

# Caveats: Numerical requirements in graph theory based quantitation of tissue architecture

J. Sudbø<sup>a,\*</sup>, R. Marcelpoil<sup>b</sup> and A. Reith<sup>a</sup>

<sup>a</sup> *Section for Digital Pathology, The Norwegian Radium Hospital, Oslo, Norway*

<sup>b</sup> *University of Grenoble, Université Joseph Fourier, Grenoble, France*

Accepted 30 October 2000

Graph theory based methods represent one approach to an objective and reproducible structural analysis of tissue architecture. By these methods, neighborhood relations between a number of objects (e.g., cells) are explored and inherent to these methods are therefore certain requirements as to the number of objects to be included in the analysis. However, the question of how many objects are required to achieve reproducible values in repeated computations of proposed structural features, has previously not been addressed specifically.

After digitising HE stained slides and storing them as grey level images, cell nuclei were segmented and their geometrical centre of gravity were computed, serving as the basis for construction of the Voronoi diagram (VD) and its subgraphs. Variations in repeated computations of structural features derived from these graphs were related to the number of cell nuclei included in the analysis.

We demonstrate a large variation in the values of the structural features from one computation to another in one and the same section when only a limited number of cells (100–500) are included in the analysis. This variation decreased with increasing number of cells analyzed. The exact number of cells required to achieve reproducible values differ significantly between tissues, but not between separate cases of similar lesions. There are no significant differences between normal and malignantly changed tissues in oral mucosa with respect to how many cells must be included.

For graph theory based analysis of tissue architecture, care must be taken to include an adequate number of objects; for

some of the structural features we have tested, more than 3000 cells.

Keywords: Graph theory, caveats, reproducibility, numerical requirements

**Abbreviations:** CCD, Charge coupled device (digital camera); CV, Coefficient of variation, given by the equation  $SD/\text{mean}$ ; DT, Delaunay Triangulation; GG, Gabriel Graph; HE, Hematoxylin-Eosine; IOD, Integrated optical density; MST, Minimum Spanning Tree; SD, Standard Deviation, given by the equation  $SD = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2}$ , where  $n$  is the number of observations (here: computations),  $x_i$  is the observed value (e.g., of a structural feature), and  $\bar{x}$  is the mean of all the (computed) values; UT, Ulam Tree; VD, Voronoi Diagram; 2D, Two-dimensional.

## 1. Introduction

Despite the introduction of simplified grading systems, inter- and intraobserver disagreement still is a challenge to diagnostic pathology [1–6]. As this lack of consistency may lead to reduced prognostic value, computer assisted quantitation of tissue architecture has been attempted, and numerous papers on mathematical modelling of tissue structure have been published over the past decades [7–14]. As one of several possible approaches, graph theory based methods have been applied both on a tissular [8,15,16], cellular [9,17] and subcellular level [18–21]. The Voronoi diagram (VD) is frequently chosen as the primary graph, from which a number of subgraphs to explore neighborhood relations between cells or other tissue components may be developed [22]. Such graphs include the Delaunay Triangulation (DT) [23], the Minimum Spanning Tree (MST) [24], the Gabriel Graph (GG) [23] and the Ulam Tree (UT) [25]. For a meaningful investigation of neighborhood relations, a pre-

\*Corresponding author: Jon Sudbø, MD, DDS, Department of Pathology, Section for Digital Pathology, The Norwegian Radium Hospital, Montebello, 0310 Oslo, Norway. Tel.: + 47 22934230; Fax: + 47 22935627; E-mail: jon.sudbo@rh.uio.no.

requisite is that a certain number objects (i.e., cells within the scope of this paper) are included. To our knowledge, the question of how structural features derived from the VD and its subgraphs depend on the number of objects analysed has not been addressed specifically in previous papers on this topic. Conceivably, there are requirements as to the minimum number of objects to be included in order to achieve reproducible and thereby valid results of the analysis. We have previously [26] presented a number of structural features for describing tissue architecture. This report investigates how the number of cell nuclei included in graph theory based analysis influences the results of the analysis and relates this to prognostic groups in carcinomas of the tongue, larynx, prostate and cervix. Detailed results are presented for an arbitrarily chosen subset of the features.

## 2. Material and methods

Analysis of altogether 27 structural features (Table 6) [26] was done on HE stained sections from normal oral mucosa, as well as carcinomas of the tongue, larynx, prostate and cervix. Details on 6 of these 27 structural features are presented (Table 1). For investigating the numerical requirements of separate structural features in specific tissues, 4 blocks from separate cases, with 4 slides from each block were investigated. The inter-individual variation in similar lesions was investigated by the same procedure. For the analysis, slides with at least 10,000 cells were included in the analysis. In squamous cell epithelium only epithelial cells deep to stratum granulosum were included, as only these cells have nuclei with a definition adequate for a reasonably precise segmentation.

### 2.1. Image acquisition

Grey level images from 5–7  $\mu\text{m}$  thick HE stained sections were digitised using a charged coupled device (CCD) camera (Philips LDH 0670/00 equipped with a Hamamatsu AC Adaptor, type A3472) mounted on a Zeiss Axioplan 2 microscope using a Plan-Neofluar 40 $\times$  objective and a numerical aperture of 0.65–0.75 lens in addition to a Prior microscope HI52V2 stage (Prior Scientific Instruments, UK). The final magnification was 400 $\times$  at an estimated resolution of 876 nm (0.9  $\mu\text{m}$ ) per pixel.

