BioCentury
REPRINT FROM MARCH 20, 2017

TOOLS & TECHNIQUES

CALIBRATING PD-L1

BY STEPHEN HANSEN, ASSOCIATE EDITOR

Flagship Biosciences Inc.'s Computational Tissue Analysis platform won't provide a common cutoff point across the five major assays used to measure PD-L1 expression. But it could calibrate the output so that results from one assay can be used to predict the scores the other assays would produce for the same sample.

That would allow oncologists to make decisions about using any marketed PD-1 or PD-L1 therapy without necessarily having to order the corresponding assay.

At the American Association for Cancer Research (AACR) meeting in April, Flagship will present data from a pilot study that demonstrates the approach is feasible.

There are three FDA-approved PD-L1 assays. Each uses a different anti-PD-L1 antibody clone to stain biopsy tissue, and each was developed for use with a specific therapy: PD-L1 IHC 22C3 pharmDx assay was developed for Keytruda pembrolizumab from Merck & Co. Inc.; PD-L1 IHC 28-8 pharmDx assay was developed for Bristol-Myers Squibb Co.'s Opdivo nivolumab; and PD-L1 (SP142) assay was developed for Tecentriq atezolizumab from the Genentech Inc. unit of Roche.

Two other tests are still investigational: PD-L1 (SP263) assay, which is being developed for use with AstraZeneca plc's durvalumab, and the PD-L1 E1L3N XP assay from Cell Signaling Technology Inc., which is a laboratory-developed test (LDT) currently used for research.

The pharmDx assays are from the Dako A/S unit of Agilent Technologies Inc. The SP142 and SP263 assays are from Roche's Ventana Medical Systems Inc. unit.

The challenge for pathologists and oncologists is that because each assay was developed based on clinical data for a specific therapeutic, the reagents, scoring algorithms and resulting interpretations are different. As a result, each diagnostic can be used only with its corresponding therapy (see "Differential Staining").

These differences also have resulted in each assay having different expression cutoff points for determining whether a patient is PD-L1-positive or -negative for a given indication. For example, in second-line urothelial carcinoma, the cutoff for PD-L1-positivity for patients treated with Tecentriq was ≥5% PD-L1 expression in immune cells using the PD-L1 (SP142) assay. In the Phase II CHECKMATE -275 trial that supported Opdivo's approval in the same indication, the cutoff for PD-L1 positivity was ≥1% expression in tumor cells.

According to Flagship CSO Joseph Krueger, the variability in the cutoffs used for the assays is largely due to the limitations of human pathologists who conduct the tests.

He said there may be 1 million or more cells on a slide that a pathologist must look at through a microscope. He said while all pathologists should be able to differentiate between a PD-L1-negative sample and a sample with 50% PD-L1 expression, determining smaller differences in expression levels among a million cells is very difficult.

"Because this is very hard, they've had to really simplify the scoring paradigms on these diagnostics to match the level of human ability," Krueger said. "A pathologist can't determine if someone expresses 12.3% PD-L1. In many cases, the pathologist cannot tell the difference between 10% and 20%."

This lack of precision may be one of the reasons why in clinical trials some patients who fall below the cutoff for PD-L1-positivity still respond to therapy.

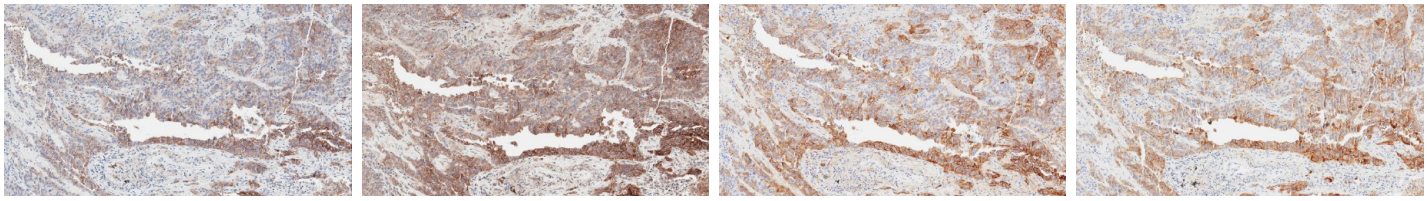
Flagship's solution is to remove human subjectivity from the equation using its Computational Tissue Analysis (cTA) software.

The process starts by using a standard slide scanner to take a high resolution image of the whole slide. The image is then uploaded to the cTA software for analysis.

Krueger said the software is programmed with the parameters for the staining intensity and scoring algorithms specific to each assay. The computer can measure each cell, determine if it is a tumor or immune cell, and then score

DIFFERENTIAL STAINING

The images show the differences in staining intensity for four different PD-L1 assays on the same biopsy sample from a non-small cell lung cancer (NSCLC) patient. **Merck & Co. Inc.** (NYSE:MRK) uses the PD-L1 IHC 22C3 pharmDx assay with Keytruda pembrolizumab. **Bristol-Myers Squibb Co.** (NYSE:BMY) uses the PD-L1 IHC 28-8 pharmDx assay with Opdivo nivolumab. Genentech Inc. uses the PD-L1 (SP142) assay for Tecentriq atezolizumab. **Cell Signaling Technology Inc.** markets the PD-L1 E1L3N XP assay for research use. The pharmDx assays are from the Dako A/S unit of **Agilent Technologies Inc.** (NYSE:A), while the SP142 assay is from the Ventana Medical Systems unit of **Roche** (SIX:ROG; OTCQX:RHHBY).



