Follow-up Study of Patients With Cervical Intraepithelial Neoplasia Grade 1 Overexpressing p16<sup>ink4a</sup>

Stefania Cortecchia, PhD,* Giuseppe Galanti, MD,* Cecilia Sgadari, PhD,† Silvano Costa, MD,‡ Margherita De Lillo, PhD,* Licia Caprara, PhD,* Giovanni Barillari, MD,§ Paolo Monini, PhD,† Roberto Nannini, MD,* Barbara Ensoli, PhD,† and Lauro Bucchi, MD||

Objective: The p16<sup>ink4a</sup> (p16) tumor-suppressor protein is a biomarker for activated expression of human papillomavirus oncogenes. However, data are insufficient to determine whether p16 overexpression predicts the risk for progression of low-grade cervical intraepithelial neoplasia (CIN). This study was aimed at evaluating the risk for progression to CIN2 or worse during a 3-year follow-up of an unselected series of 739 patients with CIN1 biopsy specimens tested for p16 expression.

Methods: Positivity of p16 was defined as a diffuse overexpression in the basal/parabasal cell layers. Selection biases were ruled out using a control group of 523 patients with CIN1 biopsies not tested for p16 expression. Analysis was based on the ratio of progression rates.

Results: In the first year of follow-up, the 216 patients (29%) with p16-positive CIN1 had a higher progression rate (12.3%) than did the 523 patients with p16-negative CIN1 (2.2%) (rate ratio, 5.5; 95% confidence interval [CI], 2.59–11.71). In the second and third years, differences were smaller (rate ratio, 1.32 and 1.14, respectively) and not significant. The patients with p16-positive CIN1 also had a lower risk for regression to normal in the first year of follow-up (rate ratio, 0.55; 95% confidence interval, 0.42–0.71) and nonsignificant changes in the second and third years (rate ratio, 0.81 and 0.84, respectively).

Conclusions: The patients with p16-positive CIN1 had an increased risk for progression that was concentrated in the first year of follow-up. Immunostaining of p16 could have a role in short-term surveillance of patients with CIN1. Further research should focus on midterm/long-term outcomes of p16-positive CIN1.

Key Words: p16<sup>ink4a</sup>, Cervical intraepithelial neoplasia, Progression, Regression

Received June 8, 2013.
Accepted for publication August 4, 2013.

(Int J Gynecol Cancer 2013;23: 1663–1669)
CDK4/6, and the affected cells overexpress p16 to counteract the irregular cell cycle activation. Because p16 no longer has any influence on CDK4/6 and cell cycle, it accumulates in the nucleus and the cytoplasm and can be detected by immunostaining.

Most published studies on p16 expression in cytological and histological specimens from the cervix have explored its value in the diagnostic process on the basis of a cross-sectional design. Their objectives were to determine whether the biomarker can increase the accuracy in detecting high-grade cervical intraepithelial neoplasia (CIN) and whether it has a role in the triage of equivocal and mildly abnormal cytology results. At present, p16 immunostaining is recommended in the differential diagnosis between high-grade CIN and processes not related to neoplastic risk, in cases in which there is disagreement in histological interpretation and the differential diagnosis includes a precancerous lesion and in cases in which there is a major discrepancy between cytological and histological findings.

Some small-sized studies have addressed the role of p16 as a marker of the risk for progression of low-grade CIN. The rationale of this second approach is that there is a need to develop techniques capable of identifying women at risk for progressive disease to offer them appropriate surveillance. Current data suggest that p16-overexpressing CIN1 is more likely to progress to CIN2, CIN2 or worse (CIN2+), and CIN3 during varying time periods. However, additional longitudinal clinical studies are needed before definitive conclusions can be drawn.

Here, we report a 3-year follow-up study of a large unselected series of patients with biopsy diagnosis of CIN1, primarily aimed at determining whether p16 overexpression may be a predictor of the risk for progression to CIN2+

MATERIALS AND METHODS

Study Rationale and Design

Except for the development of frankly invasive carcinoma, the progression of CIN1 to more severe lesions is a subclinical event. This has 2 major implications: first, the distinction of an incident (newly developing) CIN from a prevalent one (an already existing disease undiagnosed at baseline) is impossible to make in practice; second, in the absence of spontaneous symptoms for the overwhelming majority of outcome lesions, the standard time-to-event analysis is inappropriate for investigating the natural history of the disease. The cumulative risk for progression would be closely dependent on the frequency and timing of follow-up visits and the related selection factors.