Table 1

The feature values as related to the number of cells included in the analysis for the structural features DEL\_av, WGC\_av, A\_dis, DFRAC\_av, NNRR and RMPB. The CV varies between 20.2% at 100 cells analysed and 0.17% at 5000 cells analysed when the analysis was performed on normal oral mucosa (Table 1). The value of 0.17 was an extreme value, that was not reproduced for the other tissues. However, the pattern was the same; a small number of objects included yielded a large CV (typically 15–20%), whereas a large number of cells included in the analysis yielded a CV of less than 5% in all cases

Computations	Number of cells analysed		
	100	500	5000
(A) DEL_av*			
1	29.72	19.33	29.12
2	19.62	24.31	29.02
3	32.59	22.44	29.03
4	28.32	25.52	29.11
(B) WGC_av*			
1	0.3471	0.3156	0.3233
2	0.4250	0.2355	0.3025
3	0.2371	0.4154	0.3363
4	0.3647	0.2758	0.3126
(C) A_dis*			
1	0.5471	0.5156	0.5233
2	0.7250	0.5255	0.4925
3	0.3371	0.5454	0.5363
4	0.4647	0.5458	0.5126
(D) DFRAC_av*			
1	1.55	1.82	1.78
2	1.73	1.52	1.80
3	1.34	1.54	1.83
4	1.17	1.65	1.81

\* DEL\_av: Average Delaunay Edge Length; WGC\_av: Average Weighted Global Compacity; A\_dis: Area disorder (of the Voronoi polygons); DFRAC\_av: Average fractal dimension of the Ulam trees.

### 2.2. Building a composite picture

For composite pictures generated by manual movement of the microscope stage an algorithm using the binary mask of nuclear profiles was employed. For further details on building a composite picture, the reader is referred to [26].

### 2.3. Segmentation and computations of structural features

For our purpose, we have employed local segmentation, and developed an algorithm based on the size

of the elements to be detected and their contrast to the background. Thresholding was based on the pixel darkness measured as integrated optical density (IOD). Any pixel with an IOD within a given range was turned ON, otherwise OFF. Algorithms for constructing nuclear profiles were based on mathematical morphology [27] and from these, the geometrical center of gravity was computed. The resulting data were stored as files of coordinates, where the coordinates represented a center of gravity. From the 2D swarm of pointlike seeds thus obtained, VD and subgraphs (DT, MST, GG and UT) were constructed. These graphs, in turn, were the basis for computation of a total of 27 structural features [26].

#### 2.4. Structural analysis

For a complete description of the structural features developed and tested (Table 6) as well as the 6 features presented in more detail, see [26]. Analysis of 6 arbitrarily chosen and independent structural features (A\_dis [Area disorder], WGC\_av [average weighted global compacity], DEL\_av [average Delaunay edge length], DFRAC\_av [average fractal dimension of the Ulam Trees], NNRR [Number of Nearest Neighbours within a Restricted radius of 75 pixels], RMPB [maximum radius of the percolating ball]) was performed on carcinoma of the tongue, larynx, prostate and cervix.

#### 2.5. Selection of objects to analyse

A sample of approximately 50 objects (cell nuclei) placed along the epithelium-stroma interface was randomly selected. The number of object to analyse was increased in increments of 50, moving bilaterally along the epithelium and the number of cells at which the CV of the analysis was acceptable was assessed. For comparing the numerical requirements of different tissues, the average of four such computations in 4 slides of every tissue analysed was computed (Fig. 2). To eliminate biological variance, the structural analysis performed in this study was performed on single sections of normal oral mucosa, cervical carcinoma, prostate carcinoma and laryngeal carcinoma. For each specimen, the analysis was performed in as many increments as was necessary to achieve reproducible values (CV < 5% in 15–20 repeated computations) in 8–10 consecutively incremental computations. The window of analysis was defined digitally, by defining a closed contour with a digitizing pad and storing the coordinates of the contour. Only pointlike seeds within the

contour were included in the analysis. This was typically done for epithelial islands bordering onto the underlying stroma, or for the basalmost part of the squamous cell mucosa. The coordinates of the part of the contour crossing a marginal polygone were defined as the new edge in the polygone.

#### 2.6. Estimating the robustness of the methods

The number of nuclei included in the analysis was varied by setting gradually increasing inclusion thresholds for pixel values to be included in the analysis. This was done interactively on the image analysis system prior to segmentation. The percentage of missed nuclei was estimated by evaluating a number of slides and manually counting how many nuclei of epithelial cells did not generate a Voronoi polygone. After grouping the slides according to approximate percentages of missed nuclei (1, 5, 10, 15 and 30% missed nuclei) in 15–20 repeated computations in a selected digitally window of analysis consisting of a *total* number of epithelial nuclei corresponding to the number of objects at which reproducible data were obtained (Table 6) was performed.

#### 2.7. Statistical evaluation

The statistical evaluation was done using SPSS 9.0 for Windows (Chicago, Illinois, USA, 1999). Comparison of computed values in different groups of lesions was done by Student's *t*-test. All *P*-values were two-tailed, and values less than 0.05 were considered to indicate statistical significance.

### 3. Results

There was a significant variation in the values of the form parameters when only a limited number of cells are included in the analysis (Figs 1–3, Table 1).

#### 3.1. Stability in values of structural features as a function of number of cells analysed for different structural features in the same tissue

For the structural features DEL\_av, A\_dis, WGC\_av and DFRAC\_av the number of objects required to achieve reproducible values (estimated in carcinoma of the prostate) varied considerably (Fig. 1). Thus, for DEL\_av approximately 850 cells are required, for A\_dis the corresponding figure was 1500 cells, for WGC\_av the number was between 4500 and 5000, and for DFRAC\_av the number was approximately 850.

