Histological 'progression' from low (LSIL) to high (HSIL) squamous intraepithelial lesion is an uncommon event and an indication for quality assurance review

Eleanor Y Chen, Ann Tran, Caroline J Raho, Chandler M Birch, Christopher P Crum and Michelle S Hirsch

Division of Women's and Perinatal Pathology, Department of Pathology, Brigham and Women's Hospital, Boston, MA, USA

An accurate assessment of 'progression' from a low (LSIL) to high (HSIL) grade squamous intraepithelial lesion (cervical intraepithelial neoplasia (CIN)2 or CIN3) of the cervix is critical to ascertaining HSIL outcome risk, the value of predictive biomarkers, and the need for excisional therapy. We obtained biopsy outcome data on a series of initially diagnosed LSIL to assess this risk. Consecutive biopsy diagnoses of LSIL were obtained from the archives, and the frequency of HSIL biopsy outcomes were ascertained by record and histological review. Then, a 'numerical severity score' was recorded for each diagnosis: LSIL (1–2), CIN2 (3–4) and CIN3 (5–6) with lower and higher values corresponding to the degree (low vs high) of histological severity within each category, respectively. Of 264 initial LSILs, 29 (11%) were reported with an HSIL outcome. However, histological review of 21 of these HSILs confirmed only 8 (38%) HSIL diagnoses by review with the numerical severity score: three cases scored as 5, three cases scored as 4, and two cases scored as 3; the remaining 13 cases were assigned a numerical severity score of 1 or 2. P16 immunostains of corresponding previous and subsequent biopsies were discordant in 4 of 12 cases (33%). In a blind review of a randomly selected series of HSILs from the same practice, HSIL was significantly more likely to be confirmed on re-review (10 of 13 (77%), *P* = 0.024). These findings show that confirmed HSIL outcomes (on review) following an LSIL biopsy are infrequent (∼3%). A diagnosis of HSIL following an LSIL should always be reviewed, as this diagnostic pairing may more likely be associated with a diagnostic error.
Modern Pathology (2010) **23**, 1045–1051; doi:10.1038/modpathol.2010.85; published online 14 May 2010

Histological evaluation remains the basis for treatment and follow-up of women with cervical intraepithelial neoplasia (CIN). The fundamental premise for treating or following young women with CIN hinges on the risk of an eventual outcome of CIN2 or CIN3 (high-grade squamous intraepithelial lesion or HSIL), diagnoses for which management by cone biopsy or LEEP will be required.[1] Hence, the progression rate from CIN1 to CIN2 or CIN3 is germane to conservative management of low-grade squamous intraepithelial lesions (LSILs) (CIN1 or condyloma). In a review by Ostör[2] summarizing the findings of studies dealing with progression of CIN since 1950, the composite data indicated that the likelihood of progression from CIN1 to CIN3 was ∼10%. In a recent study by Pretorius et al[3], the subsequent risk of CIN3 or cancer after a colposcopic diagnosis of CIN1 or less was found to be 1.9%, a rate that is lower than the Ostör study. In a study from the ALTS trial, Cox et al[4] showed that a biopsy diagnosis of CIN1 was equivalent to a negative colposcopic biopsy following an HPV-positive smear showing LSIL or ASCUS, despite that fact that 11–13% of cases were followed by an HSIL outcome within 2 years. Thus, although these

studies compute different rates of HSIL outcome, they convey the impression that CIN1 carries a risk level that requires follow-up.

Precisely determining which CIN1 cases will progress to CIN2 or CIN3 has been a focus of study. Investigators have attempted to identify biomarkers of risk that would enrich the CIN1 population for those in need of ablative therapy. Conversely, those excluded by a biomarker analysis conceivably could be spared increased scrutiny. Two biomarkers that have received attention include Ki-67 and P16.[5–7] Investigators have proposed that the use of these and other biomarkers can help identify which lower-grade lesions will progress over time.[8,9] For example, in a study examining p16, Negri *et al*[10] reported that CIN1 cases with diffuse p16 staining had a significantly higher tendency to progress to a high-grade lesion than p16-negative cases. Hence, they recommended p16 as a useful surrogate marker for progression likelihood.

