Ki67 predicts progression in early CIN:
Validation of a multivariate progression-risk model

Arnold-Jan Kruse a, Jan P.A. Baak a,b,∗, Emiel A. Janssen a, Kjell-Henning Kjellevold a, Bent Fiane c, Kjell Lovslett c, Johan Bergh d and Stanley Robboy e

a Department of Pathology, Rogaland Central Hospital (=SiR), Stavanger, Norway and Free University, Norway
b Armauer Hansenvej 20, N-4068 Stavanger, Norway
c Department of Gynecology, Rogaland Central Hospital (=SiR), Stavanger, Norway and Free University, Norway
d Madla Gynecology, Amsterdam, The Netherlands
e Department of Pathology, Duke University Medical Center, Durham, North Carolina, USA

Received 18 June 2003
Accepted 21 August 2003

Abstract. This study of early CIN biopsies (25 CIN1 and 65 CIN2) with long follow-up was done to validate, in a new group of patients, the value of Ki67 immuno-quantitative features to predict high CIN grade in a follow-up biopsy (often denoted to as “progression”), as described in a previous study. Each biopsy in the present study was classified with the previously described Ki67-model (consisting of the stratification index and the % positive nuclei in the middle third layer of the epithelium) as “low-risk” or “high-risk”, and matched with the follow-up outcome (progression-or-not). Furthermore, it was studied whether subjective evaluation of the Ki67 sections by experienced pathologists, who were aware of the prognostic quantitative Ki67 features, could also predict the outcome. Thirdly, the reproducibility of routine use of the quantitative Ki67-model was assessed.

Fifteen cases progressed (17%) to CIN3, 2/25 CIN1 (8%) and 13/65 CIN2 (20%), indicating that CIN grade (as CIN1 or CIN2) is prognostic and that the percentage of CIN1 and CIN2 cases with progression in the present study is comparable to many previous studies. However, the quantitative Ki67 model had stronger prognostic value than CIN grade as none of the 40 “Ki67-model low-risk” patients progressed, in contrast to 15 (30%) of the 50 “Ki67-model high-risk” patients (p < 0.001). In multivariate analysis, neither CIN grade nor any of the other quantitative Ki67 features added to the abovementioned prognostic Ki67-model. Subjective analysis of the Ki67 features was also prognostic, although quantitative assessments gave better results. Routine application of the quantitative Ki67-model in CIN1 and CIN2 was well reproducible. In conclusion, the results confirm that quantitative Ki67 features have strong prognostic value for progression in early CIN lesions.

Keywords: Cervix, CIN, Ki67, image analysis, progression

1. Introduction

Increasing cervical intraepithelial neoplasia (CIN) grade is well known to correlate with increased progression-risk to higher CIN grade and/or invasive carcinoma, as many studies have shown (for an overview, see [19]). However, reproducibility between pathologists concerning CIN lesions is very low [20]. Over- or under treatment can be the result. Furthermore, in early CIN lesions, progression rates to higher-grade lesions or even cancer are low (from CIN1/2 to CIN3: 9% and 22%, and to invasive cancer 1% and 5%, respectively). In spite of these low risk rates, frequent cytological follow-up (in CIN1) and large biopsies (in CIN2) are often performed in women with these lesions. This means over treatment and unnecessary controls in the follow-up of the patients without progression (which are the majority of all cases) and thus an ineffective burden for the health-care system. To im-
prove this situation, more accurate additional predictors of the discourse of early CINs are needed. There has been doubt whether CIN1 lesions ever progress [11]. According to one theory, the CIN3 in the follow-up of a marker biopsy with CIN1 and CIN2 always has been there and was not sampled with the first biopsy. Thus, the finding of the CIN3 is a matter of co-existence, not true progression. This theory is also supported by the finding that CIN3 lesions are located more proximally in the cervix than CIN1 lesions and therefore may easily be missed in the biopsy [23]. According to another theory, CIN1 and CIN3 are two distinct manifestations of HPV infection with different natural histories. In support of this theory there are studies showing that most women who developed CIN3 never had evidence of a preceding CIN1 [13,25]. It is difficult to prove or disprove the progression or the co-existence hypotheses but for gynecologists it is therapeutically not very important which theory is adhered to. In both theoretical models, the detection of an early CIN lesion with accurate high-risk prognostic features requires from the gynecologist to perform careful repeated cytological or early repeated biopsy follow-up and pathologists should take more sections of the biopsy under study to exclude sampling errors. This explains the interest in high-risk indicators in early CIN lesions.