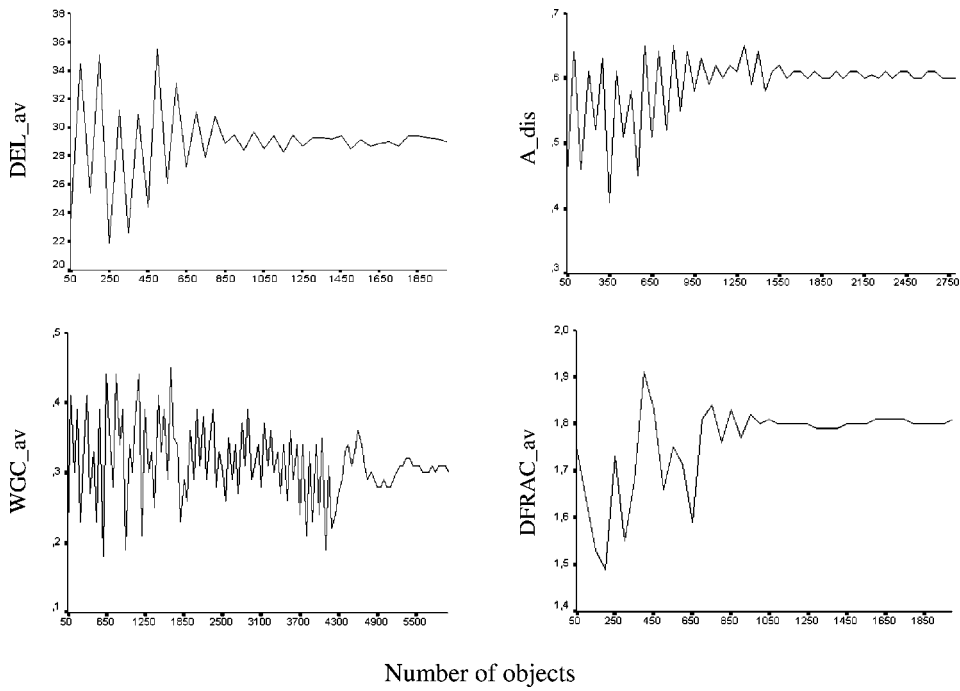


Fig. 1. The absolute values of the four structural features (DEL\_av, A\_dis, WGC\_av and DFRAC\_av) as a function of the number of cells analyzed when computed in a case of carcinoma of the prostate. The number of objects (cells) required to achieve reproducible values (CV < 5%) varies considerably from one structural feature to another.

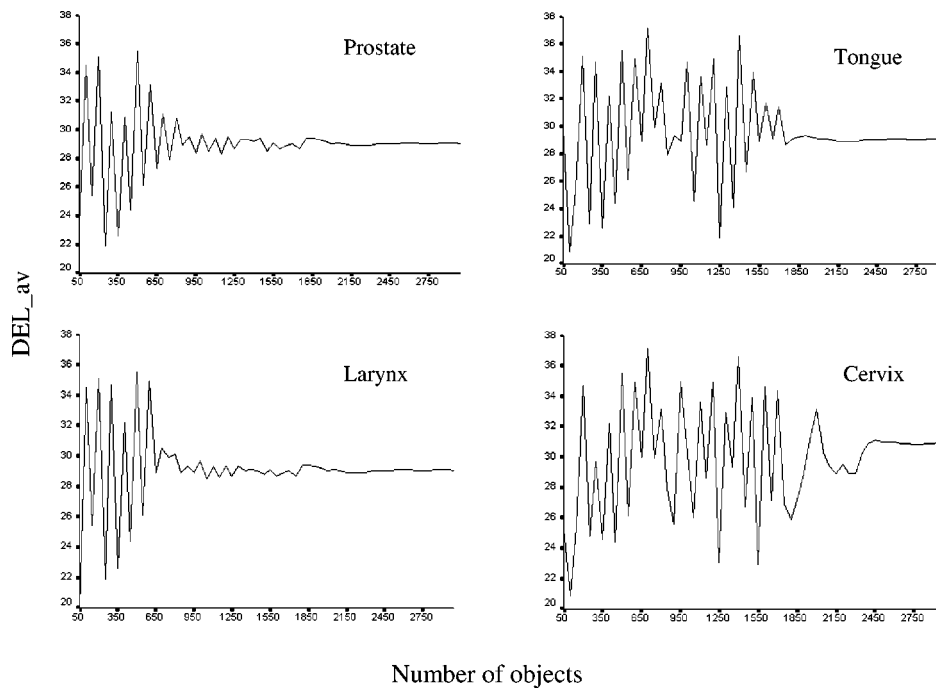


Fig. 2. The absolute values of the form parameter DEL\_av as a function of the number of cells analyzed in four different tissues (carcinoma of the prostate, tongue, larynx and cervix). The number required to achieve reproducible values for a single structural feature varies considerably between tissues (approximately 800 cells in carcinoma of the larynx, 1200 cells in carcinoma of the prostate, 1800 cells in carcinoma of the tongue and approximately 2500 cells in carcinoma of the cervix).

