Digital analysis of colposcopic images in papillomavirus infection

A Dushkin, M Afanasiev, S Afanasiev and T Grishacheva,

1 The Loginov Moscow Clinical Scientific Center is State Institution funded by Moscow Health Department, 86 Enthusiastov highway, Moscow, Russian Federation, 111123
2 I.M. Sechenov First Moscow State Medical University, 8/2 Trubetskaya str., Moscow, Russian Federation, 119991
3 First Pavlov State Medical University of St. Petersburg, L'va Tolstogo str. 6-8, Saint Petersburg, Russian Federation, 197022
4 G.N. Gabrichevsky Moscow Research Institute for Epidemiology and Microbiology, 10 Admiral Makarov str., Moscow, Russian Federation, 125212

E-mail: alex@drdushkin.ru

Abstract. Acetowhite epithelium (AWE) is one of the main clinical indicators of papillomavirus infection in squamous cell lesions. AWE has a different intensity in various degrees of the cervix papillomavirus infection. The digital approach gives the opportunity to estimate AWE intensity by numeric values. The investigation aim is a quantitative assessment of cervical surface changes in papillomavirus infection with digital analysis and computer technologies. The spread of AWE on the cervical surface area has a value of the cervix papillomavirus infection in squamous cell lesions. These two features provide to create the index intensity (IndInt) and index grey value (IndGV). Open-source software ImageJ was used to analysis of colposcopic images. The 8-bit mode was used for the estimation of the AWE grey value. The algorithm of digital analysis detected indicators that provide the severity of papillomavirus infection. The outcomes of an algorithm were the identification of the cervix surface condition severity by an objective quantification.

1. Introduction
Colposcopy investigation includes the check-up of the reproductive system. To estimate the cervical surface used an acetic test with 3–5% acetic acid and Schiller's test. Acetic acid denatures intracellular and nuclear proteins of HPV-damaged epithelial cells [2]. It is caused by the formation acetowhite epithelium (AWE) on the cervical surface. The main disadvantage of colposcopy is a subjective assessment of intensity AWE. Digital analysis is a method for images estimation with graphic software and graphic processing unit (GPU). Modern video recording devices with high-resolution lens record procedures and investigations. Video-colposcope makes colposcopy and video-recording simultaneously. These video files are used for analysis by GPU and graphic software. Some graphic software provides the ability to convert RGB-type images into 8-bit images. ImageJ is a software for this procedure [1]. The 8-bit function of ImageJ software used a grey value (GV) tint to create an image. Intensity of AWE demonstrates a degree of squamous cell lesion. Mosaic and puncture are predictors of malignant tissues. These features of acetic acid were used to estimate the square of the cervical surface. The GV on the cervical surface images has a range of bits from 0 to 255. Measures of 255 and 0 mean white and black respectively. Measures of pixels provide a possibility to estimate colposcopic images with a quantitative approach. This approach decreases subjective assessment. The aim is a quantitative assessment of cervical surface changes in papillomavirus infection with digital analysis and computer technologies.
2. Digital approach
Digital approach based on three main steps. The first step is a preparing colposcopic image. Images made by screenshot function in recorded video of colposcopy investigation after test with 3–5% acetic acid. Video was recorded in FullHD dimension (1080×1920 pixels). Video-colposcope "Kernel KN-2200-A" used for recording of colposcopy investigation.

The second step is a converting RGB image to 8-bit image (shown in figures 1, 2). Screenshots were uploaded to ImageJ 1.52q software. “Type” function was used to 8-bit mode. “Plugins” function, “3D” subfunction and “Interactive 3D Surface Plot” subsubfunction were used to create the interactive plots and evaluate visual changes by “Spectrum LUT” feature. This feature helps to visualize digital tissue heterogeneity (shown in figure 3).

Figure 1. RGB image of cervical surface. 9.22X image enlargement.

Figure 2. 8-bit image. 9.22X image enlargement.