On the basis of these considerations, we designed this study as follows: (1) we used the term progression in a purely clinical sense to refer to the detection of a higher-grade lesion at any date after the diagnosis of CIN1, with no assumptions about whether it was a prevalent or an incident lesion; (2) we divided the follow-up time into annual intervals, that is, into clinically meaningful periods; (3) we computed the progression rate for each interval by dividing the number of actively detected progressions by the number of patients followed up during that interval; and (4) we compared the annual progression rate of patients with p16-positive CIN1 versus p16-negative CIN1.

Regression of CIN1 to normal is most often documented by 1 or more negative Papanicolaou test results without histological confirmation, which makes its level of certainty less than that of progression. For this reason, regression of CIN1 was included into this study as a secondary, although complementary, endpoint.

Study Setting

The study was done using data from the Department of Pathology at the Health Care District Hospital of Imola, Italy (hereby simply referred to as the Department). The Department covers all population-based cervical screening centers and all public and private gynecology clinics in the area. According to agreed local guidelines, follow-up of an untreated patient with CIN1 includes a Papanicolaou test and a colposcopy every 6 to 12 months for 2 to 3 years.

In accordance with the Italian legislation for retrospective observational studies, no ethical approval was sought from the Health Care District review board.

Immunostaining of p16

Immunohistochemical analysis of p16 expression was performed on formalin-fixed, paraffin-embedded tissue sections prepared from CIN1 biopsies using the Clonetics p16INK4a Histology Kit (mtn Laboratories AG [currently Roche mtn Laboratories], Heidelberg, Germany) according to the manufacturer's instructions. Briefly, sections were cut at 3 to 4 mm and deparaffinized. After antigen unmasking for 10 minutes at 95°C to 99°C in citrate buffer (pH 6.0), the slides were cooled for 20 minutes at room temperature (RT), and endogenous peroxidase was blocked by 3% hydrogen peroxide solution for 5 minutes at RT. The slides were first stained for 30 minutes at RT using the primary antibody (mouse antihuman p16 monoclonal antibody, clone E6H4, included in the Clonetics p16INK4a Histology Kit), and then using the visualization reagent also provided in the kit for 30 minutes at RT. After chromogenic visualization using the 3’-diaminobenzidine chromogen, the slides were counterstained with hematoxylin and coverslipped. For each staining run, sections from a biopsy with known positive immunoreactivity were included as a positive control.

Immunostaining of p16 was implemented in 2004, and, until mid-2007, the results were appraised using a scoring method derived from a published study. Subsequently, following the manufacturer's recommendations, a negative/positive classification was adopted. A diffuse immunostaining reaction, defined as a continuous staining of cells in the basal and parabasal layers, with or without staining of superficial squamous cell layers, was considered a positive result.

Study Patient Database

Between 2004 and 2010, the annual proportion of CIN1 specimens not undergoing p16 immunostaining varied between 5% and 20%. To assess whether the composition of
patients with CIN1 biopsy specimens tested for p16 expression (hereby referred to as study group) was biased by selection factors, a group of patients with nontested CIN1 biopsies were used as a control group. The 2 subpopulations were compared for patient age distribution at biopsy and for the annual progression rate.

From the database of the Department (Pathwin system, K2 Informatica, Ravenna, Italy), we identified the patients who had at least 1 histology diagnosis of CIN1 between 2000 and 2010, with or without analysis of p16 expression. Of these, we extracted an anonymous data file containing all cervical cytology and histology records of any date (range, 1982–2010). Records dating before 2000 were used to identify the patients with a history of CIN, a feature associated with long-term persistence of HPV infection and the risk for progression.10 A specifically designed software11 was used to convert the Snomed codes,12 including the nonstandard Snomed-like codes created by the Department, into simplified text descriptions.

When more than 1 diagnosis of CIN1 (2000–2010) was listed in a patient’s records, the first one was selected as the baseline lesion. Patients in whom the diagnosis of CIN1 was preceded by a diagnosis of CIN2+ and patients in whom the diagnosis of CIN1 was made on conization or hysterectomy specimens were excluded.