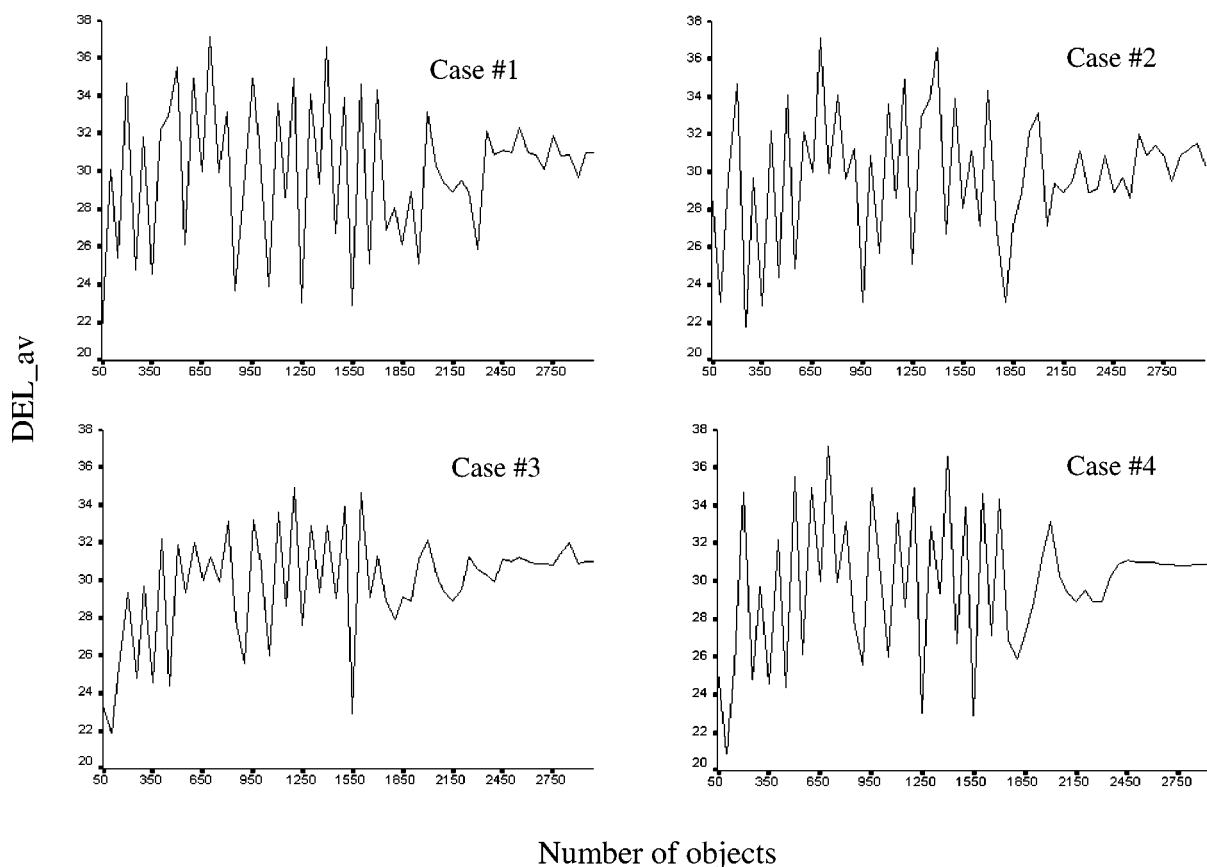


Fig. 3. Reproducibility as a function of number of cells analysed for one single structural feature in different cases of the cervical carcinomas. The number of cells required to achieve reproducible values for the structural feature  $DEL_{av}$  varies somewhat between cases. However, this variation was significantly less than the variation observed between different tissues ( $P = 0.01$ ).

### 3.2. Stability in values of structural features as a function of number of cells analysed for a single structural feature in different tissues

The absolute values of the form parameter  $DEL_{av}$  stabilises at an included number of objects that varies considerably between carcinomas of the prostate, tongue cervix and larynx (Fig. 2). Thus, approximately 800 cells were required for the analysis in carcinoma of the larynx, 1200 cells in carcinoma of the prostate, 1800 cells in carcinoma of the tongue and approximately 2500 cells in carcinoma of the cervix ( $P = 0.01$ ).

### 3.3. Stability as a function of number of cells analysed for one single structural feature in different cases of the cervical carcinomas

In 4 cases of cervical carcinomas (2 with good and 2 with poor prognosis) the values obtained at var-

ious number of cells included for structural feature  $DEL_{av}$  showed some variation when we investigated this structural feature as a function of the number of cells analysed (Fig. 3). For the structural feature  $DEL_{av}$ , the number of cells required to achieve reproducible values of the analysis ( $CV < 5\%$ ) varied between 2200 (Case #3, Fig. 3) and 2500 (Case #2, Fig. 3).

### 3.4. CV as related to the number of cells analysed

For the structural features  $DEL_{av}$ ,  $WGC_{av}$ ,  $A_{dis}$  and  $DFRAC_{av}$ , the CV at four different computations including 100, 500 and 5000 cells varies between 20.2 and 0.17% when the analysis was performed on normal oral mucosa (Table 1). The value of 0.17 was an extreme value, not reproduced for the other tissues. However, the pattern was the same; a small number of objects included yielded a large CV (typically 15–20%), whereas a large number of cells included in the analysis yielded a CV of less than 5% in all cases.

Table 2

Reproducibility of 4 structural features (DEL\_av, WGC\_av, A\_dis and DFRAC\_av) when 20 separate computations have been performed on three separate cases of carcinomas of the tongue, all with a favourable prognosis. For each case, values from 4 of the 20 computations are shown. Approximately the same values of the CV were obtained when computations were performed on 3 cases with poor prognosis

	Case #1	Case #2	Case #3
<b>DEL_av*</b>			
1	29.52	25.32	28.12
2	28.62	27.31	27.03
3	29.64	26.31	27.54
4	27.33	27.52	29.11
<b>WGC_av*</b>			
1	0.52	0.45	0.42
2	0.49	0.39	0.40
3	0.54	0.41	0.36
4	0.51	0.44	0.38
<b>A_dis*</b>			
1	0.54	0.52	0.52
2	0.53	0.53	0.49
3	0.49	0.54	0.53
4	0.56	0.51	0.51
<b>DFRAC_av*</b>			
1	1.65	1.59	1.52
2	1.63	1.54	1.49
3	1.64	1.55	1.53
4	1.66	1.56	1.51

\* DEL\_av: Average Delaunay Edge Length; WGC\_av: Average Weighted Global Compacity; A\_dis: Area disorder (of the Voronoi polygones); DFRAC\_av: Average fractal dimension of the Ulam trees.