We and others have previously shown that during intervals of follow-up, alterations in histopathological grade are not uncommon, and are not always resolved within the context of the same HPV infection.[11–13] Moreover, in young women, HPV infections fluctuate widely over time, explaining further the differences in morphology observed in biopsies taken over a long interval.[14] These realities are compounded further by the fact that the reproducibility of a diagnosis of CIN2 will vary between different observers. Thus, documenting 'progression' requires both accurate diagnoses for the initial and subsequent biopsies, as well as ensuring that the initial and outcome biopsies reflect the same biological process.[15] The purpose of this study was to address this issue by examining the rate of CIN2 or CIN3 outcome following a histological diagnosis of LSIL (CIN1), including a critical review of the original and outcome diagnoses, and comparison of p16 immunostaining in the two. The intention was to ascertain what factors contribute to perceived histological 'progression' of LSIL to HSIL.

Materials and methods

Consecutive cases of LSIL based on biopsy diagnosis were obtained from the archives of the Division of Women's and Perinatal Pathology at Brigham and Women's Hospital, Boston, MA, USA. Pathology and cytology records were reviewed with attention to those with a subsequent HSIL (CIN2 or 3) diagnosis based on biopsy, LEEP, or cone material. In all such cases in which an HSIL outcome was documented, the initial and subsequent pathology was blindly re-reviewed by a single observer, and diagnoses were confirmed or re-classified as LSIL (CIN1) or HSIL (CIN2 or CIN3) using published criteria.[16,17]

To account for nuances in interpretation of lesion grade, a 'numerical severity score' was recorded for each case, with the lower number reflecting less and the higher number more severe histological change within a given grade. Numerical severity scores were as follows: LSIL (1–2), CIN2 (3–4), and CIN3 (5–6), again with lower and higher values corresponding to the degree (low *vs* high) of perceived histological severity within each category, respectively. For reference, 53 consecutive biopsy diagnoses from outside the study in which the pathology report noted an abnormality were also blindly reviewed. The reason for this added review is discussed in the Results section below.

Where tissue was available, blocks were re-cut and sections were immunostained for p16 as previously described.[7] The staining patterns were scored based on the parameters described by Keating *et al*[7]; positive if continuous staining was seen in the horizontal plane, either partial or full thickness. If the staining was interrupted in this plane so that <80% of the epithelium stained positive, the staining was scored as patchy or focal. Weakly diffuse staining or no staining was scored as negative.

Parameters scored included frequency of HSIL outcome, agreement with the previously recorded diagnosis, numerical severity score, and p16 immunostaining patterns.

Results

Of the 264 consecutive biopsy cases of LSIL reviewed, 29 (11%) were associated with a subsequent HSIL diagnosis, and material for review (both initial and subsequent cases) was available in 24 of these cases. The follow-up period was 1 month to 5 years.

Table 1 summarizes the pathology reviews. SIL was confirmed in all initial biopsies with 22 confirmed LSILs (3 with numerical severity score = 2, 19 with numerical severity score = 1); 2 initial cases were re-classified as CIN2 (1 with numerical severity score = 3 and 1 with numerical severity score = 4) and were therefore excluded from the follow-up calculations. For the group of 22 initial biopsies with a confirmed LSIL diagnosis on re-review, follow-up samples were available in 21. Of these, eight (38%) were confirmed as HSIL (CIN2 or CIN3) on review (three with numerical severity score = 5; three with numerical severity score = 4; and two with numerical severity score = 3); the remaining thirteen cases were re-classified as LSIL (see examples in Figure 1). Twelve cases were originally classified as CIN2 (three cones/LEEPs and nine biopsies) with only eleven available for review: only two were confirmed as HSIL (CIN2), each with a numerical severity score of 3 (Table 1). Six of these eleven CIN2 cases were re-classified as LSIL with a numerical severity score of 2, and three additional re-classified cases had a numerical severity score of 1. Essentially, the histological

Table 1 Comparison of review of HSIL diagnoses following LSIL and random review of practice cases