Previous studies have shown that application of Ki67 immunoquantitative analyses of CIN1 and CIN2 in histological biopsies has strong independent predictive value for grade, presence of oncogenic HPV and progression [3,14,15,17]. The best Ki67-feature combination to predict whether a subsequent higher CIN grade or cancer will be detected in the follow-up, is the 90th percentile of the stratification index (Si90) and the percentage of Ki67 positive cells in the middle third layer of the epithelium (MIDTHIRD) [17]. The Ki67 features are easily assessed and inexpensive; their prognostic value exceeds that of CIN grade (as CIN1 or CIN2) and presence of high-risk HPV types assessed with PCR [17]. This could be clinically important. However, following Good Laboratory Practice standards [1], such predictive results must be validated in an independent test set of new patients. This is the purpose of the current study. We also analyzed whether pathologists aware of the prognostic value of the quantitative Ki67 features, could predict progression by subjective evaluation of these Ki67 features. Thirdly, the reproducibility of the quantitative Ki67 prognostic model determinations was studied.

2. Materials and methods

2.1. Patients and follow up

Originally there were 157 consecutive small (<5 millimeters maximal diameter) biopsies available, classified as CIN1 and CIN2 at the Department of Pathology, Rogaland Central Hospital (=SiR), Stavanger, Norway. The biopsies were taken because of abnormal cervical smears. The tissues were fixed in buffered 4% formaldehyde, embedded in paraffin, cut at 4 micrometer and stained with haematoxylin–eosin (H&E). 47 cases were excluded because of short follow-up (within 4 months, average 1.1 months). Six patients were lost to follow-up and 14 more were excluded, as Ki67 immuno-quantitation was not possible for technical reasons. The remaining 90 biopsies were unanimously classified at careful review by two experienced gynecological pathologists as CIN1 (n = 25) and CIN2 (n = 65) (the percentage CIN1 cases was relatively low, as contrasting most CIN2s the majority of CIN1s did not have a histological biopsy as required for this study but a cytological smear in the follow-up). The diagnosis CIN was confirmed by the presence of Ki67 cell clusters [21] and p16 positivity [12] in all cases and special attention was paid that none of the cases was a metaplastic lesion [16,21]. Relying on the guidelines of the Norwegian Medical Society, treatment for CIN1 is colposcopic control with cytological or histological follow-up (the latter is much less frequently applied). Cone biopsy is used to treat CIN2 [6]. Progression was defined as a histological increase of (review) grade (which in all cases was to CIN3), n = 15 = 17%, a progression percentage in agreement with a large meta-analysis [19]; a negative cytological or histological follow-up was regarded as “non-progression” (all other cases, n = 75).

2.2. Immunohistochemistry and image analysis

Paraffin sections 4 μm thick, taken adjacent to the H&E sections used for the CIN grade assessment, were analyzed for the nuclear Ki67 proliferation antigen [14]. The serial H&E section immediately following the Ki67 section was used to confirm the CIN. The QPRODIT (version 6.1) interactive image analysis system (Leica, Cambridge, UK) was used for the measurements [14]. Briefly, in each case, the lumen and basal membrane of non-tangentially cut, most severely dysplastic epithelium was marked electronically. Subsequently, starting at the left margin and moving to the
Fig. 1. Illustration of the image analysis method. The microscopic image of the cervical epithelium is shown on the monitor of the image analysis system. The operator demarcates with the mouse a diagnostic epithelium strip, carefully avoiding tangentially cut areas. The demarcation lines are shown as black lines (luminal surface, basal membrane, left, right). Within the demarcated strip, the operator then clicks with the mouse on all Ki67 positive nuclei. After each click, the system automatically draws a perpendicular line from that point to the basal membrane and over the full thickness of the epithelium (these thin dotted lines are barely visible for all nuclei) and calculates several quantitative features such as thickness = $T$ of the epithelium at that point, distance = $D$ of the point to the basal membrane, the stratification index $= \frac{D}{T}$, and others. In the figure, points N are marked Ki67 positive nuclei with $D = Y-Y'$ and $T = X-X'$. This procedure is further illustrated for two other Ki67 positive nuclei. The Si's are 0.80, 0.19 and 0.67 respectively. This procedure is further illustrated for three Ki67 positive nuclei. The Sis are 0.80, 0.19 and 0.67 respectively.