### 3.5. Reproducibility in repeated computations on three different cases when 5000 cells are included in the analysis

We found a high degree of reproducibility when analyzing the 4 structural features (DEL\_av, WGC\_av, A\_dis and DFRAC\_av) in 20 separate computations on three separate cases of carcinomas of the tongue, all with a favourable prognosis (Table 2 shows the data for 4 computations arbitrarily chosen out of a total of 20 computations). Approximately the same values of the CV were obtained when computations were performed on 3 cases with poor prognosis.

### 3.6. CV as a function of number of cells, 4 structural features in 4 tissues

The CV of four structural features (DEL\_av, WGC\_av, A\_dis, DFRAC\_av, NNRR and RMPB) was

Table 3

The CV of four structural features (DEL\_av, WGC\_av, A\_dis and DFRAC\_av) in carcinoma of the tongue, larynx, prostate and cervix when 100, 500 and 5000 epithelial cells are included in the analysis in 20 repeated computations. For all the tissues and structural features investigated, the CV becomes acceptably low between 850 and 5000 cells analysed

Structural features	Number of objects (cell nuclei) analysed			
	100	500	1500	5000
<b>Carcinoma of the tongue</b>				
DEL_av*	20.20%	11.30%	4.79%	2.27%
WGC_av†	21.30%	13.60%	5.02%	3.30%
A_dis‡	14.80%	7.90%	4.27%	2.43%
DFRAC_av¶	14.60%	9.40%	6.16%	4.54%
NNRR	22.19%	11.67%	7.09%	4.87%
RMPB	18.76%	10.98%	6.79%	4.15%
<b>Carcinoma of the larynx</b>				
DEL_av*	19.40%	12.80%	5.15%	3.14%
WGC_av†	25.30%	11.70%	4.93%	2.94%
A_dis‡	18.80%	9.60%	4.73%	3.73%
DFRAC_av¶	17.40%	10.40%	4.91%	3.45%
NNRR	19.19%	10.38%	7.44%	4.87%
RMPB	17.76%	10.23%	5.23%	4.15%
<b>Carcinoma of the prostate</b>				
DEL_av*	21.90%	10.50%	7.15%	4.23%
WGC_av†	22.40%	13.70%	5.93%	4.06%
A_dis‡	14.30%	10.90%	6.73%	3.67%
DFRAC_av¶	21.70%	11.10%	5.91%	4.38%
NNRR	23.05%	12.17%	6.35%	4.87%
RMPB	18.70%	11.03%	5.49%	4.15%
<b>Carcinoma of the cervix</b>				
DEL_av*	24.10%	11.80%	4.00%	2.28%
WGC_av†	21.30%	13.60%	5.19%	2.96%
A_dis‡	14.80%	6.90%	6.27%	3.37%
DFRAC_av¶	18.90%	7.30%	7.16%	4.97%
NNRR	21.19%	10.47%	7.00%	4.37%
RMPB	19.26%	9.98%	6.73%	3.35%

\* Average Delaunay edge length of the considered Delaunay network.

† Average compacity.

‡ Area disorder.

¶ Average (Hausdorff) fractal dimension of the Ulam trees.

estimated in four separate cases from carcinoma of the tongue, larynx, prostate and cervix (Table 3). When 100, 500, 1500 and 5000 epithelial cells were included in the analysis, the CV becomes acceptably low between 500 and 5000 cells analysed for all of the tissues and structural features investigated. A similarly acceptable value was found for all the other 23 structural features (Table 6).

Table 4

The CV for each tissue was estimated as a mean of altogether 20 computations, in 4 separate locations of 4 separate slides, each representing different cases of similar lesions. This was repeated in 4 separate tissues; carcinoma of the tongue, larynx, prostate and cervix. The data shown are from carcinomas of the tongue. For all the 4 tissue investigated, the CV was acceptable when 5000 cells were included in the analysis

Structural features	Average CV (range)
DEL_av*	2.98% (2.27–4.32%)
WGC_av†	3.30% (2.94–4.06%)
A_dis‡	3.30 (2.43–3.73%)
DFRAC_av¶	4.34 (3.45–4.97%)

\*Average Delaunay edge length of the considered Delaunay network.

†Average weighted global compacity.

‡Area disorder.

¶Average (Hausdorff) fractal dimension of the Ulam trees.

Table 5

Two cases of carcinomas of the prostate where computations have been done with an increasing percentage of cells missed in the segmentation. The percentage missed was estimated by manual counting of how many nuclei of epithelial cells did not generate a Voronoi polygone. Table A shows the CV as a function of cells missed in the segmentation of a total of initially 1000 cells in carcinoma of the prostate for the structural feature DEL\_av. The CV represents a mean of five computations in separate slides from a case with good prognosis (relapse free survival more than 10 years after radical prostatectomy). A similar pattern is seen in Table B, for a case with poor prognosis (relapse free survival less than 3 years after radical prostatectomy)

Percentages of cells missed	Mean values with range	CV
<b>A</b>		
1% missed	29.1 (28.7–29.5)	2.23%
5% missed	28.6 (27.9–29.1)	2.41%
10% missed	30.1 (28.7–30.8)	3.57%
15% missed	27.6 (26.8–28.4)	5.91%
30% missed	29.4 (27.1–33.9)	16.76%
<b>B</b>		
1% missed	33.8 (28.7–29.5)	3.47%
5% missed	33.6 (31.9–35.1)	4.11%
10% missed	32.1 (31.7–34.1)	4.83%
15% missed	31.4 (26.8–34.4)	8.32%
30% missed	29.4 (25.9–35.8)	25.23%