Outcome HSIL biopsies following LSIL			*Random HSIL review*		
Original diagnosis	*Review*	*Score*	*Original diagnosis*	*Review*	*Score*
HSIL (CIN2–3)	CIN3	5	HSIL	CIN3	6
CIN2	None	0	HSIL	CIN3	5
CIN2	CIN1	2	CIN2	CIN2	3
CIN2	CIN2	3	HSIL	CIN2	4
CIN2	CIN1	1	HSIL	CIN1	1
CIN2	CIN1	2	HSIL	CIN2	4
HSIL (CIN2–3)	CIN1	1	HSIL	CIN1	2
CIN2	CIN1	1	HSIL	CIN2	3
CIN2	CIN1	2	HSIL	CIN3	6
CIN2	CIN1	2	HSIL	CIN3	6
HSIL (CIN2–3)	CIN1	2	CIN2	CIN2	3
CIN2	CIN1	2	HSIL	CIN2	4
HSIL (CIN2–3)	CIN1	2	HSIL	CIN3	5
HSIL (CIN2–3)	CIN1	2	CIN2	CIN2	3
HSIL (CIN2–3)	CIN2	4	HSIL	CIN2	4
HSIL (CIN2–3)	CIN3	5	HSIL	CIN1	2
CIN2	CIN1	1			
HSIL (CIN2–3)	CIN3	5			
CIN2	CIN1	2			
CIN2	CIN2	3			
HSIL (CIN2–3)	CIN2	4			

CIN, cervical intraepithelial neoplasia; HSIL, high-grade squamous intraepithelial lesion; LSIL, low-grade squamous intraepithelial lesion.

grades of over 60% of cases with a reported HSIL outcome could not be readily distinguished from the initial biopsy in the severity of grade.

To determine whether the low level of agreement between original and reviewed outcome biopsy diagnosis (8 of 21 or 38%) reflected the pathology practice as a whole, 53 consecutive cases with a history of an abnormal cytology were pulled from the files and reviewed without knowledge of the diagnosis or proportion that were either LSIL or HSIL. In this group, 13 had an original diagnosis of HSIL (CIN2 or CIN3). On review, agreement with this diagnosis was achieved in 10 cases (77%). The difference in rate of HSIL confirmation between the two groups ('progression' study *vs* random; 38 *vs* 77%) was significant at $P = 0.024$ (Fisher's exact test) (Table 1).

Immunostains for p16 were performed on 12 corresponding sets of cases and are summarized in Table 2. Of the twelve initial and follow-up biopsy pairs, eight displayed identical staining patterns; the remaining four cases were distinctly different in staining distribution (Figure 2). Of the latter, three displayed negative staining on initial biopsies and diffuse positive staining on subsequent/follow-up material. Two cases show diffuse positive staining on initial biopsies and focal and negative staining patterns on follow-up material, respectively.

Discussion

Several findings in this study call into question the assumption that progression from LSIL to HSIL is a common event. First, as demonstrated in previous studies, HPV types can fluctuate over time, leading to the false impression that a given LSIL (or HSIL) has actually progressed (or regressed) to a higher- (or lower) grade lesion, respectively.[12] Detailed HPV testing data on women who have been followed with a biopsy-confirmed HSIL have shown shifts in HPV type over time.[12] In a significant percentage of cases in which HSIL was followed by an LSIL outcome biopsy, additional HPV types emerged or replaced the original, implying that HSILs do not actually regress in severity but are replaced by new lesions. Similarly, the impression that an LSIL has actually progressed to HSIL can be erroneous if the latter signifies a new lesion with a completely different HPV type. The differences in p16 immunostaining in four of the twelve initial and follow-up biopsy pairs in this study imply that a new lesion arose following the initial biopsy that could be associated with a different HPV type. This assumption is based on the fact that distinctly different HPV types or groups produce diffuse *vs* patchy p16 immunostaining.[7]

