Negative Biopsy Specimens in the Setting of High A Priori Probability of Disease

Is It Time to Reconsider Quality Assurance Routines in a Time of New Diagnostic Capabilities?

Jessica L. Wang, MD, and David C. Wilbur, MD

More and more, newly published management guidelines are touted as evidence driven with extensive data-backed assessments of risk of disease. These risks are determined by outcomes data generated in clinical studies and formal clinical trials. As any pathologist knows, the data derived from any study are only as good as the study inputs and methodology. For instance, if the goal is to identify the risk of disease, one needs to have optimal methods and tools in place to find evidence of the disease if it is present. In addition, and looking more carefully into the process, one needs to appreciate how reproducible the targeted diagnosis actually is and how the sampling of tissue can affect detection sensitivity. It is only when these parameters are appropriately considered that reality-based assessments of risk can be input into management guidelines.

But it is not just guideline development that is important in this discussion. Individual patient results are equally affected by the process noted above. If you apply an extant management guideline to a patient, that application also relies on a sound pathology diagnosis. Hence, the use of an optimized interpretive technique must also apply equally in the setting of the primary diagnosis in order for the guidelines to actually work. Therefore, both crafting and clinical use of any guideline must use optimal—and similar—pathology procedures.

Risk-based management guidelines for cervical cytology have been recently published that were crafted using extensive evidence-based literature reviews and expert opinion. These guidelines were based on the outcomes data grounded by extensive follow-up studies aimed at identifying the risk of high-grade cervical lesions over time for the various cytologic abnormalities. Based on risk levels observed in these studies, appropriate management could be defined. For instance, in patients having Pap test interpretations of a high-grade squamous intraepithelial lesion (HSIL), the risk of biopsy-proven preneoplastic lesions (high-grade dysplasia) is 85% to 90%, a level that certainly warrants immediate colposcopic examination. On the other hand, patients having cytologic interpretations of atypical squamous cells of undetermined significance and an associated negative test for high-risk human papillomavirus (HPV) have less than a 1% risk of precancer, a level warranting no current intervention with a return to routine screening. In each of the studies reviewed in the development of these guidelines, outcomes were based on pathology diagnoses. What if such diagnoses were “clinically” incorrect, meaning that they might be technically correct but incomplete because they did not “find” the disease that was actually present? As noted above, both the guideline and the clinical outcome might therefore be flawed.

It is in the above context that the article by Carrigg et al published in this issue of the Journal gains significant importance. The identification of cervical abnormalities on Pap tests leads to biopsies, the diagnosis of which guides downstream management and follow-up. As our ability to more sensitively detect cytologic abnormalities improves, smaller precancerous lesions are being identified on colposcopic examination. Small punch biopsy specimens and, to a lesser extent, larger cone or loop electrosurgical excision procedure biopsy specimens by their very nature imply potential incomplete sampling of small lesions. Add to this the fact that histologic preparations of tissue are 2-dimensional representations of 3-dimensional specimens. Add yet again the fact that not all precancerous lesions look alike, have histologic mimics, and have inherent rates of reproducibility in the subjective art of microscopic interpretation. Negative cervical biopsy
Two-dimensional subsampling of a 3-dimensional specimen is not difficult to understand as a cause for false negativity, and deeper levels provide a clear-cut resolution. However, histologic evaluation by itself can be challenging. There is significant morphologic overlap between benign mimics and dysplastic lesions of the cervix. Interobserver agreement in cervical biopsy specimen interpretation has been shown to be quite low, an indication of this difficulty. The p16 immunohistochemical stain has been touted as a means to more reliably identify precancerous lesions and has been shown to improve interobserver agreement.

The findings of the study by Carrigg et al lend further support to the recent recommendations of the Lower Anogenital Squamous Terminology (LAST) project, a publication related to HPV-associated disease, which was cosponsored by the College of American Pathologists and the American Society for Colposcopy and Cervical Pathology. The LAST project encourages appropriate use of p16 immunohistochemical stains as part of cervical biopsy specimen evaluation. Via an extensive literature review and consensus process, LAST recommended that p16 should be used to differentiate between benign mimics or HPV infections and precancers, adjudicate differences of opinion, and identify missed areas of precancer in high-risk specimens (meaning those specimens having a high a priori probability of harboring precursor that initially had been evaluated as negative). This latter recommendation is highlighted in the current work and shows the utility of this process.

Interestingly, Carrigg et al also found that consensus review identified prior unrecognized dysplasias. This result has not been described previously but suggests that high-risk cases might benefit from a consensus review on a systematic rather than a discretionary basis.

This study highlights an issue that one might refer to as “pathology’s dirty little secrets”: that tissue obtained is subsampled, pathologists disagree with one another on a regular basis, and subjective opinions are just that—opinions. Objectification with tests such as p16, which tells us something about the biology of the case, along with systematic quality assurance processes such as examining more tissue in high-risk cases (deeper levels) or consensus review to obtain more (and perhaps better) opinions, should increase detection accuracy. On the basis of these tools, Carrigg et al suggest an algorithm for cervical cone biopsy specimens. Perhaps these principles may also be applied more broadly in the future, with consensus processes such as in LAST serving as a guide to more scientific, evidence-based approaches to pathology practice.

From the James Homer Wright Laboratory of Pathology, Massachusetts General Hospital, Harvard Medical School, Boston, MA.
Wang and Wilbur / Negative Biopsy Specimens

References


