Laboratory Management of CIN 2
The Consensus Is Consensus

Christopher P. Crum, MD

DOI: 10.1309/GCVB5C4KFV7TFC7F

Since the first associations between human papilloma-virus (HPV)-16 and cervical squamous neoplasia 25 years ago, the field of cervical cancer prevention and precursor diagnosis has continued to evolve, culminating recently with the validation of vaccines that promise to significantly reduce the incidence of this disease. The significance of HPV-16 was appreciated early by its strong association with higher grade cervical intraepithelial neoplasia (CIN) and squamous carcinoma, both of which contained this HPV in nearly 60% of cases.1,2 Subsequent studies, in vitro and in the clinical setting, have validated the carcinogenic properties of HPV-16 and its significance as a risk factor for the development of high-grade squamous intraepithelial lesions (HSILs; CIN grade 2 or 3) detected by follow-up biopsy.3,4

Although the strong theoretical link between HPV-16 and cervical premalignancies has been validated by the success of the recently developed vaccines, the link has been inconsistently translated from biologic reality to diagnostic standard. The reason for this is rooted in the historic evolution of precursor management and diagnosis. The first has had an important impact on the second. With the adoption of relatively nontraumatic approaches to cervical neoplasia in the late 1970s—cryotherapy and laser ablation—the principal issues that governed management were ensuring that the patient had a CIN and, if so, excluding CIN 3. CIN 1 was managed typically by cryoablation and CIN 3 by cone biopsy. A diagnosis of CIN 2 (or a small CIN 3 lesion) usually resulted in cryotherapy or laser ablation. In retrospect, what was immediately apparent from this philosophy was that the distinction between CIN 1 and CIN 2 was essentially irrelevant in terms of management. Either cryotherapy or laser could be administered in an outpatient setting to both, and the clinical indications did not require separating CIN 1 from CIN 2 when reviewing the biopsy specimen.5

In the late 1980s, large loop electrical excision (large loop excision of the transformation zone [LLETZ] or loop electrode excision procedure [LEEP]) was introduced and rapidly embraced by the clinical field.6 This modality brought with it a fundamental change in the management of women with early cervical neoplasia. First, because LEEP was equivalent to conization, it was not appropriate for women with low-grade squamous intraepithelial lesions (LSILs; condyloma/CIN 1), given the low risk of progression to malignancy. Thus, LSIL was relegated to cytologic follow-up, sparing countless young women the discomfort of cryotherapy or laser ablation. However, because of the strong association between CIN 2 and CIN 3 with high-risk HPVs and the reluctance to leave CIN 2 untreated, both of these entities were treated as HSIL, and managed by LEEP.7 This approach brought with it an immediate increase in the number of women who were designated for surgical (LEEP) ablation and exposed a specific problem in the selection process for who would be treated.

Pathologists have traditionally shown a range of interobserver disagreement in grade-by-grade classification of early cervical neoplasia.8,9 This disagreement is not surprising for a number of reasons. First, the cervical transformation zone harbors a wide range of HPV-induced atypias that reflect the interplay of numerous HPV types superimposed on an epithelial environment with a wide range of phenotypic plasticity (metaplasia).10 Second, how this range of epithelial changes is interpreted varies widely among experts, not to mention between experts and general practitioners.9 Third, the cutoff points for the grading scheme have never been scrutinized in a systematic manner, complicated in part by different perceptions...
of CIN across different generations and training programs. For example, the current precursor classification incorporates lesions once viewed as “koilocytosis” such as flat or exophytic condylomas of the cervix, lesions largely ignored in classic texts on cervical preinvasive disease.\textsuperscript{11,12} Incorporating these entities into a CIN or SIL classification requires a modification of the criteria and shifting most classically described CINs into the higher grade categories of CIN 2 and CIN 3.\textsuperscript{11} It is ironic that the cytologic classification of SIL, based on a very limited sampling, has met with more success than the histologic classification, directed at the morphologic end point.

At the center of the diagnostic conundrum is CIN 2. This diagnosis will generally result in a LEEP. Moreover, CIN 2 (and CIN 3) is traditionally considered a histologic outcome for success or failure in vaccine and therapeutic trials.\textsuperscript{13,14} As originally described and currently used, the diagnosis of CIN 2 has a significant association with HPV-16.\textsuperscript{15} Nevertheless, as a lesion that may share features with CIN 1 (koilocytosis) and CIN 3 (increasing atypia), CIN 2 may be interpreted as either, depending on the observer. For this reason, the value of CIN 2 as a reliable outcome for HSIL has been challenged by some, based on its weaker association with high-risk HPVs, creating the perception that CIN 2 is sufficiently different from CIN 3 to be separated from it.\textsuperscript{15} The separation of CIN 2 from CIN 3 is, thus, the issue that must be addressed.

The recent study by Galgano and colleagues\textsuperscript{17} takes a slightly different tack relative to prior studies by asking if HPV-16 can be used as a quality assurance metric to assess agreement between observers. In essence, they determined the degree to which agreement between clinical center (site) pathologist and central consensus pathology review (quality control [QC]) correlated with HPV-16. The premise to be tested was that lesions containing HPV-16 would be more likely to be diagnosed with greater reproducibility among pathologists as CIN 2 or greater. Similar trends have been reported previously.\textsuperscript{8} The authors used data from the Atypical Squamous Cells of Undetermined Significance–Low-Grade Squamous Intraepithelial Lesion Triage Study, using the HPV results from a cervical cytology specimen to ascertain the HPV type and compare with the subsequent histologic diagnosis.