The most compelling observation in this study was the difference in interpretation of SIL grade between the original and second observers, particularly on outcome biopsy or LEEP, emphasizing both an obvious and a more subtle issue that must be addressed in determining whether an HSIL outcome has occurred. The first and central issue is the reproducibility of a diagnosis of CIN2. Some reports have argued that CIN2 is poorly reproduced and have even suggested that CIN2 is an entity that does not fit well within the HSIL spectrum.[18] However,
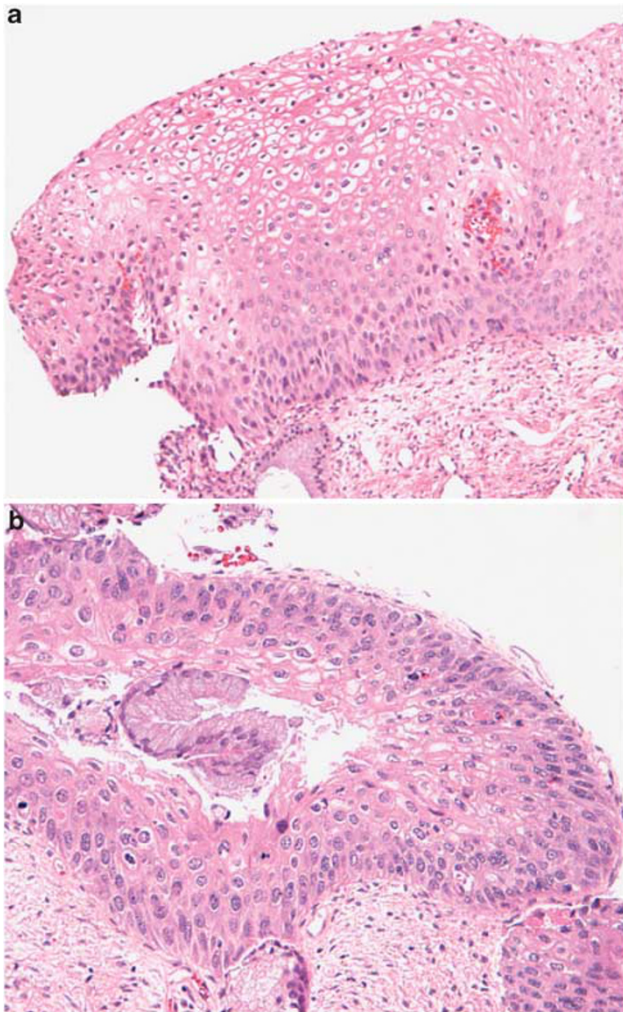
Figure 1 Examples of subsequent/follow-up specimens with original diagnosis of CIN2 re-classified as LSIL. (**a**) Numerical severity score = 1 (less severe LSIL). (**b**) Numerical severity score = 2 (more severe LSIL).

Table 2 Summary of numerical severity scores and p16 immunostaining patterns for initial and outcome specimens

Initial biopsy		Outcome biopsy	
p16	NSS	p16	NSS
Neg.	1	DP	5
DP	1	Focal	0
Neg.	1	Neg.	2
DP	1	DP	3
DP	1	DP	1
Neg.	1	DP	2
Neg.	1	Neg.	1
Neg.	1	DP	1
DP	1	DP	2
DP	1	DP	2
DP	1	DP	2
DP	3	DP	3

DP, diffusely positive; neg., negative staining.

the more compelling argument has been that the above discrepancy is much more likely for a single observer.[19] In contrast, when the diagnosis of CIN2 requires agreement by two or more observers, its association with HPV16 nucleic acids is more than 40%, and exceeds 60% when coupled with a concurrent HSIL cytology.[20] On the basis of this, we have recommended previously that all CIN2 diagnoses in a practice be verified by another observer before reporting this diagnosis, given the greater likelihood that the patient will undergo a LEEP and the questionable need for LEEP when the diagnosis is not certain, particularly in young women.[20]

The second and less obvious issue, but one that may have important implications for both patient management and research studies, is the distinct possibility that follow-up diagnoses of HSIL following an LSIL biopsy are more likely to reflect a diagnostic error when reviewed by a single individual. As shown in previous reports, the rate of cancer following an LSIL diagnosis is so sufficiently low (~1%) that any follow-up diagnosis of cancer in this circumstance would prompt both the review of the original biopsy and the outcome diagnosis of malignancy.[2] The likelihood that one of the diagnoses was in error would be presumed to be higher than if the initial and outcome diagnoses were HSIL and cancer, respectively, as the latter two diagnoses have a strong etiological link. In this study, the high level of nonconfirmed HSILs in the LSIL follow-up study (71%) was significantly higher than that seen in randomly reviewed HSILs from the practice (38%). This suggests that similar to the aforementioned example of cancer following an LSIL biopsy, there is significant risk that an HSIL diagnosis following an LSIL biopsy is a diagnostic error and should be reviewed by an additional observer. Less likely is that the initial LSIL biopsy is an undercalled HSIL (CIN2), which only happened in 2 of the 24 cases reviewed (8%). In this study, all the histological diagnoses were not reviewed, but the outcome rate of verified HSIL following a diagnosis of LSIL could be estimated at 3% (8 of 264), which is in line with the report of Pretorius *et al.*[12]