Figures 3. Interactive 3D Surface Plot. Digital tissue heterogeneity.
The third step is a calculation of index intensity (IndInt) and index grey value (IndGV). This step is shown in figures 4–8. Function “Polygon selection” used to calculate area square. IndInt is a ratio of pathological tissue area (%Area) to healthy tissue area (Area). Measures of area in pixels were converted measures of area in cm² by “Analyze” function and “Set Scale…” subfunction. IndGV is a multiplication IndInt to dGV (1). Value of GV calculated by “Image” function, “Adjust” subfunction, “Threshold” subfunction. dGV is a deviation between MAX (255 bit – white color) and MIN (0 bit – black color) measures of “Threshold” subfunction. Maximal (MAX) and minimal (MIN) measures of AWE should be used from “Threshold” window. Changes of AWE area have a different GV in an 8-bit mode image.

\[
\text{IndGV} = \left( \frac{\% \text{Area}}{100} / A \right) \cdot (\text{MAX} - \text{MIN}) \quad (1)
\]

Figure 4. Converting from pixels² to cm² by “Set Scale…” function.

Figure 5. Cervical surface area by “Polygon selections” function selection.

Figure 6. “Threshold” function selection.
2.1 Equipment for digital approach
Video-colposcope "Kernel KN-2200-A" used for recording of colposcopy investigation. Macbook Pro (16-inch, 2019) with processor – 2.3 GHz 8-core Intel Core i9, main memory – 16 GB 2667 MHz DDR4, graphics Intel UHD Graphics 630 1536 MB and AMD Radeon 5500M 4GB was used for digital analysis.

2.2 Statistical analysis
The IndGV data were assessed by MS office 365 (Excel v16.46) and RStudio v1.2.5042 statistical packages. Student's t-test (p<0.05) and descriptive statistics with confidential intervals (CI95%) used to estimate results.

3. Results
Digital approach detected statistically significant results between patient’s groups and subgroups with papillomavirus infection manifestations.

3.1. Results between groups
IndGV has an equal mean (M) value between groups with PVI and CIN II. It increases from PVI to CIN I, from CIN II to CIN III and decreases from CIN I to CIN II, from CIN III to cervical cancer. IndGV helps to identify the cervical papillomavirus infection as undulation infection leading to marinization of the cervix (shown in table 1). PVI and CIN II can be attributed to key stages in the transition to cervical dysplasia and malignancy.

| Table 1. IndGV between groups. |
|-----------------------------|
| n   | PVI | CIN I | CIN II | CIN III | CIS | ICC |
|-----|-----|-------|--------|---------|-----|-----|
| M   | 67  | 76    | 80     | 90      | 30  | 16  |
| SD  | 2.9 | 3.15  | 3.51   | 1.01    | 4.9 | 4.46|
| CI95% (±) | 3.48 – 4.91 | 3.97 – 5.39 | 3.34 – 4.88 | 4.89 – 6.91 | 3 – 6.5 | 2.8 – 7.18 |

PVI – papillomavirus infection without squamous intraepithelial neoplasia; CIN – cervical intraepithelial neoplasia; CC – cervical cancer.

Figure 7. Detection of range grey value by “Threshold” function.

Figure 8. Automatically calculation of measures to put on the formula (1).
Digital approach provides to divide groups by visual changes. There are patients with AWE, AWE + mosaic + punctation and without AWE (shown in tables 2–4). Approximation trend lines ($R^2$) for groups with/without AWE are 0.605 and 0.3585 respectively. Approximation trend line ($R^2$) for groups with AWE + P + M is 0.8329.

