between the results so that clinicians can understand and make the right treatment decisions for patients.

The labs that position themselves as the interface between this complex, coming multi-omics data and the treating oncologist will be best placed to support the next generation of precision cancer immune therapy.

Bob Barrett: That was Dr. Mitchell von Itzstein from UT Southwestern Medical Center in Dallas, Texas. He wrote a review article in the September 2026 issue of *Clinical Chemistry* summarizing biomarkers to predict immune checkpoint inhibitor response, and he's been our guest in this podcast on that topic. I'm Bob Barrett. Thanks for listening.