As in any study of this type, the issue of sampling error must be addressed. Detailed entry cytological data are not available in all cases and it is conceivable that as many as 25% of the LSIL biopsies in this study followed an HSIL smear. When this is the case, an eventual HSIL outcome has been recorded in ~30%.[21] However, the impact of such an occurrence would seem minimal given the overall low frequency of HSIL biopsy outcomes. The ultimate risk of CIN3 cannot be determined in this population. The risk following LSIL cytology in older women has been estimated at <2% by some; nevertheless, periodic follow-up would be advised.[22]

In summary, the most likely explanations for 'progression' from LSIL to HSIL, based on record and histological review, are, in increasing frequency,

(1) actual progression, (2) underdiagnosis of HSIL on initial biopsy, (3) change in HPV type over time (based on p16 stain discrepancy), and (4) over-

diagnosis of HSIL on follow-up biopsy/cone. The possibility of the latter circumstance is accentuated when the follow-up diagnosis is CIN2. The unu-
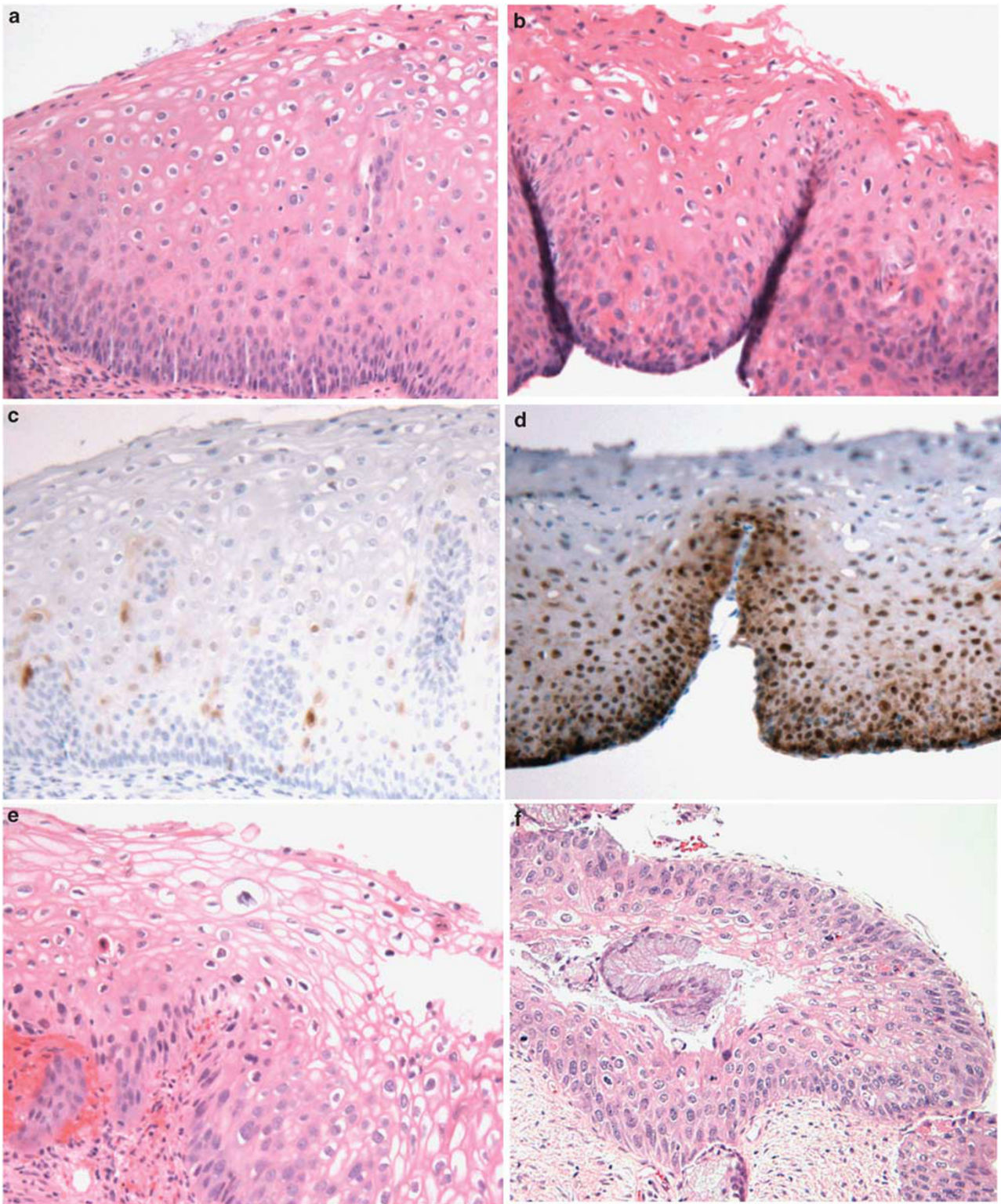
Figure 2 P16 immunostains of initial and subsequent/follow-up specimens. Panels (**a–d**) show H&E and p16 stains from a set with discrepant p16 immunostain patterns. The initial biopsy (**a**) has numerical severity score of 1 with a negative p16 immunostain, whereas the subsequent specimen (**c**) has numerical severity score of 2 with a positive p16 immunostain (**d**). Panels (**e–h**) show H&E and p16 stains from a set with similar p16 immunostaining pattern. Panel (**e**) represents the initial biopsy and (**g**) represents the subsequent specimen.
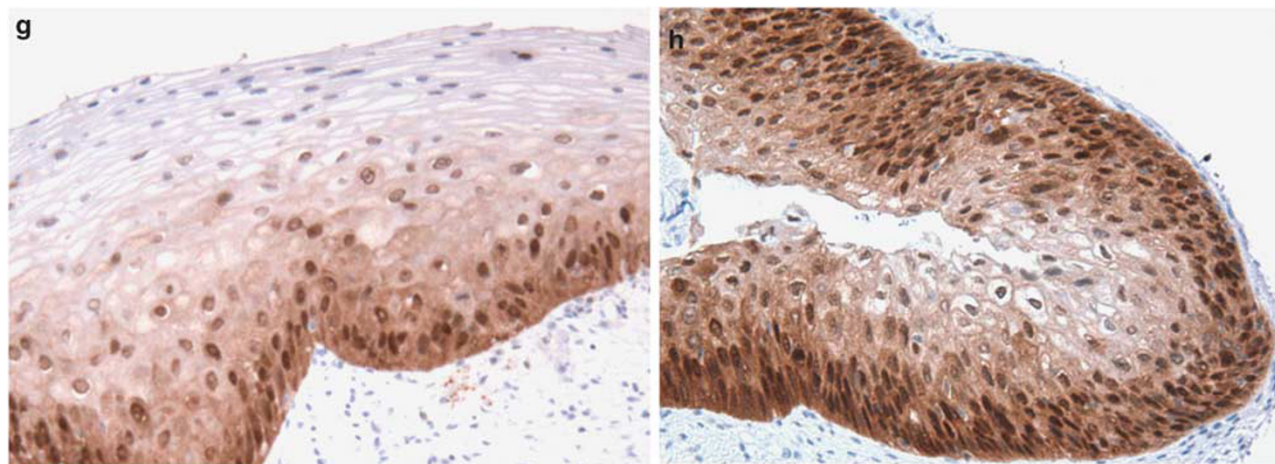
Figure 2 Continued.