Cytology and Histology Data

During the study period, cervical cytology diagnoses were reported using the Bethesda System of 1991 and 2001.13,14 The original reports of high-grade squamous intraepithelial lesion (HSIL) and carcinoma were grouped as HSIL+. All diagnosis codes less than atypical squamous cells of undetermined significance were grouped as negative. The histology diagnoses of CIN2, CIN3, adenocarcinoma in situ, and invasive carcinoma were grouped as CIN2+.

Definition of Follow-up, Progression, and Regression

Follow-up time was calculated from the date of diagnosis of CIN1. In each year of follow-up, the patients with at least 1 satisfactory Papanicolaou test or a cervical histology diagnosis were considered to be followed up.9 After conization or hysterectomy, the observation was truncated.

Follow-up data of each patient were reviewed to classify the outcome of CIN1 into progression, regression, and persistence. The result of p16 immunostaining was encrypted. Disease progression was defined as 1 or more consecutive diagnoses of HSIL+ or a histology diagnosis of CIN2+. Regression was defined as 1 or more consecutive negative Papanicolaou test results or a negative histology diagnosis.

When multiple patterns of progression were present, the date of the one with the highest level of certainty (eg, a histology diagnosis of CIN2+ rather than a cytology diagnosis of HSIL+) was taken as the date of progression. The same criterion (eg, 2 consecutive negative Papanicolaou test results rather than 1) was applied to define regression. If a given pattern of progression was present twice or more, the earliest one was considered. When regression was based on 2 or more consecutive negative Papanicolaou test results, the date of the first one was taken as the date of regression. The patients observed in a second dysplastic state after regression were not considered to contribute a second time to the study population.9

Data Quality Assurance

We designed 2 levels of data quality assurance. First, all biopsy specimens undergoing p16 immunostaining in the years 2004–2007 were reclassified according to current criteria for positivity.2,7 The reviewer was blinded to follow-up findings.

Second, the procedures of an important study,15–19 in which the Department was involved during the study years were followed. A random sample of Papanicolaou tests from all participating centers was reviewed by 3 international experts. The sensitivity and the specificity of this panel for CIN2+ differed only marginally from those of original cytoscreeners.17 The CIN specimens collected during the trial were also reviewed by independent pathologists.18 When diagnoses were dichotomized as either CIN1 or CIN2+, the κ for agreement between the Department and the reviewers was 0.69.

Data Analysis

The patient subgroups were compared for proportions using the Fisher exact test, the Pearson χ² test, or the stratified Mantel-Haenszel χ² test, as appropriate. Trends in proportions were evaluated using the χ² test for linear-by-linear association. Group comparisons for continuous variables were performed using the Mann-Whitney U test. All tests were 2-tailed with the α value set at 0.05. According to Gardner and Altman,20 progression and regression rates were compared by the calculation of their ratio and 95% confidence interval (CI).

RESULTS

Patient Characteristics

There were 1262 eligible patients, with a total of 5962 single baseline and follow-up cytology and histology records. These included 739 patients with CIN1 biopsy specimens tested for p16 expression (study group) and 523 patients with nontested CIN1 biopsies (control group). Figure 1 shows that there was no significant age difference between the 2 groups (median, 36 years). Table 1 shows that the patients in the study group had a greater progression rate to CIN2+ in the third year of follow-up. In the years 1 and 2, there was no evidence that they were a selected subset of patients.

In the study group, 216 patients (29.2%) were positive for p16 immunostaining. Figure 2 shows that patient age was inversely associated with the proportion of p16-positive biopsies. By implication, the patients with a p16-positive CIN1 were significantly younger than the patients with a p16-negative CIN1 (Fig. 1).

To further check for selection biases in the study group, we compared the prevalence of history of CIN1 between the patients with p16-positive CIN1 and those with p16-negative
No difference whatsoever was observed (5% versus 6%, $P = 0.44$).

**Follow-up**

Of the total of 1262 eligible patients, 1030 underwent follow-up (Fig. 3). The patients with p16-positive CIN1 were more likely to be followed up, although the difference was no longer significant when patient age was adjusted for ($P = 0.25$).

In the subset of patients followed up, the median time to the first visit was shorter for the 188 patients with p16-positive CIN1 (199 days versus 217 days, $P = 0.001$).

