The relationship between syntactic structure analysis features, histological grade and high-risk HPV DNA in cervical intraepithelial neoplasia

Arnold-Jan Kruse a, Susanne Buhr-Wildhagen a, Emiel A. Janssen a and Jan P.A. Baak a,b,*

a Department of Pathology, Rogaland Central Hospital (SiR), Stavanger, Norway
b Free University, Amsterdam, The Netherlands

Received 10 October 2003
Accepted 16 April 2004

Abstract. Aim: To assess the correlation between syntactic structure analysis (SSA) features, revised dysplasia grade and the presence of high-risk human papillomavirus DNA in cervical intraepithelial neoplasia (CIN). Materials and methods: HPV polymerase chain reaction (PCR) was assessed in 101 consecutive biopsies and consensus in CIN grade between the experts occurred in 88 cases (CIN1 = 16, CIN2 = 27, CIN3 = 45). SSA was performed in the diagnostic histological section of the CIN lesions in these patients and SSA features were compared with the blind review CIN grade, and presence/absence of high-risk HPV DNA. Results: One of the SSA features (points from which the surrounding surfaces has 4 edges, PECO-4) was significantly different between all three consensus CIN grades. Many more features revealed significant differences between CIN1 and CIN2 or between CIN2 and CIN3 cases. With stepwise discriminant analysis, the best multivariate combination of features to distinguish the different CIN grades were the Maximum MST Line Length (MML) and the Area Disorder. Crude overall classification of the consensus grades with these features was 69%. The MML and the Area Disorder is also the best combination to distinguish cases with and without high-risk HPV DNA (77% correct classifications). Conclusions: SSA features are correlated with both CIN grade and presence of high-risk HPV DNA, but the discrimination power is not good enough to be used as a routine method for quality control of subjective grade or as a surrogate marker for high-risk HPV DNA presence.

Keywords: Cervix, dysplasia, CIN, HPV, SSA, differentiation

1. Introduction

Cervical intraepithelial neoplasia (CIN) is a precancerous lesion of cervical cancer and is histopathologically classified into three grades (CIN1, CIN2 and CIN3). Increasing CIN grade roughly corresponds to increasing progression risk to invasive carcinoma of the uterine cervix [7]. Therefore, CIN grade is often used as a guide in the management of patients with CIN lesions. However, many CINs spontaneously regress, while a small subset of early CINs show high-grade lesions and invasive cancer in their follow-up [7]. Furthermore, benign and reactive conditions may be mistaken for CIN and pathologist’s in intra- and interobserver reproducibility in the grading of cervical lesions is not perfect [3,8,14]. Finally, many CIN3s regress spontaneously [7]. Therefore, additional methods are required to improve reproducibility of CIN grade and perhaps also for the identification of biologic favorable unfavourable CIN cases. Such methods should be fast and easy to apply. Several studies have indicated that the development of CIN is preceded by human papillomavirus (HPV) infection of the normal epithelial cells of the uterine cervix [11,16]. Moreover, an association exists between persistent infections with high-risk HPV types and the progression of premalig-
nant cervical lesions to CIN3 or invasive carcinoma [9, 12]. One could thus say that most, if not all CIN3s are positive for high-risk HPV.

The importance of high-risk HPV in CIN is therefore obvious. However, certain CIN1s are also high-risk HPV positive and most of these never progress to CIN3, whereas CIN2s can be negative (these often harbour non-high-risk HPV). Thus, high-risk HPV alone cannot be used as an exclusive diagnostic criterion for low- and high-grade CIN, nor is koilocytosis a good marker for low and high grade CIN. Mitotic activity has a classical role in classifying CIN lesions as high grade [17], which fits with the role of HPV. However, next to proliferation, cellular differentiation is also disturbed in CIN and is widely used as a diagnostic criteria for grading of the dysplasia [17]. Consequently, in standard histological sections decreased maturation of nuclei and cytoplasm in the suprabasal and superficial cell layers in CIN is a widely used dysplasia grading criterion. At the protein level, loss of cellular differentiation in CIN is evident from the up-and down regulation of certain cytokeratins (CKs) in high grade but not in low-grade CINs [13].