### 3.7. CV of selected structural features at 5000 cells analysed

The CV was estimated for each tissue as a mean of altogether 16 computations, in 4 separate locations of 4 separate slides, each representing different cases of similar lesions (Table 4). This was repeated in 4 sepa-

Table 6

The total of 27 structural features, with their corresponding number of cells at which the CV was less than 5%, and the CV at 5000 cells analysed. The data shown are from 4 separate cases of normal oral mucosa. Similar data (not shown) were found in carcinomas of the tongue and cervix. All but one of the features (DEL\_av) require more than 1000 cells analysed in order to achieve reproducible values (CV < 5%). Thirteen of the investigated structural features require more than 1500 cells in order to achieve reproducible values. Seven features require more than 2000 cells, and 3 (DKNN\_av, WGC\_av and WGC\_dis) features require 3000 cells or more

Structural features	Number of objects at which CV < 5%	CV at 5000 object analysed
A_dis	1450	3.30
DEL_av	850	2.98
DEL_dis	1350	3.98
DENS	1400	4.15
DEP_av	1250	3.96
DEP_dis	1150	4.89
DFRAC_av	950	2.34
DKNN_av	3000	4.77
DKNN_dis	2350	4.90
DRT	1200	3.15
ELH_av	1900	4.18
HA_av	1750	3.98
HA_dis	1450	4.75
MSPDG	1050	2.35
MSTEL_av	1200	3.70
MSTEL_dis	1450	2.15
NNRR_av	2100	4.67
NNRR_dis	1700	3.77
PTS	2000	4.56
PTS_av	1850	3.77
PTS_dis	1350	4.45
RMPB	1200	4.98
RF_av	2050	4.90
RF_dis	1800	3.10
WGC	1500	2.30
WGC_av	4950	3.30
WGC_dis	3400	5.90

\*For a description of the structural features, see [26].

rate tissues; carcinoma of the tongue, larynx, prostate and cervix. For all the 4 tissues investigated, the CV was acceptable when 5000 cells were included in the analysis. The CV did not extend 5% in any of the cases that included 5000 cells.

### 3.8. CV as a function of inaccuracies of segmentation

We have chosen a cut-off for the CV at 5%. The CV of the computations did not exceed this value before 15% of the cells were missed, as shown for initially

Table 7

The prognostic value of the two independent structural features DEL\_av and ELH\_av as related to the number of cells included in the computations. These two structural features attain a prognostic value ( $P \leq 0.05$ ) approximately at the same values at which their reproducibility was acceptable ( $CV \leq 5\%$ ). The ranges of the computations are given in brackets

		Number of cells included in the computations			
		100	500	2000	5000
Carcinoma of the prostate					
DEL_av	Good prognosis	24.8 (22.7–28.9)	29.7 (25.4–30.9)	26.7 (25.3–27.9)	27.3 (26.1–28.3)
	Poor prognosis	25.1 (23.9–27.3)	27.1 (24.4–29.6)	24.8 (22.9–25.9)	25.4 (24.6–26.3)
	<i>P</i> -value	0.25	0.23	0.02	0.01
ELH_av	Good prognosis	0.43 (0.28–0.59)	0.41 (0.31–0.49)	0.39 (0.32–0.47)	0.40 (0.35–0.48)
	Poor prognosis	0.41 (0.23–0.51)	0.38 (0.29–0.45)	0.32 (0.26–0.37)	0.32 (0.27–0.37)
	<i>P</i> -value	0.28	0.17	0.05	0.005
Carcinoma of the larynx					
DEL_av	Good prognosis	25.6 (22.7–28.3)	26.1 (23.9–28.1)	24.3 (24.4–26.5)	24.1 (23.9–26.3)
	Poor prognosis	25.3 (21.9–28.7)	26.8 (22.2–31.7)	27.6 (26.3–29.0)	27.3 (25.9–28.7)
	<i>P</i> -value	0.29	0.23	0.04	0.01
ELH_av	Good prognosis	0.46 (0.33–0.59)	0.44 (0.33–0.54)	0.49 (0.46–0.52)	0.51 (0.46–0.58)
	Poor prognosis	0.39 (0.23–0.51)	0.40 (0.27–0.52)	0.44 (0.42–0.47)	0.42 (0.36–0.48)
	<i>P</i> -value	0.35	0.25	0.04	0.01
Carcinoma of the cervix					
DEL_av	Good prognosis	27.6 (25.7–28.5)	29.1 (25.9–31.5)	29.5 (25.1–28.3)	30.1 (26.9–31.7)
	Poor prognosis	26.3 (21.9–30.7)	28.8 (22.2–31.7)	24.8 (22.9–25.9)	26.8 (25.2–27.7)
	<i>P</i> -value	0.35	0.33	0.05	0.01
ELH_av	Good prognosis	0.28 (0.28–0.59)	0.33 (0.26–0.47)	0.34 (0.32–0.37)	0.31 (0.28–0.34)
	Poor prognosis	0.34 (0.22–0.55)	0.35 (0.28–0.42)	0.41 (0.38–0.44)	0.39 (0.36–0.41)
	<i>P</i> -value	0.39	0.23	0.05	0.0001

1000 cells in carcinoma of the prostate for the structural feature DEL\_av (Table 5). The CV represents a mean of 20 computations in separate slides from a case with good prognosis (relapse free survival more than 10 years after radical prostatectomy). A similar pattern was seen in a case with poor prognosis, i.e., with a relapse free survival less than 3 years after radical prostatectomy (Table 5, Panel B).

