

## Supplemental Text: Immunohistochemistry proficiency testing and central review

### Methods

For PTEN IHC, interlaboratory scoring proficiency was carried out using two controls and ten patient samples over two rounds, with interim evaluation of agreement and intervening feedback to five participating pathologists. Following staining and local review, slides were scanned to a minimum of 200X magnification for central review. Two thoracic pathology experts performed independent central review of approximately 370 PTEN-stained scanned images each. The two reviewers scored both a training set of 10 cases prior to review to establish scoring parameters and a validation cohort of 30 cases at the end of the review to confirm consistency.

For MET, pathologists from 7 participating institutions participated in training and 3 rounds of proficiency testing using scanned images MET expression according to H-score (Teresa Boyle, personal communication).

### Results

PTEN immunohistochemistry proficiency testing involving five site pathologists demonstrated concordance (PTEN expression score as intact versus lost) for 21/25 results (84%); removing results from one outlier improved concordance to

19/20 (95%). Concordance between the central reviewers was 90% in a training set (9 of 10 cases) and 91% in a validation set (27 of 30 cases).

PTEN immunohistochemistry results were reported by the participating sites for 840 patients: 615 (73%) were performed at individual sites and 222 (27%) were performed and interpreted at a central testing site (University of Colorado). 824 were eligible for analysis; PTEN was considered lost in 120 (15%). A total of 745 PTEN-stained slides were submitted for central pathology review. Of these, 48 were removed from analysis due to poor image quality and 22 were removed because of patient ineligibility for other reasons. Agreement between distributed and central PTEN scoring was only fair ( $\kappa = 0.33$ ), likely reflecting a training and experience effect for the central reviewers. PTEN scores from the central review were used for analysis purposes.

MET IHC will be reported in detail in a separate manuscript, in preparation.