It is therefore important to evaluate cellular differentiation for CIN grading, but one difficulty is that the assessment in standard histological sections is fairly intuitive and not well reproducible. This may be improved by using syntactic structure analysis (SSA), a technique based on the graph theory and which may be useful for this purpose. A detailed discussion of the SSA technique can be found elsewhere [6]. Advantages of SSA are that it is very time-effective (a typical determination takes only 2–5 minutes) and can be applied to standard histological sections. In breast cancer, SSA-derived features are related to tumor differentiation [2]. In colonic polyps, different histologic patterns produce different graphs and SSA features may be predictive of metachronous cancer [5]. In CIN, SSA features may be helpful to distinguish between different CIN grades, but the relationship between SSA and the presence of high-risk HPV DNA is unknown. Therefore, SSA was evaluated for grading support in CIN and we have also investigated whether SSA features are correlated with the presence of high-risk HPV DNA.

2. Materials and methods

2.1. Patients

Specimens of 101 consecutive cervical biopsies, originally classified by 6 different pathologists as CIN were used. The tissues had been fixed in buffered 4% formaldehyde, embedded in paraffin, cut at 4 micrometer thickness and stained with haematoxylin–eosin (H&E). At blind review and using well established criteria, there was complete consensus between two experienced gynecopathologists on the CIN grade in 88 cases (87%), including CIN1 (n = 16), CIN2 (n = 27), and CIN3 (n = 45) (Kappa 0.98; p < 0.001) [17].

2.2. Detection of high-risk HPV

Five paraffin sections of 5-micrometer thickness following the diagnostic H&E section were used for the detection of high-risk HPV as described below. The presence of the CIN lesion in these consecutive sections was confirmed by review of a serial H&E section immediately following these 5 “HPV” sections. Testing for HPV was done by the enzyme-immunoassay polymerase chain reaction (EIA PCR), according to published protocols [18]. Briefly, cervical biopsies were first pre-screened by a beta globulin PCR using the primer combination PCO3 and PCO3 to check the quality of the target DNA [10]. All 88 cervical biopsies showed intact beta-globulin sequences and were subjected to GP5+/βiG6+ PCR. GP5+/βiG6+ PCR products derived from the beta-globulin PCR-positive cervical biopsies were analysed in the EIA. One assay was used to test for 14 high-risk types of HPV (16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, 66 and 68).

2.3. Syntactic structure analysis

The QPRODIT (version 6.1) interactive image analysis system (Leica, Cambridge, UK) was used for the measurements. In each case the most severely dysplastic part of the epithelium in the H&E section used for the routinely assessed diagnosis was selected. In this area, the upper border, the basal membrane, and the borders between the two deepest layers of the epithelium were marked (the deepest layer excluding the basal cell layer) (see Fig. 1). SSA was performed in fields of vision chosen randomly in the said demarcated area by constructing a Voronoi diagram (VD) and a minimum spanning tree (MST) per field. Briefly, with a ×40 objective (final magnification ×1200) the centres of gravity of at least 75 nuclei in the lower deep half of the epithelium per field of vision were interactively marked on a video screen by setting a point with the cursor of the system. This sample size minimizes the influence of boundary effects [6]. Using this set of points, the computer program composed the VD
Fig. 1. Diagram illustrating the different layers in which the measurements were performed.

and the MST. The Voronoi tessellation splits the image plane into polygons, each containing one nuclear centre of gravity in its centre and in such a way that each point in the polygon is closer to this centre nucleus than to any other point in the plane. The MST interconnected a set of nuclei in such a way that the total length of the lines was minimal and no loops were formed. Figure 2 illustrates this image analysis method. The following SSA features were derived from the VD:

1. total of the points clicked on with the cursor,
2. the sum of all polygons surfaces,
3. average surface of the polygons,
4. minimum surface of the polygons,
5. maximum surface of the polygons,
6. standard deviation of the polygons,
7. average of all roundness factors,
8. standard deviation of all roundness factors,
9. area disorder,
10. the number and percentage of selected points from which the surrounding surface has n edges,
11. the sum of all distances between points clicked on with the cursor,
12. average, minimum, maximum and standard deviation of all distances, and
13. density of selected points per 10000 μm².

From the MST, the following features were calculated:

1. number of points,
2. total, average, minimum and maximum MST line length,
3. number and percentages of points with one neighbour, two neighbours, three neighbours, four neighbours, and five neighbours composed in each field.

The values per case were the total of the measurements in the five fields taken together and for each of the features, the mean and standard deviation (SD) of the five fields analysed were calculated per case.