**Progression**

Table 2 shows that the progression rate during the first year of follow-up was 5.5-fold greater among the patients with p16-positive CIN1. In the following 2 years, there was a limited and nonsignificant increase. The increase observed in the first year was nonsignificantly higher among the patients 36 years or older (median age), with a progression rate ratio of 7.26 (95% CI, 2.65–19.94) versus 4.26 (95% CI, 1.37–13.25).

The patients with p16-positive and p16-negative biopsy specimens did not differ significantly in distribution by type of diagnostic evidence supporting the progression (ie, 1 cytology diagnosis of HSIL+, 2 or more consecutive diagnoses of HSIL+, and histology diagnosis of CIN2+) ($P = 0.64$). Four (7%) of the total 56 cases of disease progression were based on the weakest level of evidence (one cytology diagnosis of HSIL+). No case of progression to invasive carcinoma was observed.

**Regression**

The temporal pattern of regression reflected that of progression. The patients with p16-positive CIN1 had a 50% decrease in the regression rate in the first year of follow-up, with smaller changes thereafter (Table 3).

The patients with p16-positive and p16-negative biopsy specimens did not differ significantly in the level of evidence supporting regression (1 negative Papanicolaou test result,
DISCUSSION

Overexpression of p16 led to a 5.5-fold increase in the progression rate during the first year of follow-up, coupled with a lower regression rate, and to a nonsignificant increase during the second and third years. The increase was earlier and 2-fold larger than in an important study of HPV-positive patients free of CIN, which associated p16 positivity with a relative risk of 2.6 for developing CIN2+ within 3 years. In that study, as in our own, the risk increase was non-significantly larger for patients older than 35 years.

Assuming that the observed excess rate of progression to CIN2+ was related to p16 overexpression, the data would indicate that the interval between the detection of a p16-overexpressing CIN1 and the related progression is less than 1 year. By implication, this would suggest that progression has a short latency and that p16 overexpression is thus a late event in progression. Apart from anecdotal exceptions, however, the literature indicates the opposite, that is, that p16 overexpression occurs in an early phase and that further molecular events are needed to fully develop the potential for progression. Accordingly, there are data suggesting that p16 overexpression may precede by years the development of high-grade CIN.6,23 Besides, our observation of an inverse association between patient age and the proportion of p16-positive biopsies is compatible with the view that p16 positivity, which reflects a persistent high-risk HPV infection, is only the prelude to a chain of events that may subsequently lead to a higher grade of atypia.24

An alternative explanation for the high rate of CIN2+ in the first year of follow-up is that it is accounted for by prevalent lesions. There is evidence that even systematic multimodality evaluations by experts can miss a prevalent disease,7,9 such as an area of CIN2+ with indistinct colposcopic features, a colposcopically occult focus of CIN2+ already developing within the baseline CIN1 or in a different ectocervical site, a focus of CIN2+ located in the endocervical canal and difficult to reach with the biopsy device, and a CIN2+ downgraded by the reporting pathologist.

Unfortunately, except for the latter type, these prevalent conditions cannot be identified by reviewing baseline and follow-up histological specimens. However, the hypothesis that they correlate with p16 overexpression is plausible based on current knowledge and is compatible with the fact that the biomarker increases the accuracy of diagnosis.7,25 Besides, it is conceivable that the first follow-up for a p16-positive CIN1 is conducted with greater care, allowing the detection of prevalent lesions missed at baseline. In our data, for example, the patients with p16-positive CIN1 were more often

<table>
<thead>
<tr>
<th>Year of Follow-up</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Followed up, n</td>
<td>422</td>
<td>188</td>
<td>610</td>
</tr>
<tr>
<td>Total</td>
<td>523</td>
<td>216</td>
<td>799</td>
</tr>
<tr>
<td>Lost to follow-up</td>
<td>101</td>
<td>38</td>
<td>129</td>
</tr>
<tr>
<td>Follow-up</td>
<td>420</td>
<td>229</td>
<td>103</td>
</tr>
<tr>
<td>Total</td>
<td>523</td>
<td>254</td>
<td>1262</td>
</tr>
</tbody>
</table>

A positive p16 expression was defined as a diffuse immunostaining reaction.
followed up, and the median time to their first appointment was shorter.