### 3.9. Comparing the prognostic values of ELH\_av and DEL\_av at increasing number of objects included in the analysis

We have previously presented the prognostic value of these two structural features [26]. When analysing a limited number of cells (100–500), no significant prognostic value was obtained. However, when increas-



Table 8

The prognostic in altogether 30 cases of carcinomas of the prostate of 6 structural features as related to the CV at 20 repeated computations. The number of cells included that yield a CV  $\approx 5\%$  (Panel A) gives a significant prognostic value ( $P < 0.05$ ). For all the structural features except ELH\_av, the prognostic value is no longer significant at a CV of 10% (Panel B)

Structural feature*	Carcinoma of the prostate good prognosis ( $n = 15$ )	Carcinoma of the prostate poor prognosis ( $n = 15$ )	$P$ -value Student's $t$ -test
A Prognostic value at CV $\approx 5\%$			
RF_dis	0.61	0.74	0.01
A_dis	0.55	0.35	0.02
DEL_av	27.1 <sup>†</sup>	25.8	0.001
ELH_av	0.41	0.32	0.001
NNRR	29.8	33.8	0.01
RMPB	13.4 <sup>†</sup>	16.2	0.03
B Prognostic value at CV $\approx 10\%$			
RF_dis	0.61	0.74	0.09
A_dis	0.55	0.35	0.12
DEL_av	27.1 <sup>†</sup>	25.8	0.07
ELH_av	0.41	0.32	0.01
NNRR	29.8	33.8	0.13
RMPB	13.4 <sup>†</sup>	16.2	0.15

\*For further description of the structural features, see [26].

ing the number of cells included in the analysis, these structural features attain a significant prognostic value (Table 7).

### 3.10. Prognostic value as related to the accepted CV

The prognostic value in prostatic carcinoma of 6 structural features was no longer significant as judged by the Student's  $t$ -test performed on two outcome groups (good and poor prognosis) when the CV increases from 5 to 10% (Table 8). The same pattern was observed for the other 10 structural features that we previously have shown to have a prognostic potential [26].

## 4. Discussion

This study investigated the numerical requirements as to how many objects are required to obtain an acceptable degree of reproducibility in repeated computations on a specific set of structural features. A prerequisite for prognostication in pathology is that repeated observations on a given entity do not show a coefficient of variation (CV) beyond a certain limit. A reasonable indication of where this limit lies, is the ability of a computed structural feature to predict the outcome of a lesion. We found that for our specific set of structural features, the CV critically depended on the number of cells included in the computations and for no structural feature was there an CV of 5% until at least 850 cells were included in the computations. For 13 of the 27 investigated structural features, a minimum of 1500 cells had to be included in the computations in order to achieve reproducible results. When analysing a limited number of cells (100–500), no significant prognostic value was found. However, when increasing the number of cells included in the analysis, these structural features attain a prognostic value (Table 7). The prognostic value of our proposed structural features was diminished when the CV of repeated computations changes from 5 to 10% (Table 8). Our study included relatively few cases, with few (8–10) computational runnings for any number of cells included in each specimen. We nevertheless observe a substantial variation in the computed values of the structural features when analysing less than 500–1000 objects. There is no reason to assume that this variation would decrease after an increasing number of repetitions have been performed. Most likely, an increase in the CV would be observed. Therefore, the numerical requirements we propose most likely represent an underestimate.

We have chosen cell nuclei as the object of interest, and our results apply specifically to analysis based on identification of cell nuclei from carcinomas of the tongue, larynx, cervix and prostate. However, there is no reason to assume that computations based on identification of other objects (e.g., lumina of vessels, nucleolar organising regions, fragments of chromatin, etc.) should not show similar requirements. Our segmenting algorithm represents an optimisation between precision and speed. Inevitably, increasing the speed of the computations for the segmentation part will reduce the precision of the segmentation, hence the structural analysis. Cell nuclei may be left out, and structures not corresponding to nuclei may be included. This is in analogy to reducing the number of cells for the analysis. Failure to include cell nuclei or inclusion of non-nuclear structures may yield values of the structural features that do not reflect the architecture and biology of the tissues.

We have not investigated systematically the variation in feature values for number far exceeding 5000 objects. Conceivably, these values could change con-

siderably when the scale of the analysis was increased to, e.g., 15 000 cells or more. However, for all the structural features investigated, there was an interval between 1500 and 10,000 cells where the values of the extracted features are reproducible to an acceptable degree. Most likely, this is a number of cells (or other structural components) that adequately represents the tissue in question. In addition, very few of the investigated tissue blocks contain more than 10–15,000 epithelial cells. If reproducible values are not acquired after 5–10,000 cells are analysed, most likely the use of the structural feature is not feasible.

Theoretically, in an invasive carcinoma, an island consisting of no more than 1–3 layers of cells can be detected and be of prognostic significance. By graph theory based methods, all but one layer of cells would be eliminated as a result of border effects. This makes the analysis without meaning, as graph theory based methods relate to neighbourhood relations between a number of cells. However, when numerous islands of epithelial cells are available for to analysis, data may be integrated over all the islands included in the analysis.

Whether the numerical requirements on a structural feature can be predicted a priori may depend on how the feature is defined. Thus, features that are based on the Voronoi polygons of subgraphs of the VD could possibly require more nuclei included in the analysis than features based on straight averages. However, the actual feature values may also be modified by properties of the tissue investigated, and there is no obvious way to predict how the feature values are influenced. Therefore, the numerical requirements on structural features must be investigated in a learning set when new tissue types are explored.

We have developed and tested 27 structural features on tissue architecture. All features were derived from the Voronoi Diagram and its subgraphs. There is no reason to assume that these structural features differ in principle from any other structural features that may be derived from these graphs. We therefore conclude that graph theory based structural analysis of biological entities should include a number of cells adequate to yield a CV of 5% or less, and that the adequate number of objects must be estimated experimentally for each specific tissue studied. For some structural features, this may imply including several thousand objects of interest.

## Acknowledgements

The present work was supported by a grant from the Norwegian Cancer Society. Discussions with Prof. E. Skovlund on the statistical analysis, are acknowledged.