Merck | Bristol-Myers | Genentech | Cell Signaling

it for PD-L1-positivity based on the scoring algorithm required for the given assay and indication.

The output is a positive or negative score — as an oncologist would receive for the available assays — but with additional information on the precise PD-L1 expression levels in tumor and immune cells.

COMPUTATIONAL BENEFITS

Krueger said the improved precision of the cTA platform could provide two main benefits: calibration of results across the various assays, and optimization of cutoff points for patient selection.

He said it is already burdensome for pathologists to have to use three different assays on two different IHC platforms.

At least two studies — the Blueprint Project and a second prospective study funded by Bristol-Myers and the National Comprehensive Cancer Network (NCCN) Oncology Research Program — have demonstrated analytical similarity across three of the assays, meaning each assay showed a similar percentage of tumor cells stained compared with the other assays.

However, in each study the SP142 assay was analytically different.

With more than 600 immuno-oncology clinical trials ongoing that use a tissue-based biomarker, Krueger estimated that there could be 20 or more new immuno-oncology companion diagnostics over the next 10 years, each with different reagents and scoring algorithms for multiple indications.

Conducting prospective studies to determine analytical similarity of all the assays could be extremely time consuming and expensive. He said industry could agree to only use one assay for each biomarker, but he considered such collaboration highly unlikely.

The cTA platform could allow a pathologist to use his or her assay of choice to get a predicted score for the same biopsy sample on another assay.

For example, the cTA software could analyze a slide image stained using the 22C3 assay, and the pathologist could select for a readout based on the SP142 assay. The software would then predict how that same biopsy sample would be scored for PD-L1 expression if the SP142 assay parameters were applied.

"Because I understand all the mechanical differences between these assays, I can make these correlations between the tissue analysis," Krueger said.

"I could run one of these assays that's out there for any one indication, and get the result for all the other assays across all the other indications just by running that one slide and assay," he added.

A pilot study of 20 NSCLC patients compared biopsies from each patient stained using the 28-8, 22C3, SP142 and E1L3N assays. Data won't be available until the AACR presentation on April 2.

Krueger said the next steps for the cTA platform for PD-L1 harmonization include larger analytical validation studies to demonstrate the software can correctly predict the score of other assays using a single assay, followed by a clinical validation study that shows the platform can correctly predict scores that correlate with patient outcomes.

While he declined to disclose a timeline for when the next trials would start, Krueger did say Flagship recently met with FDA to discuss a strategy for studies that would be acceptable to the agency for using cTA to measure PD-L1 expression.

OPTIMIZING CUTOFFS

Because it enables the detection of smaller differences in the amount of PD-L1 expression than can be detected by the human eye, the cTA platform can also optimize the cutoff point for an assay, either during clinical development or retrospectively once a drug is on the market.

"We could set up an entirely new digital scoring paradigm that better selects patients than the original companion diagnostic did, and potentially resolve some of these problems with variability between pathologists," Krueger said.

For example, cTA could be used in a retrospective bridging study on the same samples that were used to determine an assay's original cut-off. By correlating the more precise PD-L1 expression level generated by cTA with individual patient outcomes, a new cutoff point could be found that better differentiates responders from non-responders.

Partnering with companies to use the cTA platform for quantitative analysis of biomarkers in clinical development is part of Flagship's business model.

Krueger said Flagship has worked with more than 100 industry customers to support biomarker analysis in clinical development on a fee-for-service basis. Flagship also works with *in vitro* diagnostic (IVD) manufacturers to integrate the cTA system into their devices.

Krueger said Flagship eventually hopes to earn product revenue, milestones and or royalties by providing the cTA software as part of an approved diagnostic.

He said the potential to more accurately predict which patients will respond to therapy should allow diagnostics using cTA to command a better price, and reimbursement, than existing companion diagnostics that require manual assessment by a pathologist. bc

COMPANIES AND INSTITUTIONS MENTIONED

Agilent Technologies Inc. (NYSE:A), Santa Clara, Calif.
American Association for Cancer Research (AACR), Philadelphia, Pa.
AstraZeneca plc (LSE:AZN; NYSE:AZN), London, U.K.
Bristol-Myers Squibb Co. (NYSE:BMY), New York, N.Y.
Cell Signaling Technology Inc., Danvers, Mass.
Flagship Biosciences Inc., Westminster, Colo.
Genentech Inc., South San Francisco, Calif.
Merck & Co. Inc. (NYSE:MRK), Kenilworth, N.J.
Roche (SIX:ROG; OTCQX:RHHBY), Basel, Switzerland
U.S. Food and Drug Administration (FDA), Silver Spring, Md.

REFERENCES

Hansen, S. "The devil is in the assay." *BioCentury* (2016)

Rimm, D., et. al. "A prospective, multi-institutional, pathologist-based assessment of 4 immunohistochemistry assays for PD-L1 expression in non-small cell lung cancer." *JAMA Oncology* (2017)

BIOCENTURY INC.

NEWSROOM
pressreleases@biocentury.com

SAN CARLOS, CA
+1 650-595-5333; Fax: +1 650-595-5589

CHICAGO
+1 312-755-0798; Fax: +1 650-595-5589

WASHINGTON, DC
+1 202-462-9582; Fax: +1 202-667-2922

UNITED KINGDOM
+44 (0)1865-512184; Fax: +1 650-595-5589