The authors found the following: (1) There was a strong overall association between increasing lesion grade and HPV-16, a fact that has never been challenged and reflects the fundamental association between HPV-16 and higher grade precursors.\textsuperscript{2,18} (2) There was considerable variation among pathologists in the study with respect to associations between HPV-16 and CIN grade, which is depicted in their Figure 2, which illustrates the ratio of HPV-16 positives between a site pathologist and the central consensus (QC) review.\textsuperscript{17}

It is clear that the associations between HPV-16 and diagnosis vary greatly among the site pathologists, including negative (5%–12%), CIN 1 (9%–23%), CIN 2 (30%–60%), and CIN 3 (48%–79%). One wonders to what degree the QC pathologists agreed with one another with respect to HPV-16 before establishing the consensus diagnosis. On examination of Figure 2, it is also clear that site pathologists were more likely to overclassify CIN 1 relative to the QC diagnosis, based on the generally lower frequency of HPV-16 association, a level of discrepancy that is commonly encountered for this diagnosis across multiple observers. What is of interest is that 4 of the site pathologists were considerably more likely to classify CIN 2 as an HPV-16–associated lesion relative to the QC diagnosis. Three of these pathologists had a higher index of association between their CIN 3 diagnosis and HPV-16, perplexing figures that imply that these site readers are more likely than the central consensus to correlate high-grade CIN with HPV-16. The implication is that although many pathologists tend to overcall CIN 1, the same is not necessarily true for CIN 2 or CIN 3, depending on the observer.

Irrespective of how the aforementioned discrepancies are interpreted, they point to a lack of consistency with which CIN 2 is classified and, in turn, correlated with HPV-16 across observers. One other important finding, however, was the observation by these authors that as the $\kappa$ values for agreement on CIN 2 between site and QC review increased, the differences in percentage of HPV-16 recorded by each for CIN 2 or CIN 3 narrowed. To me, such findings, while not solving the problem of suboptimal reproducibility, indicate that agreement and predictions of HPV-16 positivity are linked. Moreover, another statistic in this study implies that the solution to this reproducibility problem lies not in searching for an objective standard (such as HPV-16) but in recognizing a strategy that could have considerable clinical usefulness, which is relying on more than one observer to confirm a diagnosis of CIN 2.

When both observers agreed on a diagnosis of CIN 1, 17% of the samples contained HPV-16. When either disagreed, the index of HPV-16 was between 21% and 31%. When both observers agreed on CIN 2, the index of HPV-16 was 44%. This is an endorsement for a second observer opinion when confirming a diagnosis of CIN 2 and the reason why current protocols require agreement by at least 2 of 3 observers before recording a diagnosis of CIN 2 or CIN 3.

My experience is the same as that of Galgano et al.\textsuperscript{17} My colleagues and I have found that when 2 of 3 observers (1 site pathologist and 2 central pathologists) agreed on a diagnosis of CIN 2, the index of HPV-16 was 43%. Similar to the study by Galgano et al,\textsuperscript{17} the index of HPV-16 for a majority diagnosis of CIN 3 was 61%. Moreover, similar to their study, the presence of high-risk HPV (HPV-16 or other HPVs) increased the concordance rates of CIN 2 (but not CIN 3). When the biopsy followed an HSIL cytologic diagnosis, the indices of HPV-16 for a majority diagnosis of CIN 2 and CIN 3 were 64% and 71%, respectively. The fundamental
question that comes up in comparisons of this type is whether CIN 2, which is often poorly reproduced, is more likely to be confused with CIN 1 or CIN 3. We found that the former was more likely when validating the diagnosis of a site pathologist (which varied), in keeping with the findings of Galgano et al. In contrast, in many cases in which central pathologists disagreed on CIN 2, it was because one observer’s CIN 2 was another’s CIN 3 and vice versa, such that 86% of CIN 2 diagnoses were corroborated as HSIL by the second central pathologist. In general, when CIN 2 and CIN 3 are grouped together, the ability of observers to separate them from anything less than CIN 2 is substantial. To be fair, such levels of concordance are observer pair–dependent and may not be translated to all observer pairs who happen to concur on a diagnosis of CIN 2. Nevertheless, the cumulative experience strongly suggests that when 2 observers independently agree on a diagnosis of CIN 2, the correlation with HPV-16 increases significantly.17,19,20

In my opinion, the findings by Galgano et al and others support a laboratory management algorithm in which all biopsy diagnoses of CIN 2 are reviewed by a second pathologist and disagreements (as in a second opinion of CIN 1 or less) are thoughtfully resolved. What is at stake is whether the patient who is waiting for the report will be told she will need a procedure that is going to cause considerable anxiety, cost, inconvenience, and discomfort. Given the emergence of biomarkers that could correlate with a CIN 2 or CIN 3 outcome following high-risk HPV infection, it is not far-fetched to envision one or more immunostains that would predict which lesions in the CIN 1–CIN 2 range would persist or progress in grade and merit ablation. However, until such a milestone is reached, our patients deserve the best that can be achieved by their health care providers. In the case of CIN 2, a second opinion is highly recommended.

References