3. Statistical analysis

Statistical analysis was performed by SPSS version 11.0 (SPSS Inc., Chicago, USA) for each SSA feature using the consensus CIN grade and presence of high-risk HPV DNA as the independent variables (“gold standards”). To investigate significant differences between the three CIN grades, a univariate analysis using the Kruskal–Wallis test with multiple comparisons was performed. The difference between the individual groups was evaluated with a Mann–Whitney U-test. The level of significance was set to $p < 0.05$. Subsequently, the intra-observer reproducibility of the variables was assessed by calculating the mean coefficient of variation (MCV) for all SSA variables from five specimens that were measured five times by the same observer. A jack-knifed stepwise discriminant analysis was used to find the best discriminating set of features between the three CIN grades. Since in a discriminant analysis the number of variables used should not exceed one fifth of the number of cases, a selection had to be made. Variables with a CV $> 10\%$ were regarded insufficiently reproducible and were therefore omitted. Seventeen variables had a CV $< 10\%$ and were regarded sufficiently reproducible for the analysis. The diagnostic consensus was checked by means of the Kappa statistic.

4. Results

Descriptive statistics of the significant SSA features for all review CIN cases are displayed in Table 1. One of the SSA features (points from which the surrounding surfaces has 4 edges, PECO-4) was significantly different between all the three CIN grades but many other features revealed significant differences between CIN1 and CIN2 or between CIN2 and CIN3 cases. The maximum MST line length (MML) was the strongest discriminator between the different CIN grades, however with considerable overlap. With stepwise discriminant analysis, the best multivariate combination of features to distinguish the different CIN grades was the MML and the Area Disorder (Fig. 3). CIN1 cases in general have higher values than the CIN3 cases, and
Fig. 2. Illustration of the image analysis method. In each case the most severely dysplastic part of the epithelium in the H&E section used for the routinely assessed diagnosis was selected. In this area, the upper border, the basal membrane, and the borders between the two deepest layers of the epithelium were marked (the deepest layer excluding the basal cell layer). SSA was performed in fields of vision chosen randomly in the already demarcated area by constructing a Voronoi diagram (VD) and a minimum spanning tree (MST) per field. Briefly, with an ×40 objective (final magnification ×1200) the centres of gravity of all nuclei in the lower deep half of the epithelium per field of vision were interactively marked on a video screen by setting a point with the cursor of the system. Using this set of points, the computer program composed the VD and the MST. The Voronoi tessellation splits the image plane into polygons, each containing one nuclear centre of gravity in its centre and in such a way that each point in the polygon was closer to this centre nucleus than to any other point in the plane. The MST interconnected a set of nuclei in such a way that the total length of the lines was minimal and no loops were formed. The quantitative image analysis system automatically calculated a large number of quantitative features for each case. The values per case were the total of the measurements in the five fields taken together and for each of the features, the mean and standard deviation (SD) of the five fields analysed were calculated per case.

many of the CIN2 cases fall in between the grade 1 and the grade 3 cluster but with considerable overlap. Tables 2 and 3 show the confusion matrices of the consensus CIN grades and the prediction based on SSA quantitative pathology (QP) classification. Crude overall classification of the consensus CIN grades was 69%. Of the CIN1 consensus cases, 4/16 (25%) were misclassified as CIN2 and 1 (7%) as CIN3. Of the 27 CIN2 cases, 7 (26%) were misclassified as CIN1 and 10 (37%) as CIN3. Five (10%) of the 51 CIN3 cases were misclassified as CIN2 but none as CIN1.

Eighty percent (70/88) of the CIN cases studied were positive for high-risk HPV. The 18 patients with high-risk HPV-negative CIN lesions were marginally older than those with positive HPV lesions (38.9 versus 35.7 years), but with a considerable variation and overlap. Several SSA features were significantly different between the high-risk HPV positive and negative group, but the maximum MST line length and the area disorder is the best combination to distinguish cases with and without high-risk HPV DNA (Fig. 4). Of the 62 CIN cases with max MST line length > 15, 6 (10%) were negative and 56 (90%) were positive for high-risk HPV. In contrast, of the 26 CIN cases with max MST line length ≤ 15 or max linelength > 15 and areadisorder < 0.25], 14 cases (54%) were high-risk
Table 1

