

Clinical progress of human papillomavirus genotypes and their persistent infection in subjects with atypical squamous cells of undetermined significance cytology: Statistical and latent Dirichlet allocation analysis

YEE SUK KIM¹, SUNGIN LEE², NANSU ZONG³ and JIMIN KAHNG⁴

¹Department of Anesthesiology and Pain Medicine, College of Medicine, The Catholic University of Korea, Bucheon St. Mary's Hospital, Bucheon, Kyunggi-do 14647; ²Biomedical Knowledge Engineering Laboratory, School of Dentistry, Seoul National University, Seoul 08826; ³Department of Biomedical Informatics, School of Medicine, University of California, San Diego, CA 92093, USA; ⁴Department of Laboratory Medicine, College of Medicine, The Catholic University of Korea, Bucheon St. Mary's Hospital, Bucheon, Kyunggi-do 14647, Republic of Korea

Received October 1, 2015; Accepted November 4, 2016

DOI: 10.3892/etm.2017.4295

Abstract. The present study aimed to investigate differences in prognosis based on human papillomavirus (HPV) infection, persistent infection and genotype variations for patients exhibiting atypical squamous cells of undetermined significance (ASCUS) in their initial Papanicolaou (PAP) test results. A latent Dirichlet allocation (LDA)-based tool was developed that may offer a facilitated means of communication to be employed during patient-doctor consultations. The present study assessed 491 patients (139 HPV-positive and 352 HPV-negative cases) with a PAP test result of ASCUS with a follow-up period ≥ 2 years. Patients underwent PAP and HPV DNA chip tests between January 2006 and January 2009. The HPV-positive subjects were followed up with at least 2 instances of PAP and HPV DNA chip tests. The most common genotypes observed were HPV-16 (25.9%, 36/139), HPV-52 (14.4%, 20/139), HPV-58 (13.7%, 19/139), HPV-56 (11.5%, 16/139), HPV-51 (9.4%, 13/139) and HPV-18 (8.6%, 12/139). A total of 33.3% (12/36) patients positive for HPV-16 had cervical intraepithelial neoplasia (CIN)2 or a worse result, which was significantly higher than the prevalence of CIN2 of 1.8% (8/455) in patients negative for HPV-16 ($P < 0.001$), while no significant association was identified for other genotypes in terms of genotype and clinical progress.

There was a significant association between clearance and good prognosis ($P < 0.001$). Persistent infection was higher in patients aged ≥ 51 years (38.7%) than in those aged ≤ 50 years (20.4%; $P = 0.036$). Progression from persistent infection to CIN2 or worse (19/34, 55.9%) was higher than clearance (0/105, 0.0%; $P < 0.001$). In the LDA analysis, using symmetric Dirichlet priors $\alpha = 0.1$ and $\beta = 0.01$, and clusters (k) = 5 or 10 provided the most meaningful groupings. Statistical and LDA analyses produced consistent results regarding the association between persistent infection of HPV-16, old age and long infection period with a clinical progression of CIN2 or worse. Therefore, LDA results may be presented as explanatory evidence during time-constrained patient-doctor consultations in order to deliver information regarding the patient's status.

Introduction

The precancerous stage of cervical cancer is 7-20 years, after which it may progress to invasive cancer. Due to this long latency period, there is ample time for early detection (1). The association between human papillomavirus (HPV) infection as a primary etiological cause of cervical cancer has been well established (1-4). At present, Papanicolaou (PAP) is the most prevalent screening method; ~10% of the women subjected to PAP present with atypical squamous cells of undetermined significance (ASCUS) and 5-10% exhibit malignant lesions (5). Women with high-risk HPV genotypes and persistent HPV infection have an increased risk of being diagnosed with high-risk lesions following biopsy (6,7), but for Korean patients, limited studies exist regarding this. Few studies have been completed on the clinical progress and prognosis of patients with HPV infection with regard to the effects of high-risk HPV genotype infection in patients presenting with ASCUS whose follow-up or biopsy results exhibited low-grade squamous intraepithelial lesions or a better result than cervical intraepithelial neoplasia (CIN)1 (8).

Correspondence to: Dr Jimin Kahng, Department of Laboratory Medicine, College of Medicine, The Catholic University of Korea, Bucheon St. Mary's Hospital, 327 Sosa-ro, Wonmi-gu, Bucheon, Kyunggi-do 14647, Republic of Korea
E-mail: jmkahng@catholic.ac.kr

Key words: human papillomavirus, atypical squamous cells of undetermined significance, latent Dirichlet allocation, high-risk lesions, patient-doctor communication

Regular PAP and/or HPV DNA tests for cervical cancer greatly increase the chance of early detection of lesions. This in turn leads to minimal surgical intervention and promotes longevity and increased quality of life for patients (9-11). However, in a number of countries, regular visits to an obstetrician or gynecologist are not feasible due to cultural or economic barriers. This is the case throughout Korea, apart from metropolitan areas and in patients who were diagnosed with cervical cancer or metastatic cancer at their first visit (12).