### Table 2. IndGV between groups with visual changes.

|     | PVI | CIN I | CIN II | CIN III | CIS | ICC |
|-----|-----|-------|--------|---------|-----|-----|
| 1   | n   | 7     | 7      | 7       | 6   | 2   |
|     | M   | 1.45  | 0.58   | 0.65    | 0.54 | 0.99 | 0.53 |
|     | SD  | 1.4   | 0.45   | 0.32    | 0.33 | 0.92 | 0.73 |
|     | CI95% (±) | 0.42 – 2.48 | 0.25 – 0.92 | 0.42 – 0.89 | 0.27 – 0.8 | -0.29 – 2.26 | -0.49 – 1.55 |
| 2   | n   | 21    | 30     | 37      | 39  | 17  | 8   |
|     | M   | 4.94  | 3.3    | 3.14    | 3.49 | 2.47 | 3.6  |
|     | SD  | 3.06  | 2.42   | 2.94    | 2.19 | 0.91 | 1.88 |
|     | CI95% (±) | 3.62 – 6.25 | 2.43 – 4.16 | 2.19 – 4.19 | 2.8 – 4.17 | 2.03 – 2.9 | 2.3 – 4.9 |
| 3   | n   | 39    | 39     | 36      | 45  | 11  | 6   |
|     | M   | 4.3   | 6.48   | 5.78    | 8.71 | 8.95 | 8.32 |
|     | SD  | 2.9   | 2.66   | 3.55    | 5.19 | 6.09 | 5.52 |
|     | CI95% (±) | 3.38 – 5.2 | 5.64 – 7.31 | 4.62 – 6.94 | 7.2 – 10.23 | 5.38 – 12.55 | 3.91 – 12.74 |

LPVI – latent papillomavirus infection; APVI – acute papillomavirus infection; CIN – cervical intraepithelial neoplasia; CIS – cancer in situ; ICC – invasive cervical cancer; WAPVI – with acute papillomavirus infection; WOAPVI - without acute papillomavirus infection; WC – with condyloma; WOC – without condyloma; 1 – without AWE; 2 – with AWE; 3 – with AWE+punctuation+mosaic.

IndGV has a direct linear relationship in groups with AWE + P + M and inverse linear relationship in groups with/without relationship. Acetowhite epithelium with punctation and mosaic has more intensity than acetowhite epithelium. Intensity AWE depends on clinical stage of papillomavirus infection.

#### 3.2. Results between subgroups

Groups were divided into subgroups in according to clinical manifestations of papillomavirus infection (shown in table 1). IndGV in subgroups of PVI and CC groups have no differences in mean (M) value (shown in table 3). Subgroups of CIN I-III groups have difference of IndGV in mean (M) value (shown in table 3). Subgroup division based on visual changes: 1 – without AWE; 2 – with AWE and 3 – with AWE + P + M.

### Table 3. IndGV between subgroups.

|     | PVI | CIN I | CIN II | CIN III |
|-----|-----|-------|--------|---------|
|     | n   | 14    | 53     | 55      |
|     | M   | 4.03  | 4.24   | 3.92    |
|     | SD  | 2.02  | 3.19   | 2.75    |
|     | CI95% (±) | 2.96 – 3.37 | 3 – 4.83 | 4.3 – 6.35 |

LPVI – latent papillomavirus infection; APVI – acute papillomavirus infection; CIN – cervical intraepithelial neoplasia; WAPVI – with acute papillomavirus infection; WOAPVI - without acute papillomavirus infection; WC – with condyloma; WOC – without condyloma.

Approximation trend line ($R^2$) for groups with AWE + P + M is 0.7234. Subgroups with/without AWE tend to low intensity of AWE on the cervical surface. Approximation trend line ($R^2$) for groups with/without AWE are 0.0652 and 0.1559 respectively. These visual changes have a weak approximation trend line. Visual changes and clinical manifestations have a weak linear correlation.

Subgroups with AWE + P + M tend to high intensity of AWE on the cervical surface (shown in table 4). Ambiguous results were in subgroups with AWE.