## References

- [1] G. Anneroth and L.S. Hansen, A methodologic study of histologic classification and grading of malignancy in oral squamous cell carcinoma, *Scand. J. Dent. Res.* **92** (1984), 448–468.
- [2] G. Anneroth, J. Batsakis and M. Luna, Review of the literature and a recommended system of malignancy grading in oral squamous cell carcinomas, *Scand. J. Dent. Res.* **95** (1987), 229–249.
- [3] M. Bryne, H.S. Koppang, R. Lilleng, T. Stene, G. Bang and E. Dabelsteen, New malignancy grading is a better prognostic indicator than Broders' grading in oral squamous cell carcinomas, *J. Oral Pathol. Med.* **18** (1989), 432–437.
- [4] T. Bundgaard, S.M. Bentzen, J. Wildt, F.B. Sorensen, H. Sogaard and J.E. Nielsen, Histopathologic, stereologic, epidemiologic, and clinical parameters in the prognostic evaluation of squamous cell carcinoma of the oral cavity, *Head Neck* **18** (1996), 142–152.
- [5] A. Karabulut, J. Reibel, M.H. Therkildsen, F. Praetorius, H.W. Nielsen and E. Dabelsteen, Observer variability in the histologic assessment of oral premalignant lesions, *J. Oral Pathol. Med.* **24** (1995), 198–200.
- [6] J.A. Morris, Information and observer disagreement in histopathology, *Histopathology* **25** (1994), 123–128.
- [7] D.C. Aziz and R.B. Barathur, Quantitation and morphometric analysis of tumors by image analysis, *J. Cell. Biochem. (Suppl.)* **19** (1994), 120–125.
- [8] F. Darro, A. Kruczynski, C. Etievant, J. Martinez, J.L. Pasteels and R. Kiss, Characterization of the differentiation of human colorectal cancer cell lines by means of Voronoi diagrams, *Cytometry* **14** (1993), 783–792.
- [9] C. Godreche, I. Kostov and I. Yekutieli, Topological correlations in cellular structures and planar graph theory, *Phys. Rev. Letters* **69**(2) (1992), 2674–2677.
- [10] K. Kayser and W. Schlegel, Pattern recognition in histopathology: basic considerations, *Methods Inf. Med.* **21** (1982), 5–22.
- [11] K. Kayser and H. Stute, Minimum spanning tree, Voronoi's tessellation and Johnson–Mehl diagrams in human lung carcinoma, *Pathol. Res. Pract.* **185** (1989), 729–734.
- [12] K. Kayser, K. Sandau, G. Bohm, K.D. Kunze and J. Paul, Analysis of soft tissue tumors by an attributed minimum spanning tree, *Anal. Quant. Cytol. Histol.* **13** (1991), 329–334.
- [13] K. Kayser, O. Hagemeyer and T. Runtsch, Morphologic lesions in non-neoplastic bronchial mucosa associated with bronchial carcinomas, *Zentralbl. Pathol.* **137** (1991), 425–429.
- [14] K.K. Kayser, B. Kiefer, H. Toomes and H.U. Burkhard, Analysis of adenomatous structures in histopathology, *Anal. Quant. Cytol. Histol.* **9** (1987), 273–278.

- [15] E. Raymond, M. Raphael, M. Grimaud, L. Vincent, J.L. Binet and F. Meyer, Germinal center analysis with the tools of mathematical morphology on graphs, *Cytometry* **14** (1993), 848–861.
- [16] R. Kiss, I. Camby, I. Salmon, P. Van Ham, J. Brotchi and J.L. Pasteels, Relationship between DNA ploidy level and tumor sociology behavior in 12 nervous cell lines, *Cytometry* **20**(1) (1995), 118–126.
- [17] J.M. Prewitt and M.L. Mendelsohn, The analysis of cell images, *Ann. N.Y. Acad. Sci.* **128**(31) (1966), 1035–1053.
- [18] S.M. Patra and S. Vishveshwara, Backbone cluster identification in proteins by a graph theoretical method, *Biophys. Chem.* **84**(14) (2000), 13–25.
- [19] R. Samudrala and J. Moult, A graph-theoretic algorithm for comparative modeling of protein structure, *J. Mol. Biol.* **279**(29) (1998), 287–302.
- [20] C.C. Maley, DNA computation: theory, practice, and prospects, *Evol. Comput.* **6** (1998), 201–229.
- [21] K. Kayser and H.J. Gabius, Graph theory and the entropy concept in histochemistry. Theoretical considerations, application in histopathology and the combination with receptor-specific approaches, *Prog. Histochem. Cytochem.* **32** (1997), 1–106.
- [22] G.T. Toussaint, Pattern recognition and geometrical complexity, in: *Proceedings of the 5th International Conference on Pattern Recognition*, IEEE Catalogue No 80CH1499-3, 1980, pp. 1324–1347.
- [23] R. Marcelpoil, Normalization of the minimum spanning tree, *Anal. Cell. Pathol.* **5** (1993), 177–186.
- [24] D. Cheriton and R.E. Tarjan, Finding minimum spanning trees, *SIAM J. Comput.* (1976), 724–742.
- [25] S. Ulam, Patterns of growth of figures: mathematical aspects, in: *Proceedings of Symposia in Applied Mathematics, XIV: Mathematical Problems in the Biological Sciences*, American Mathematical Society, 1962, pp. 64–65.
- [26] J. Sudbø, R. Marcelpoil and A. Reith, New algorithms based on the Voronoi Diagram applied in a pilot study on normal mucosa and carcinomas, *Anal. Cell. Pathol.* (2000), submitted.
- [27] J. Serra, *Image Analysis and Mathematical Morphology*, London, Academic Press, 1982.