<table>
<thead>
<tr>
<th>Feature</th>
<th>CIN1 (Mean (SD))</th>
<th>P 1 vs. 2</th>
<th>CIN2 (Mean (SD))</th>
<th>P 2 vs. 3</th>
<th>CIN3 (Mean (SD))</th>
<th>P 1 vs. 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minimum area</td>
<td>56.4 (34.7)</td>
<td>0.07</td>
<td>66.3 (26.3)</td>
<td>0.003</td>
<td>40.0 (13.2)</td>
<td>0.4</td>
</tr>
<tr>
<td>Maximum area</td>
<td>1098.1 (688.8)</td>
<td>0.2</td>
<td>771.0 (244.6)</td>
<td>0.11</td>
<td>599.9 (142.0)</td>
<td>0.003</td>
</tr>
<tr>
<td>Standard deviation of areas</td>
<td>184.8 (107.2)</td>
<td>0.4</td>
<td>137.9 (55.2)</td>
<td>0.01</td>
<td>83.6 (35.6)</td>
<td>0.003</td>
</tr>
<tr>
<td>Area disorder</td>
<td>0.44 (0.05)</td>
<td>0.005</td>
<td>0.38 (0.05)</td>
<td>0.47</td>
<td>0.39 (0.03)</td>
<td>0.03</td>
</tr>
<tr>
<td># Points from which the</td>
<td>7.8 (1.8)</td>
<td>0.03</td>
<td>6.0 (1.7)</td>
<td>0.001</td>
<td>12.2 (4.2)</td>
<td>0.02</td>
</tr>
<tr>
<td>surrounding surface has 4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>edges</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td># Points from which the</td>
<td>23.8 (6.7)</td>
<td>0.04</td>
<td>17.7 (7.1)</td>
<td>0.01</td>
<td>34.4 (17.8)</td>
<td>0.2</td>
</tr>
<tr>
<td>surrounding surface has 5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>edges</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average of all distances</td>
<td>15.2 (4.3)</td>
<td>0.8</td>
<td>15.0 (3.0)</td>
<td>0.02</td>
<td>11.8 (2.2)</td>
<td>0.02</td>
</tr>
<tr>
<td>Minimum of all distances</td>
<td>3.8 (1.2)</td>
<td>0.06</td>
<td>4.6 (1.3)</td>
<td>0.001</td>
<td>2.9 (0.6)</td>
<td>0.04</td>
</tr>
<tr>
<td>Maximum of all distances</td>
<td>42.9 (12.9)</td>
<td>0.1</td>
<td>34.1 (7.0)</td>
<td>0.04</td>
<td>28.3 (7.2)</td>
<td>0.01</td>
</tr>
<tr>
<td>Standard deviation of all</td>
<td>16.8 (4.8)</td>
<td>0.9</td>
<td>16.2 (3.4)</td>
<td>0.01</td>
<td>12.8 (2.6)</td>
<td>0.02</td>
</tr>
<tr>
<td>distances</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average MST line length</td>
<td>10.5 (2.9)</td>
<td>0.5</td>
<td>10.6 (1.9)</td>
<td>0.003</td>
<td>8.0 (1.4)</td>
<td>0.01</td>
</tr>
<tr>
<td>Minimum MST line length</td>
<td>3.4 (1.1)</td>
<td>0.03</td>
<td>4.2 (1.1)</td>
<td>0.001</td>
<td>2.7 (0.5)</td>
<td>0.1</td>
</tr>
<tr>
<td>Maximum MST line length</td>
<td>25.8 (7.5)</td>
<td>0.2</td>
<td>21.7 (5.8)</td>
<td>0.003</td>
<td>15.0 (2.8)</td>
<td>0.001</td>
</tr>
<tr>
<td>Standard deviation of MST</td>
<td>4.4 (1.5)</td>
<td>0.2</td>
<td>3.6 (1.2)</td>
<td>0.004</td>
<td>2.4 (0.5)</td>
<td>0.002</td>
</tr>
<tr>
<td>length</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td># Points with one neighbour</td>
<td>26.0 (6.6)</td>
<td>0.04</td>
<td>19.6 (6.9)</td>
<td>0.01</td>
<td>34.6 (16.4)</td>
<td>0.3</td>
</tr>
<tr>
<td># Points with two neighbours</td>
<td>57.1 (15.5)</td>
<td>0.05</td>
<td>(12.7)</td>
<td>0.006</td>
<td>78.4 (36.2)</td>
<td>0.2</td>
</tr>
<tr>
<td># Points with three neighbours</td>
<td>20.7 (5.5)</td>
<td>0.07</td>
<td>15.9 (6.1)</td>
<td>0.01</td>
<td>29.8 (14.9)</td>
<td>0.2</td>
</tr>
</tbody>
</table>

Table 2

Confusion matrix, sensitivity, specificity, positive and negative predictive value of consensus CIN grade and SSA quantitative pathology (QP) classification based on the maximum MST line length (MML) and the area disorder (AD) (the agreement cases are shaded grey)

<table>
<thead>
<tr>
<th>Consensus CIN diagnosis</th>
<th>SSA QP classification</th>
<th>Total (MML and AD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CIN1</td>
<td>11</td>
<td>4</td>
</tr>
<tr>
<td>CIN2</td>
<td>7</td>
<td>10</td>
</tr>
<tr>
<td>CIN3</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>Total</td>
<td>18</td>
<td>19</td>
</tr>
</tbody>
</table>

Overall correct classification = (11 + 10 + 40)/88 = 77%.