---

5
Table 4. IndGV between subgroups with visual changes.

|        | LPVI | APVI | WOAPVI | WAPVI | WOC | WC | WC |
|--------|------|------|--------|-------|-----|----|----|
|        |      |      |        |       |     |    |    |
| CIN I  | n    | 4    | 3      | 7     | 0   | 7  | 0  |
|        | M    | 2.06 | 0.63   | 0.58  | 0   | 0.65 | 0   | 0.54 | 0 |
|        | SD   | 0.51 | 0.8    | 0.45  | 0   | 0.32 | 0   | 0.33 | 0 |
| CI95%  | ±    | 0.58 – 3.54 | -0.27 – 1.54 | 0.25 – 0.92 | 0 | 0.42 – 0.89 | 0 | 0.27 – 0.8 | 0 |
| CIN II | n    | 6    | 15     | 17    | 13  | 27  | 10 |
|        | M    | 4.83 | 4.98   | 4.15  | 2.17 | 2.88 | 3.82 | 3.22 | 5.81 |
|        | SD   | 1.13 | 3.6    | 2.71  | 1.41 | 2.85 | 3.22 | 1.96 | 3.01 |
| CI95%  | ±    | 3.93 – 5.73 | 3.15 – 6.8 | 2.86 – 5.44 | 1.4 – 2.94 | 1.81 – 3.96 | 1.83 – 5.82 | 2.57 – 3.87 | 2.86 – 8.77 |
| CIN III| n   | 4    | 35     | 11    | 28  | 19  | 17 |
|        | M   | 4.39 | 4.23   | 5.68  | 6.79 | 5.91 | 5.64 | 8.15 | 10.69 |
|        | SD  | 2.6  | 2.97   | 1.6   | 2.94 | 4.26 | 4.26 | 4.61 | 6.78 |
| CI95%  | ±   | 2.36 – 7.21 | 3.24 – 5.22 | 4.73 – 6.63 | 5.7 – 7.88 | 4.7 – 8.83 | 4.37 – 6.9 | 6.62 – 9.67 | 6.49 – 14.9 |

LPVI – latent papillomavirus infection; APVI – acute papillomavirus infection; CIN – cervical intraepithelial neoplasia; WAPVI – with acute papillomavirus infection; WOAPVI - without acute papillomavirus infection; WC – with condyloma; WOC – without condyloma.

4. Conclusion

Clinical diagnostics include HPV testing, liquid-based cytology/histological investigation and colposcopy. A general inspection by a gynecologist used a colposcopic investigation with an acetic test and Schiller’s test. The intensity of AWE is used as a pathological process indicator after the acetic test. Ruan, Yetian et al. conclude that the area of AWE has different intensity, contour and surface [7]. These indicators have different visual changes in HPV infection, CIN I-III and cervical cancer [6]. Digital analysis of colposcopy images gives the ability to estimate visual changes of AWE with a quantitative approach. It makes an objective evaluation of clinical manifestations in HPV infection. The quantification approach detects the severity of AWE of the cervical surface. IndGV correlates with CIN stages and clinical manifestations. The clinical manifestations of HPV infection in CIN have a direct dependence on digital measurement as grey value (GV). These indicators give the opportunity to evaluate clinical manifestations in digital analysis. Digital analysis is methodic in real-time mode. IndGV might be used as a predictive factor of the severity of HPV infection in CIN I-III, cervical cancer. Approach cannot be used for detecting and predicting papillomavirus infection outcomes in patients without visual changes. Patients with AWE visual changes should be clearly investigated. Digital approach might be used as additional method for colposcopy, which provide the objective assessment intensity of AWE and can be used as indicator for prediction of papillomavirus manifestation.

Acknowledgments

Authors declare no conflict of interests. No funding sources supported this research.

References

[1] Rasband W S 2019 ImageJ, U. S. National Institutes of Health, Bethesda, Maryland, USA, https://imagej.nih.gov/ij/
[2] Bauer H 2015 Color Atlas of Colposcopy (Moscow: GEOTAR-Media) p 288
[3] Dushkin A et al. Application for an invention №2020120982 «Method for determining the degree of the cervical dysplasia» (priority 25.06.2020 RU).
[4] Rauch T et al. 2020 Discrimination analysis of breast calcifications using x-ray dark-field radiography Med Phys. 47(4) 1813–26 doi:10.1002/mp.14043
[5] Natalahle B et al. 2013 World Health Organization guidelines for screening and treatment of precancerous lesions for cervical cancer prevention (Geneva: World Health Organization) p 40
[6] Girardi F, Reich O and Tamussino K 2015 Burkhardt’s Colposcopy and Cervical Pathology (Germany: Thieme Verlagsguppe 4th ed) p 240
[7] Ruan Y, Liu M, Guo J, Zhao J, Niu S and Li F 2020 Evaluation of the accuracy of colposcopy in detecting high-grade squamous intraepithelial lesion and cervical cancer Arch Gynecol Obstet. 302(6) 1529–38 doi:10.1007/s00404-020-05740-x