right border of the thus demarcated area, all the Ki67 positive nuclei were marked interactively by the investigator using the system cursor. Figure 1 illustrates this. The QPRODIT system then automatically calculated multiple quantitative features (including numerous descriptive statistics for each feature) for each case: distance between the nucleus and basal membrane ($= DBM$), epithelial thickness ($= T$) at the location of the nucleus indicated, distance between the nucleus and the lumen ($= DL$), stratification index (SI) ($= DBM/T$, where DBM is the distance between the nucleus and basal membrane, divided by the epithelial thickness $T$), total number of Ki67 positive nuclei per 100 µm basal membrane and percentage Ki67 positive nuclei in the deep third, the middle third, and the upper third layer of the epithelium. Each case was classified as “low-risk” or “high-risk” using the Ki67 features, as follows: [combined $Si_90 < 0.57$ and $MIDTHIRD < 30$] = low-risk; all others = high-risk (further denoted to as Ki67 progression-risk model or Ki67-model). Figure 2 gives a graphic illustration of the model.

2.3. Statistical analysis

SPSS version 10 (SPSS Inc., Chicago, USA) was used to analyze the prognostic value of CIN grade, the Ki67-model and the other (quantitative and subjectively assessed) Ki67 features. Progression-or-not...
in the follow-up was the independent variable. Cross tables were made to compare progression with CIN grade and the Ki67 progression-risk model. The Mann–Whitney U-test compared differences in the features between the progression and non-progression cases. Univariate survival analysis (Kaplan–Meier) was performed to evaluate the prognostic value of each of the features studied. Survival curves were evaluated using the log-rank test (LR). Multivariate comparison of all the features was performed with the Cox model. Sensitivity, specificity, positive and negative predictive values were also calculated. To assess the predictive value of subjective Ki67 feature analysis, Ki67 sections were independently subjectively evaluated by two of us (J.B., K.K.) who were aware of the prognostic value of the quantitative Ki67 features). Reproducibility of duplicate routine quantitative assessments of the Ki67 model and its constituting features, by two independent observers was studied in 25 consecutive cases with cross-tables and the Pearson’s test.

3. Results

Fifteen cases (17%) progressed (all to CIN3), 2/25 CIN1 (8%) and 13/65 CIN2 (20%). The mean follow-up of the non-progression and progression cases was 3.4 years (range 0.4–5.6) and 1.0 year (range 0.4–3.3) (Table 1). Age was not significantly different ($p = 0.12$). Nearly all Ki67 immunoquantitative features differed significantly between the cases with and without progression, but as in a previous study [17], Si90 was the strongest prognostic feature followed by MIDTHIRD (Fig. 2) (although Si90 plus UPTHIRD was nearly as strong). Among the 40 “Ki67-model low-risk” cases (40 of 90 cases, 44%), none did progress (negative predictive value = 100%) (Table 2). In contrast, 15 of the 50 “Ki67-model high-risk” cases progressed (positive predictive value = 30%).

Table 3 summarizes the sensitivity, specificity, positive and negative predictive values and percentage overall correct classified cases with the subjective CIN grade, the Ki67-model and combined CIN grade-Ki67; they are highest for the quantitative Ki67 model (consisting of Si90 and MIDTHIRD, see Fig. 2). Survival analysis shows that the combination of MIDTHIRD and Si90 was highly significant (Fig. 3, LR = 13.1, $p < 0.001$). Combining the Ki67-model with the other Ki67 features or subjective CIN grade did not improve prediction (Cox regression, see also Table 3). Subjective estimations of the Ki67 features were also prognostic, but just missed statistical significance ($p = 0.06$) (Table 4). Reproducibility among two independent observers of the Ki67-model and, also its individual constituting features was good (overall classification agreement of the Ki67 model as low-risk versus high-risk = 100%, correlations coefficients of Si90 = 0.99, $p < 0.0001$ and of MIDTHIRD = 0.98, $p < 0.0001$).

4. Discussion

The present study confirms that the prognostic value of the Ki67 progression-risk model exceeds the value of histopathological CIN grade (as CIN1 or CIN2) to predict progression to higher CIN grade. While other Ki67 features were also prognostically significant, their utility was negated in multivariate analysis. Intra- and inter-observer reproducibility of the essential individual prognostic quantitative Ki67 features (MIDTHIRD and Si90), and of the Ki67-model showed excellent results (100% agreement) indicating it is robust.