Table 3

Correlation table of high-risk human papillomavirus (HPV) positive and negative consensus cervical intraepithelial neoplasia (CIN) cases and SSA QP classification based on the SSA QP features maximum MST line length (MML) and area disorder (AD)

<table>
<thead>
<tr>
<th>HPV DNA status</th>
<th>HPV prediction from SSA QP (MML and AD)</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>HPV negative</td>
<td></td>
<td>6</td>
</tr>
<tr>
<td>HPV positive</td>
<td></td>
<td>56</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>62</td>
</tr>
</tbody>
</table>

Sensitivity HPV positive vs. HPV negative = 56/(56 + 14) = 80%.
Specificity HPV negative vs. HPV positive = 12/(12 + 6) = 67%.
Positive predictive value HPV negative vs. HPV positive = 56/(6 + 56) = 90%.
Negative predictive value HPV negative vs. HPV positive = 12/(12 + 14) = 46%.
Overall agreement = (12 + 56)/88 = 77%.

5. Discussion

The aim of this study was to assess the correlation between syntactic structure analysis (SSA) features, dysplasia grade and presence of high-risk human papillomavirus (HPV) in cervical intraepithelial neoplasia (CIN). Previous studies in other organs found that SSA features were correlated to grade (breast cancer) [2] and also prognosis (oral cavity cancer, colonic polyps) [5,15]. Indeed, SSA features in CIN are correlated with both grade and presence of high-risk HPV
DNA. However, there was considerable overlap for the different grades. One explanation could be that pathologists diagnoses are not well reproducible [14], but although the discrimination of the expert’s consensus grades was better there was also considerable overlap.

It must therefore be concluded that SSA features are insufficient to be used as quality control or support methods for routine grading.

It should be kept in mind, however, that grade is not the final gold standard but a morphological impression of the pathologists with a certain (fairly weak) [7] correlation with progression. Therefore, it may be more important to use certain biological cell properties as an alternative gold standard. Interestingly, the correlation of SSA features and high-risk HPV DNA was good (Fig. 3), suggesting that SSA parameters accurately reflect basic disturbances of cellular control mechanisms of the HPV-infected epithelium. It would be interesting to investigate whether SSA features also correlate to other molecular biological features reflecting important cell processes. This would also be important, as many studies have shown that quantitative microscopical features are very sensitive and accurate prognostic markers. However, while they are not very popular amongst pathologists as the underlying biological mechanism is often unclear, such molecular cell biological correlation studies are currently being undertaken.

It is tempting to hypothesize about the biologic background of the syntactic structure features described. The feature maximum MST line length decreased with increasing CIN grades. This may reflect a lower nuclear area and/or a tighter packing of nuclei in CIN3 compared with CIN2 and CIN1. The fact that the features minimum MST line length and the average MST line length also showed decreasing values in the CIN3 cases as compared to the CIN1 cases supports this. Furthermore, our finding is in agreement with the results of Boon et al. who showed that increasing CIN grade is accompanied by increasing values of the nucleus/cytoplasma ratio [1]. With respect to the meaning of the feature PECO-4 (the number of points from which the surrounding surfaces has 4 edges), in normal cervical epithelium most nuclei would have four neighbours. As expected, this feature showed decreasing values for the CIN2 cases as compared to the CIN1 cases and, interestingly, this feature showed increasing values for the CIN3 cases as compared to the CIN1 and CIN2 cases. This can be caused due to the large number of basaloid cells present in the cervical epithelium in CIN3 cases which become more tightly packed. The feature Area Disorder showed the same tendency, which may also reflect this more dense packing of basaloid cells in the epithelium. Another explanation could be that these CIN3 cases represent a group of lesions with variable degrees of differentiation. This
is reflected by a high standard deviation of this feature, indicating that there is considerable variation in the quantitative characteristics between the individual biopsies. Therefore, it would be interesting to correlate this feature with the follow-up of each patient.

We conclude that SSA features are correlated with both CIN grade and presence of high-risk HPV DNA, but the discrimination power is insufficient to be used as a routine method for quality control of subjective grade or as a surrogate marker for high-risk HPV DNA presence. Further evaluation of the prognostic value as to progression prediction in early CIN lesions is important.

Acknowledgment

Supported by grant 02-281 of the Stichting Bevordering Diagnostische Morfometrie.

References


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