Some methodological points need to be made. First, the quality of data in this study was ensured in several ways: the immunostaining of early specimens was reevaluated, the decoding of Snomed codes was made using a software designed for the Department’s data, and the accuracy of cytology and histology diagnoses was documented by independent reviewers and all checks for selection biases yielded negative results except 1. The study and the control groups were comparable for age distribution, and in the study group, the patients with p16-positive and p16-negative CIN1 were comparable for the prevalence of history of disease. More importantly, the progression rate in the first and the second year of follow-up was similar between the study and the control groups. In the third year, conversely, progression was more frequent in the study group. Although the CI around the rate ratio was very large, it was suggested that 1 or more risk factors for high-grade CIN (possibly including colposcopic findings) acted as selection factors for patient attendance.

Second, our work can be considered a population-based study. The Department covers virtually the entire population of a large area, which explains the high number of patients (n = 739) in this single-institution study. To our knowledge, the largest follow-up study to date included 138 patients with CIN1.8 The mean proportion of p16-positive biopsies we observed, 29%, was smaller than the 38% reported from a pooled literature analysis.3 This was due to the interaction of 2 factors: the inverse association of patient age with p16 positivity and the higher patient age in our case series compared with studies from selected clinical settings.

Third, we have pointed out why time-to-event analysis is not appropriate for this type of data. Previous studies have reported progression rates as high as 40% to 60% for p16-positive CIN15–8 and 15% to 30% for p16-negative CIN1.5,6 All of these percentages are greater than the mean ones reported for CIN1 not otherwise specified,26 which confirms that follow-up studies for this lesion are prone to biases.7,9 Our finding of a mean rate of 5% to 8% for total study group in the first 2 years of follow-up is more plausible in light of major literature data.

The short length of follow-up and the rapid decrease in the proportion of attending patients were major limitations of the current study. In Italy, the recommended duration of colposcopy and cytology surveillance for a diagnosis of CIN1 does not exceed 3 years, after which the patients return to the normal screening of every 3 years. In addition, follow-up failures were more frequent than expected. A sense of reassurance after a negative result on the first visit and the lack of fail-safe systems in colposcopy clinics were the most likely causes. Another factor of importance was p16 negativity, which was associated with a higher frequency of losses to follow-up.

Because of follow-up failures, results for the second and third years require caution in interpretation because the observed nonsignificant excess risk for progression might be a false-negative result. Moreover, our data do not exclude an increased progression rate in the midterm/long-term. Because routine follow-up schedules are generally not compatible with the correct conduct of observational studies, a prospective study of sufficient duration and with active follow-up is needed to clarify the midterm/long-term outcome of p16-positive CIN1. With a more complete follow-up for the first 3 years, it might also be possible to conclude that patients with p16-negative CIN1 can be returned to the normal screening interval because their risk for progression is quite low—or at least as low as to require less intensive follow-up.8

In conclusion, the patients with p16-positive CIN1 had an increased risk for CIN2+ that was concentrated in the first year of follow-up. Although this finding will have to be confirmed at other institutions, it suggests that p16 immunostaining could have a role in the planning of short-term surveillance of patients with CIN1. Further studies are also needed to assess whether this biomarker can predict midterm/long-term outcomes of the disease.

ACKNOWLEDGMENTS

The authors thank Carlo Pierri (K2 Informatica, Ravenna, Italy) for his invaluable computer assistance and Arrigo Bondi (Unità Operativa di Anatomia Patologica, Ospedale Maggiore, Bologna, Italy) for his encouragement.

**TABLE 3.** Regression to normal of patients with p16-positive CIN1 versus p16-negative CIN1

<table>
<thead>
<tr>
<th>Patients</th>
<th>Year of Follow-up</th>
</tr>
</thead>
<tbody>
<tr>
<td>Followed up, n</td>
<td>1</td>
</tr>
<tr>
<td>With p16-positive CIN1</td>
<td>179</td>
</tr>
<tr>
<td>With p16-negative CIN1</td>
<td>403</td>
</tr>
<tr>
<td>Regressing to normal, n (%) rate</td>
<td></td>
</tr>
<tr>
<td>With p16-positive CIN1</td>
<td>47 (26.3)</td>
</tr>
<tr>
<td>With p16-negative CIN1</td>
<td>194 (48.1)</td>
</tr>
<tr>
<td>Rate ratio of p16-positive to p16-negative CIN1</td>
<td>0.55</td>
</tr>
<tr>
<td>95% CI</td>
<td>0.42–0.71</td>
</tr>
</tbody>
</table>

A positive p16 expression was defined as a diffuse immunostaining reaction.
REFERENCES