

Comment to the FDA in follow up to Public Workshop - Complexities in Personalized Medicine: Harmonizing Companion Diagnostics Across a Class of Targeted Therapies, March 24, 2015

Written by David L. Rimm

Comments endorsed by a group of Pathologists attending the workshop, including:

Ken Bloom MD – GE / Clariant Labs

Elizabeth Hammond – University of Utah

Debra Leonard – University of Vermont

Given that the goal is (quoted from the meeting) to “Deliver a data / information package to inform the medical practice community on PD-L1 IHC testing”; the following are key concerns that should be considered in the design of the proposed *Industry Working Group Proposal on Comparing the Analytical Performance of the PDL-1 tests*.

**1. The proposal does not include outcome information or address clinical utility.** In the best case scenario, comparison of two platforms with 4 antibodies would show a high degree of concordance suggesting any assay would be suitable. This outcome is extremely unlikely. It is much more likely, based on the published literature to date, that each assay will show substantially different characteristics, but without any mechanism to determine which is biologically correct or which is clinically useful. We have already done a quantitative assessment of PD-L1 expression comparing 2 antibodies on a single platform on serial sections from 49 cases and shown differences between E1L3N from Cell Signaling and SP142 from Spring Biosciences (see figure 1). Unfortunately, since we do not have response data, we cannot tell which is more clinically useful. A similar study conducted by the consortium is likely to obtain similar results that will thus not address the issue of clinical utility and thus not help resolve which assay should be selected by pathology labs that wish to run companion diagnostic tests. If the medical practice community needs to know which test to use, their decision should be driven by outcome data. Since all parties are holding tissues with outcome information, it is a flaw in this proposal design that none is included.

**2. The proposal does not include other common platforms beyond Ventana and Dako.** For the detection of BRAF, KRAS and EGFR mutations, the FDA cleared device on the label of vemurafenib and erlotinib is only used in a relatively small minority of molecular diagnostic labs surveyed by the College of American Pathologists (CAP). Almost certainly, we will see similar use of lab derived tests (LDTs) for assessment of PD-L1. Thus it would represent a significant oversight to not include an arm, in any comparison study, that used a laboratory validated platform that might represent what could be used as an LDT (for example the antibody E1L3N on a Leica Bond platform). There is no doubt that the vast majority of the medical practice community will not purchase a “system” for each companion diagnostic test.

**3. The proposal does not include measurement of PD-L1, but rather simply reading of the assay.** At a time we are trying to increase the level of evidence and the level of the science in medicine and pathology, it is surprising that 30 year old methods of assessment of protein expression in tissue are exclusively proposed. While it is true that existing companion diagnostic tests are based on traditional immunohistochemistry(IHC), the recent failure of IHC in the EGFR setting and most recently in the MetMab trial should be a warning beacon that this technology is obsolete. While it is fine to include conventional methods, it is a significant omission to not include quantitative IHC, quantitative fluorescence or mass spectroscopy methods in at least one arm of this study.

**4. The proposal does not include an open system with defined assay conditions.** It is surprising that the proposal includes a closed-box system with proprietary reagents rather than an open system with defined, specified, chemicals and biological reagents. Since the medical practice community is likely to have to choose a set of optimal conditions and define an LDT that works for their lab, it is critical that all conditions and reagents be completely and scientifically disclosed. While this may be a bad business strategy for large companies intent on maximizing their market share it is bad science and potentially bad for patients.

**5. The proposal does not include standards to prove the assays are measuring PD-L1.** The current proposal would compare assays only on tissue sections. Studies have shown that many antibodies to PD-L1 cross react with other unknown proteins [1]. While we assume that the antibodies produced by each company (28.8, 22c3, SP142 and SP263) are validated according to current best practices[2, 3], there are no peer reviewed public data supporting that assumption. As such, if they are to be used in a comparison study, it is critical to validate the antibodies and then to use validated standards (possibly including cell line arrays or other methods) with each preparation batch (each run of the autostainer) to show antibody reproducibility.

**6. The proposal does not include biostatistical considerations for determination of study size and statistical power.** This is easily remedied.

#### References:

1. Velcheti, V., et al., *Programmed death ligand-1 expression in non-small cell lung cancer*. Lab Invest, 2014. **94**(1): p. 107-16.
2. Fitzgibbons, P.L., et al., *Principles of Analytic Validation of Immunohistochemical Assays: Guideline From the College of American Pathologists Pathology and Laboratory Quality Center*. Arch Pathol Lab Med, 2014.
3. Bordeaux, J., et al., *Antibody validation*. Biotechniques, 2010. **48**(3): p. 197-209.

Figure 1: PD-L1 protein distribution using quantitative immunofluorescence (QIF) from 49 individual cases of NSCLC. Score distributions are shown for each case where each field of view(FOV) is a blue dot and with mean score shown as the black bar for each individual tumor assessed with E1L3N or SP142. QIF scores measured in arbitrary units (AU). Inset of H&E shows an example whole tissue section with the portion analyzed by QIF shown in black box. Note that many cases with some fields showing expression of PD-L1 with one antibody show no FOVs with expression when assessed by the second antibody. See figure on next page:

