Anatomic Pathology Abstracts, 11/15

Editors: Michael Cibull, MD, professor emeritus, University of Kentucky College of Medicine, Lexington; Rouzan Karabakhtsian, MD, attending pathologist, Department of Pathology, Montefiore Medical Center, Albert Einstein College of Medicine, Bronx, NY; Thomas Cibull, MD, dermatopathologist, Evanston Hospital, NorthShore University HealthSystem, Evanston, Ill.; and Rachel Stewart, DO, resident physician, Department of Pathology and Laboratory Medicine, University of Kentucky.

Variation in reporting extraprostatic extension after radical prostatectomy

Extraprostatic extension of prostate cancer in radical prostatectomy specimens significantly affects patient management. The authors evaluated the degree of interobserver variation between uropathologists at a tertiary referral teaching hospital in assessing the extraprostatic extension of prostate cancer in radical prostatectomy specimens. Histopathological data from a consecutive series of 293 radical prostatectomy specimens (January 2007–December 2012) were reviewed. A subset of 50 consecutive radical prostatectomy cases originally staged as tumors confined to the prostate (pT2) or tumors extending into periprostatic tissue (pT3a) were assessed by four specialist uropathologists. Five consultant histopathologists reported on these specimens, with significant differences in the reported stage (P=0.0164) between pathologists. Double-blind review of the 50 consecutive radical prostatectomy cases by the four uropathologists showed a lack of consensus in 16 of 50 (32 percent) cases (κ score, 0.58; moderate agreement). A consensus meeting was held, but a consensus still could not be reached in nine of the 16 cases. The authors' findings highlight variability in the reporting of pT stage in radical prostatectomy specimens, even by specialist uropathologists. Assessment of extraprostatic extension has important implications for patient management, and more precise guidance is needed.

Bryant RJ, Schmitt AJ, Roberts IS, et al. Variation between specialist uropathologists in reporting extraprostatic extension after radical prostatectomy. *J Clin Pathol.* 2015;68:465–472.

Correspondence: Dr. Clare Verrill at clare.verrill@ouh.nhs.uk

[hr]

Prostate biopsy concordance in a large population-based sample

The authors used the Surveillance, Epidemiology, and End Results database to evaluate prostate biopsy concordance in a large population-based sample. They identified 34,195 men who were diagnosed with prostate cancer and underwent radical prostatectomy between 2010 and 2011. All patients had to have clinical and pathological Gleason scores available for analysis. The concordance of the biopsy Gleason score to the pathological

Gleason score was analyzed using the coefficient of agreement (κ). Univariate and multivariate logistic regression analyses were performed to determine potential factors that may impact concordance of Gleason score. Overall, the clinical and pathological Gleason scores matched in 55.4 percent of patients. The concordance rates were 55.3 percent for Gleason 6, 66.9 percent for Gleason 3+4, 42.9 percent for Gleason 4+3, and 24.8 percent for Gleason 8, with frequent downgrading to Gleason 7. The κ for Gleason score concordance was 0.36 (95 percent confidence interval [CI], 0.35–0.37), indicating fair agreement. The weighted κ for Gleason score concordance was 0.51 (95 percent CI, 0.50–0.52), indicating moderate agreement. The Bowker tests of symmetry were highly significant (P<0.001), indicating that when discordant findings were present, pathological upgrading was more common than downgrading. The authors concluded that this study is, to their knowledge, the largest contemporary study of prostate biopsy concordance. They found that there continues to be significant Gleason migration both upward from biopsy Gleason 6 or 3+4 and downgrading from biopsy Gleason 8 or higher. Additional studies are needed to better determine other potential genomic or biologic factors that may help increase biopsy Gleason concordance.

Schreiber D, Wong AT, Rineer J, et al. Prostate biopsy concordance in a large population-based sample: a Surveillance, Epidemiology and End Results study. *J Clin Pathol.* 2015;68:453–457.

Correspondence: Dr. David Schreiber at david.schreiber@va.gov

[hr]

Calculating Ki-67 index in pancreatic neuroendocrine tumors

Ki-67 index is an essential part of the classification of pancreatic neuroendocrine tumors. However, its adaptation to daily practice has been fraught with challenges related to counting methodology. In this study, three reviewers used the following counting methodologies to calculate Ki-67 index in 68 well-differentiated pancreatic neuroendocrine tumors: eye-ball estimation, which is considered reliable and is widely used; automated counting by image analyzer; manual eye-counting (eye under a microscope without a grid); and manual count of camera-captured/printed image. Pearson's correlation (r) was used to measure pairwise correlation among the three reviewers using the four methodologies. The average level of agreement was calculated using the mean of r values. The study found that "eye-balling" was the least expensive and fastest methodology (average time, less than one minute) but had poor reliability and reproducibility. Automated count was the most expensive and least practical and had a major impact on turnaround time (limited by machine and personnel accessibility). More importantly, it had inaccuracies in overcounting unwanted material. Manual eye count was no additional cost and averaged six minutes but proved impractical and poorly reproducible. Camera-captured/printed image was the most reliable and had the highest reproducibility, but it took longer than eye-balling. The study concluded that, based on its comparatively low cost/benefit ratio and reproducibility, camera-captured/printed image appears to be the most practical methodology for calculating Ki-67 index. Although automated counting is generally advertised as the gold standard for index calculation, in this study it was not as accurate or cost-effective as camera-captured/printed image and was highly operator dependent. Eye-balling produces highly inaccurate and unreliable results and is not recommended for routine use.

Reid MD, Bagci P, Ohike N, et al. Calculation of the Ki67 index in pancreatic neuroendocrine tumors: a comparative analysis of four counting methodologies. *Mod Pathol.* 2015;28:686–694.

Correspondence: Dr. V. Adsay at nadsay@emory.edu

[hr]

Use of biomarkers to predict detection of goblet cells for Barrett's esophagus

The diagnosis of Barrett's esophagus in the United States requires both endoscopically evident columnar-lined esophagus and the presence of goblet cells by histology. No consensus exists regarding how patients with

nongoblet columnar-lined esophagus should be followed. The authors conducted a study in which they investigated whether biomarkers can be used to predict the detection of goblet cells in follow-up biopsies. Patients with nongoblet columnar-lined esophagus were identified. In 13 of these cases, goblet cells were detected in subsequent follow-up endoscopic biopsies (Barrett's group). Twenty-six cases that remained negative for goblet cells in follow-up biopsies served as controls. Immunohistochemistry for CDX2, SOX9, BMP4, SHH, and MUC2 was performed on the initial biopsies and graded independently by at least two pathologists in a masked manner. CDX2 was positive in the nongoblet columnar epithelium of seven of 13 cases in the Barrett's group and in four of 26 controls (sensitivity, 54 percent; specificity, 85 percent; odds ratio [OR], 6.4). Strong and diffuse immunoreactivity for SOX9 was detected in 10 of 13 cases in the Barrett's group and in one of 26 controls (sensitivity, 77 percent; specificity, 96 percent; OR, 83.3). Combining CDX2 and SOX9 as a panel increased sensitivity to 85 percent, although the specificity decreased to 85 percent (OR, 30.3). SHH, BMP4, and MUC2 expression showed no significant difference between the Barrett's and control groups. The authors concluded that in patients with nongoblet columnar-lined esophagus, SOX9 and CDX2 may be useful in identifying a subset of patients who have a higher risk of being diagnosed as having Barrett's esophagus (developing goblet cells) and need closer follow-up.

Zhang X, Westerhoff M, Hart J. Expression of SOX9 and CDX2 in nongoblet columnar-lined esophagus predicts the detection of Barrett's esophagus during follow-up. *Mod Pathol.* 2015;28:654–661.

Correspondence: Dr. J. Hart at john.hart@uchos pitals.edu

[hr]

Support for a four-tiered bladder cancer grading system based on WHO classifications

The use of two classification systems for bladder cancer grade is advocated in clinical guidelines because the WHO 2004 classification has not been sufficiently validated with biological markers and follow-up. The slides of 325 primary nonmuscle invasive bladder cancers from three hospitals were reviewed by one uropathologist in two sessions for the WHO 1973 (G1, G2, and G3) and 2004 (papillary urothelial neoplasm of low malignant potential [LMP], low grade [LG], and high grade [HG]) classifications. FGFR3 status was examined with PCR-Snapshot analysis. Expression of Ki-67, P53, and P27 was analyzed by immunohistochemistry. Clinical recurrence and progression were determined. The authors performed validation and cross-validation of the two systems for grade with molecular markers and clinical outcome. Multivariable analyses were done to predict prognosis and pT1 bladder cancer. Grade review resulted in 88 G1, 149 G2, and 88 G3 lesions (WHO 1973), as well as 79 LMP, 101 LG, and 145 HG lesions (WHO 2004). Molecular validation of both grading systems indicated that FGFR3 mutations were associated with lower grades, whereas altered expression (Ki-67, P53, and P27) was found in higher grades. Clinical validation showed that the two classification systems were significant predictors for progression but not for recurrence. Cross-validation of both WHO systems showed a significant stepwise increase in biological (molecular markers) and clinical (progression) potential along the line G1–LG–G2–HG–G3. The LMP and G1 categories had a similar clinical and molecular profile. On the basis of molecular biology and multivariable clinical data, these results support a four-tiered grading system using the 1973 and 2004 WHO classifications, with one low-grade (LMP/LG/G1) category that includes LMP, two intermediate grade (LG/G2 and HG/G2) categories, and one high-grade (HG/G3) category.

Van Rhijn BW, Musquera M, Liu L, et al. Molecular and clinical support for a four-tiered grading system for bladder cancer based on the WHO 1973 and 2004 classifications. *Mod Pathol.* 2015;28:695–705.

Correspondence: Dr. B. W. van Rhijn at basvanrhijn@hotmail.com

[hr]

Link between subclinical thyroid cancer diagnosis rates and survival rates

Survival rates are commonly used to measure success in treating cancer, but they can be misleading. Modern diagnostic practices can lead to cancer survival rates appearing to improve as tumors are diagnosed earlier (lead-time bias) or as an increasing proportion are slow growing (length bias), whereas the actual burden of cancer deaths is unchanged. Increasingly, more subclinical thyroid cancers are being diagnosed. The authors conducted a study to determine whether thyroid cancer survival rates have been affected by this phenomenon. They analyzed survival data from patients with thyroid cancer who were treated at Memorial Sloan Kettering Cancer Center (MSKCC) from 1950 to 2005, as well as United States population-based incidence, prevalence, and survival data from 1973 to 2009 in the Surveillance, Epidemiology, and End Results data set. The authors found that thyroid cancer incidence in the United States increased threefold from 1975 to
2009. The proportion of thyroid cancers that are subcentimeter in size increased from 23 percent in 1983 to 36 percent in 2009. At MSKCC, the percentage rose from 20 percent in 1950 to 35 percent in 2005. The incidence rates of large tumors (greater than 6 cm) and distant metastasis have not changed. In the United States, 10-year relative survival improved from 95.4 percent to 98.6 percent (1983–1999). At MSKCC, 10-year disease-specific survival improved from 91.1 percent to 96.1 percent (1950–2005). However, when stratified by tumor size and stage, no changes in survival outcomes were observed. Thyroid cancer mortality rates have remained stable in the United States (1975–2009). The authors concluded that modern medical practices increasingly uncover small, asymptomatic thyroid cancers. Survival rates appear improved, but this finding is spurious, attributable instead to shifts in the characteristics of the disease being diagnosed. Relying on survival rates to measure success in treating thyroid cancer may reinforce inappropriately aggressive management. Treatment decisions in thyroid cancer should be made based on mortality, not survival data.

Ho AS, Davies L, Nixon IJ, et al. Increasing diagnosis of subclinical thyroid cancers leads to spurious improvements in survival rates. *Cancer.* 2015;121:1793–1799.

Correspondence: Dr. Luc G. Morris at morrisl@mskcc.org