MRMC Analysis of Mitotic Counts

Brandon D. Gallas, PhD and Weijie Chen, PhD
Division of Imaging, Diagnostics, and Software Reliability, OSEL/CDRH/FDA

Rationale

We wanted to evaluate the survival outcome prognostic ability of mitotic counts reported by pathologists evaluating glass slides stained with pHH3 (phospho-histone H3) compared to those of the standard stain, H&E (Hemotoxylin & Eosin). The pHH3 stain reacts to cells undergoing mitosis; it stains them red. Counting mitoses is expected to be easier for pHH3 as it is a color detection task rather than a challenging morphologic discrimination task. Furthermore, we wanted to compare the counts that come from evaluating the glass slides on the microscope to those from evaluating whole slide images (WSI's) on a digital display.

Methods

We conducted a study with 12 pathologists and 113 patients. The patients were canines diagnosed with oral melanoma. Survival data included date of death by melanoma for 30 patients, death by unknown cause for 27 patients, and last live contact for 10 patients. A pHH3 slide and H&E slide were prepared for each patient and these slides were scanned with an Aperio AT2 scanner to produce corresponding WSI's. The pathologists were veterinary pathologists distributed across four sites. Data collection followed clinical practice in which pathologists count mitoses in 10 consecutive, non-overlapping, high-power fields of view, starting in an area of high mitotic activity. High-power fields of view are typically those resulting from a 40X objective and a 10X eyepiece, corresponding to approximately 0.24-0.26 mm^2 of tissue. Each pathologist determines the fields of view to evaluate. Counting mitoses with WSI's was done on a digital display by creating circular annotations outlining fields of view with areas equivalent to those on the microscope. These annotations were saved to allow us to investigate the overlap of fields of view chosen by each pathologist. Prognostic performance was evaluated in terms of Harrell's C statistic, the area under the receiver operating characteristic curve (AUC), sensitivity, and specificity. To compare counts from glass slides and the microscope to counts from WSI, we used the (pairwise) probability of concordance, percent difference, and the coefficient of variation. Established multi-reader multi-case (MRMC) analysis methods were used to analyze prognostic performance, while new MRMC analysis methods were developed for the other comparisons.

Results

Preliminary findings show that there is a significant amount of within- and between-reader variability, and that different pathologists choose different fields of view. Analyses are ongoing and more results will be presented at the conference.

Conclusions

Selection of fields of view may be a significant source of the variability observed in the data collected. We intend to complement this study with one where the readers read the same fields of view in both modalities.