One may argue that the evaluation of the progression rate in CIN2 is very uncertain due to the very low reproducibility between pathologists concerning separating CIN2 and CIN3. However, most studies on pathologists’s intra- and interobserver reproducibility [8–10,18,22,24] have concluded that there is good agreement when it comes to diagnosing CIN3, but that the reproducibility of CIN1 and CIN2 is poor. It is of interest that in all these studies the reproducibility of CIN2 has been closer to CIN1 than CIN3, the CIN subgroup with which it is usually proposed that it should be linked [5]. Furthermore, in the present study consensus cases were used for the evaluation, with the goal to minimize a certain lack of reproducibility. In a previous study we evaluated the value of Ki67 immunoquantification for progression prediction for both consensus cases and routinely diagnosed cases. Results were similar for both groups. Furthermore, progression rates in this study were 8% of CIN1 and 20% of CIN2 cases, in agreement with most other studies [19]. This argues that the patients in this study are representative of most other patients with CIN lesions. Thus, Ki67 immunoquantitation can be regarded as a useful prognostic adjunct.

We cannot be sure whether finding a higher CIN grade will occur beyond the 5-year follow-up of the
that HPV infection occurs rapidly; most often within early CIN lesions. Moreover, a recent article indicates this study are representative of most other patients with many other studies [19]. This argues that the patients in 8% of CIN1 and 20% of CIN2 cases, in agreement with ing reasons. The progression rates in this study were current study, but the likelihood is low, for the following reasons. The progression rates in this study were 8% of CIN1 and 20% of CIN2 cases, in agreement with many other studies [19]. This argues that the patients in this study are representative of most other patients with early CIN lesions. Moreover, a recent article indicates that HPV infection occurs rapidly; most often within 3 months of inoculation [4] and that the risk of normal, HPV-negative epithelium developing high-grade CIN on average is 6–12 months after the first positive onco-HPV test [25]. In agreement, progression beyond 14 months was a rare event in our patients. Of course, re-infection and development of new lesions cannot be predicted with the current methods.

Table 1
Features assessing probability of progression

<table>
<thead>
<tr>
<th>Feature</th>
<th>Follow-up status</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Progression (n = 15)</td>
<td>Non-progression (n = 75)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mean (SD)</td>
<td>Mean (SD)</td>
<td>P-value*</td>
</tr>
<tr>
<td>Follow-up time (years)</td>
<td>1.0 ± 0.8</td>
<td>3.4 ± 1.3</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Age (years)</td>
<td>39.7 ± 9.8</td>
<td>35.6 ± 9.0</td>
<td>0.12</td>
</tr>
<tr>
<td>Distance nucleus to basal membrane (µm)</td>
<td>64.5 ± 23.2</td>
<td>47.0 ± 24.0</td>
<td>0.01</td>
</tr>
<tr>
<td>Stratification index</td>
<td>0.40 ± 0.05</td>
<td>0.27 ± 0.12</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Thickness of the epithelium (µm)</td>
<td>168.2 ± 62.3</td>
<td>181.9 ± 73.2</td>
<td>0.46</td>
</tr>
<tr>
<td>Distance nucleus to epithelial surface (µm)</td>
<td>103.7 ± 41.7</td>
<td>135.7 ± 65.1</td>
<td>0.09</td>
</tr>
<tr>
<td>Density of positive nuclei per 100 µm</td>
<td>35.9 ± 27.3</td>
<td>22.3 ± 13.6</td>
<td>0.13</td>
</tr>
<tr>
<td>Percentage positive nuclei in the basal membrane</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Percentage positive nuclei in the deep third layer of the epithelium</td>
<td>46.5 ± 9.5</td>
<td>68.1 ± 22.8</td>
<td>0.001</td>
</tr>
<tr>
<td>Percentage positive nuclei in the middle third layer of the epithelium (MIDTHIRD)</td>
<td>36.5 ± 8.3</td>
<td>24.7 ± 16.5</td>
<td>0.01</td>
</tr>
<tr>
<td>Percentage positive nuclei in the upper third layer of the epithelium (UPTHIRD)</td>
<td>17.1 ± 7.3</td>
<td>7.2 ± 8.6</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>90th percentile of the stratification index (Si90)</td>
<td>0.74 ± 0.09</td>
<td>0.51 ± 0.21</td>
<td>&lt; 0.001</td>
</tr>
</tbody>
</table>

*P determined by Mann–Whitney U test.