sually high percentage of apparent overdiagnosis of HSIL outcomes underscores the uncertainty of this diagnosis in the setting of a previous LSIL. Histological re-review of any case perceived by one observer to have progressed to CIN2 or CIN3 is important to determine with the greatest precision so as to guide who will require LEEP or cone biopsy, and is germane to any study in which conclusions are based on an interpretation of the outcome biopsy.

Disclosure/conflict of interest

The authors declare no conflict of interest.

References

1 Wright Jr TC, Cox JT, Massad LS, *et al.* Consensus guidelines for the management of women with cervical intraepithelial neoplasia. 2001ASCCP-sponsored Consensus Workshop. J Low Genit Tract Dis 2003;7: 154–167.

2 Ostör AG. Natural history of cervical intraepithelial neoplasia: a critical review. Int J Gynecol Pathol 1993;12:186–192.

3 Pretorius RG, Peterson P, Azizi F, *et al.* Subsequent risk and presentation of cervical intraepithelial neoplasia (CIN) 3 or cancer after a colposcopic diagnosis of CIN 1 or less. Am J Obstet Gynecol 2006;195:1260–1265.

4 Cox JT, Schiffman M, Solomon D, ASCUS-LSIL Triage Study (ALTS) Group. Prospective follow-up suggests similar risk of subsequent cervical intraepithelial neoplasia grade 2 or 3 among women with cervical intraepithelial neoplasia grade 1 or negative colposcopy and directed biopsy. Am J Obstet Gynecol 2003;188:1406–1412.

5 Sano T, Oyama T, Kashiwabara K, *et al.* Expression status of p16 protein is associated with human papillomavirus oncogenic potential in cervical and genital lesions. Am J Pathol 1998;153:1741–1748.

6 Klaes R, Friedrich T, Spitkovsky D, *et al.* Overexpression of p16(INK4A) as a specific marker for dysplastic and neoplastic epithelial cells of the cervix uteri. Int J Cancer 2001;92:276–284.

7 Keating JT, Cviko A, Riethdorf S, *et al.* Ki-67, cyclin E, and p16INK4 are complimentary surrogate biomarkers for human papilloma virus-related cervical neoplasia. Am J Surg Pathol 2001;25:884–891.

8 Baak JP, Kruse AJ, Robboy SJ, *et al.* Dynamic behavioural interpretation of cervical intraepithelial neoplasia with molecular biomarkers. J Clin Pathol 2006;59:1017–1028.

9 Baak JP, Kruse AJ, Janssen E, *et al.* Predictive testing of early CIN behaviour by molecular biomarkers. Cell Oncol 2005;27:277–280.

10 Negri G, Bellisano G, Zannoni GF, *et al.* p16 ink4a and HPV L1 immunohistochemistry is helpful for estimating the behavior of low-grade dysplastic lesions of the cervix uteri. Am J Surg Pathol 2008;32:1715–1720.

11 Nuovo GJ, Pedemonte BM. Human papillomavirus types and recurrent cervical warts. JAMA 1990;263: 1223–1226.

12 Crum CP, Beach KJ, Hedley ML, *et al.* Dynamics of human papillomavirus infection between biopsy and excision of cervical intraepithelial neoplasia: results from the ZYC101a protocol. J Infect Dis 2004;189: 1348–1354.

13 Medeiros F, Yuan L, Breslin G, *et al.* Type-specific HPV testing as a predictor of high-grade squamous intraepithelial lesion outcome after cytologic abnormalities. J Low Genit Tract Dis 2005;9:154–159.

14 Rosenfeld WD, Rose E, Vermund SH, *et al.* Follow-up evaluation of cervicovaginal human papillomavirus infection in adolescents. J Pediatr 1992;121:307–311.

15 Pinto AP, Crum CP. Natural history of cervical neoplasia: defining progression and its consequence. Clin Obstet Gynecol 2000;43:352–362.

16 Wright TC, Kurman RJ, Ferenczy A. Precancerous lesions of the cervix. In: Kurman RJ, Blaustein A (eds). Blaustein's Pathology of the Female Genital Tract 1994, 4th edn., Springer-Verlag: New York, pp 229–278.

17 Crum CP, Rose PG. Cervical squamous neoplasms. In: Crum CP, Lee KR (eds). Diagnostic Gynecologic and Obstetric Pathology. Philadelphia: WB Saunders, 2006.

18 Carreon JD, Sherman ME, Guillén D, *et al.* CIN2 is a much less reproducible and less valid diagnosis than CIN3: results from a histological review of population-based cervical samples. Int J Gynecol Pathol 2007;26:441–446.

19 Galgano MT, Castle PE, Stoler MH, *et al.* Can HPV-16 genotyping provide a benchmark for cervical biopsy specimen interpretation? Am J Clin Pathol 2008;130: 65–70.

20 Crum CP. Laboratory management of CIN 2: the consensus is consensus. Am J Clin Pathol 2008;130:162–164.

21 Brown FM, Faquin WC, Sun D, *et al.* LSIL biopsies after HSIL smears. Correlation with high-risk HPV and greater risk of HSIL on follow-up. Am J Clin Pathol 1999;11:765–768.

22 Evans MF, Adamson CS, Papillo JL. *et al.* Distribution of human papillomavirus types in ThinPrep Papanicolaou tests classified according to the Bethesda 2001 terminology and correlations with patient age and biopsy outcomes. Cancer 2006;106: 1054–1064.