Table 2
Comparison of disease progression in patients with low and high grade Ki67 quantitative risk factors

<table>
<thead>
<tr>
<th></th>
<th>CIN1 Ki67-model low-risk n (%</th>
<th>CIN1 Ki67-model high-risk n (%)</th>
<th>CIN2 Ki67-model low-risk n (%)</th>
<th>CIN2 Ki67-model high-risk n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Progression</td>
<td>0 (0%)</td>
<td>2 (40%)</td>
<td>0 (0%)</td>
<td>13 (29%)</td>
</tr>
<tr>
<td>Non-progression</td>
<td>21 (100%)</td>
<td>3 (60%)</td>
<td>19 (100%)</td>
<td>32 (71%)</td>
</tr>
<tr>
<td>Total</td>
<td>21 (23%)</td>
<td>5 (6%)</td>
<td>19 (21%)</td>
<td>45 (50%)</td>
</tr>
</tbody>
</table>

Sensitivity CIN2 versus CIN1 = 13/(2 + 13) = 87%.
Specificity CIN1 versus CIN2 = (24/(24 + 51) = 32%.
Positive predictive value CIN1 versus CIN2 = 13/(51 + 13) = 20%.
Negative predictive value CIN1 versus CIN2 = 24/(24 + 2) = 92%.
Overall agreement CIN = (24 + 13)/90 = 41%.

Sensitivity Ki67 high-risk versus low-risk = (13 + 2)/(15) = 100%.
Specificity Ki67 high-risk versus low-risk = (21 + 19)/(75) = 53%.
Positive predictive value Ki67 high-risk = (2 + 13)/(5 + 45) = 30%.
Negative predictive value Ki67 low-risk = (21 + 19)/(21 + 19) = 100%.
Overall agreement quantitative Ki67 model = (40 + 15)/90 = 61%.
Table 3
Quantitative measurement of Ki67 alone is superior to Ki67 combined with CIN grade

<table>
<thead>
<tr>
<th></th>
<th>CIN alone</th>
<th>Ki67 QP alone</th>
<th>Combined Ki67 QP &amp; CIN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sensitivity</td>
<td>87</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Specificity</td>
<td>32</td>
<td>53</td>
<td>28</td>
</tr>
<tr>
<td>Positive predictive value</td>
<td>20</td>
<td>30</td>
<td>22</td>
</tr>
<tr>
<td>Negative predictive value</td>
<td>92</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Overall correct classification</td>
<td>41</td>
<td>61</td>
<td>40</td>
</tr>
</tbody>
</table>

The results of this study show also, in agreement with a previous study [17], that subjective evaluation of Ki67 is also prognostic, but Ki67 quantitative image analysis is required to obtain the best (i.e., most accurate) prognostic results. Application of these methods is practically possible, as standard Ki67 paraffin sections can be used and the quantitative image analysis equipment is simple to use and economic. We use quantitative pathology as a routine technique in our laboratory, with around 1500 determinations per year (including early CIN Ki67 immuno-quantitation and have calculated previously that the costs per case for measurement time by a technician, equipment depreciation, pathologist supervising time, and required space should total maximally US $50 [2], which is cost-effective. Technician training takes maximally 2 months, 2 hours per day, and supervision of the technicians by the pathologist plus quality control time after the training period is approximately 30 minutes per week (in our surgical pathology practice of around 1500 QP determinations per year on 21,000 histological specimens in total).

Should all CIN biopsies be stained with Ki67? Obviously not clear CIN3 lesions, which may form the majority of all CIN lesions in many pathology practices. In case of CIN1 and 2, the decision for additional Ki67 investigation will depend on the pathology laboratory and the quality of the gynecological follow-up system. Ki67 and p16 staining can be done on early CIN lesions, also to exclude reactive lesions, which are often over-diagnosed. Both Ki67 cell clusters and p16 positivity have proven to be useful diagnostic adjuncts of CIN, and thus can improve the certainty and quality of routine early CIN diagnosis [12,16,21]. The costs are limited, certainly compared to the serious consequences of over-diagnosis and doctor’s delay. Additional Ki67 quantitation is inexpensive, as shown above.

In spite of the positive findings, Ki67 analysis has certain shortcomings. First, staining is not always informative, e.g., in small cut-through or tangentially cut lesions. Secondly, although the prognostic accuracy of Ki67 quantitation exceeds the value of histopathological CIN grade (as CIN1 or CIN2) to predict progression to higher CIN grade, our method is far from perfect in predicting the outcome. DNA-image cytometry (ICM) can potentially also be used to predict in cytological low-CIN the occurrence in follow-up cytology of high-grade and malignant CIN lesions [7]. However, DNA-ICM requires more specialized facilities than Ki67 quantitation in histological sections. Ad-
ditional studies to refine the prognostic value are thus indicated. Prospective validation in a routine setting is also important.

In summary, the quantitative Ki67 progression-risk model has strong predictive value in CIN1 and 2 lesions and is well reproducible.

Acknowledgements

This study was supported in part by grant 02/214 of the SBDM. We thank Ivar Skaland, Bianca van Diermen and Silje Lyse for their skilled technical assistance.

References


Submit your manuscripts at http://www.hindawi.com