Communicating with patients who are potentially HPV-positive and informing them of the association between HPV infection and cervical cancer is difficult, particularly when time is limited. In secondary and tertiary hospitals in Korea, patient-doctor consultations are often short and limited; thus it is difficult to persuade patients to visit practitioners more frequently. Presentation of explanatory evidence, which aids in more effective communication between doctors and potential HPV-positive patients, may increase patient compliance with doctors' advice following regular follow-up sessions, as well as helping to build a rapport between practitioners and their patients.

In the present study, a total of 491 patients whose initial PAP results revealed the presences of ASCUS and who had been subjected to follow-up examinations for ≤ 7 years, were assessed in order to identify any significant differences in prognoses regarding HPV infection, persistent infection and genotype variations. The present study used traditional statistical analysis as well as a probabilistic topic model called latent Dirichlet allocation (LDA) (13). The LDA analysis was specifically used to assess its significance as a tool for facilitated communication between doctors and patients.

Patients and methods

Patients. The present study consisted of 491 female subjects (139 HPV-positive and 352-negative cases) with a PAP result of ASCUS (14). The follow-up period was ≥ 2 years. Patients underwent simultaneous PAP and HPV DNA chip tests (HPVDNAChip; AGBIO Diagnostics Co., Ltd., Chuncheon, Korea) at Bucheon St. Mary's Hospital of the Catholic University of Korea (Bucheon, Korea) from January 2006 to January 2009. Patients positive for HPV were followed up with a minimum of two instances of PAP and HPV DNA chip tests with an average follow-up period of 1,297 days (range, 731-2,562 days).

The ages of the patients were as follows: 4 patients (0.8%) were aged ≤ 20 years, 33 (6.7%) were aged 21-30 years, 98 (20.0%) were aged 31-40 years, 240 (48.9%) were aged 41-50 years, 89 (18.1%) were aged 51-60 years, 21 (4.3%) were aged 61-70 years and 6 (1.2%) were aged ≥ 71 years. Therefore, 375 patients (76.4%) were ≤ 50 years old and 116 (23.6%) were > 50 years old. There was no significant difference in the HPV-positive or -negative rate among the age groups. The study protocol was approved by the institutional review board (IRB) of St. Mary's Hospital (The Catholic University of Korea, Bucheon, Korea) our institution and the need for informed consent was waived by the IRB, due to the retrospective nature of the present study.

Sample collection for PAP and HPV DNA chip test. A cytobrush was used to obtain samples for PAP tests and the

Bethesda system III (14) was used to assess the results. Using a sterilized vaginal speculum, cervical specimens were collected from each patient once per visit for DNA chip test by inserting a cytobrush attached to a HPVDNAChip sampler into endo- and exo-cervical areas. Once removed, the cytobrush was placed into transport medium (1X phosphate-buffered saline) and immediately refrigerated at 4°C.

DNA extraction, hybridization and chip scanning. DNA, which was extracted using the HPVDNAChip kit in compliance with the manufacturer's protocol, was subjected to immediate use or stored at 4°C. If the DNA was stored long-term (> 1 day), it was stored at -70°C. All HPV assays were performed using a PCR-based DNA microarray system (AGBIO Diagnostics Co., Ltd.) with an HPVDNAChip providing 22 HPV probes, comprising high-risk (HPV-15, -16, -18, -31, -33, -35, -39, -45, -51, -52, -56, -58, -59, -66, -68 and -69) and 7 low-risk (HPV-6, -11, -34, -40, -42, -43 and -44) types. The assays were performed following the manufacturer's protocol (15). A confocal laser scanner (GenePix 4000B, laser power 33%, PMT 880; Axon Instruments; Molecular Devices, LLC, Sunnyvale, CA, USA) was used to visualize the hybridized HPV DNA.

Statistical analysis. Fisher's exact test and the χ^2 -test were used to analyze associations between variables. For two-group comparisons, unpaired Student's *t*-tests were used for parametric analyses and Wilcoxon's rank sum test was used for nonparametric analyses. For comparisons of more than two groups, one-way analysis of variance was used for parametric analyses and the Kruskal-Wallis test for nonparametric analyses. $P < 0.05$ was considered to indicate a statistically significant difference. All statistical analyses were performed using SAS software, version 9.2 (SAS Institute, Inc., Cary, NC, USA).

LDA. LDA is a probabilistic topic model that assesses collections of text in order to identify common topics throughout the collections (13). LDA models accommodate heterogeneous data in the selection and characterization of intricate associations of data. In the present study, a document was prepared for each patient, which contained a group of clinical observations, including clinical measurements such as genotypes. LDA is a mixed-membership model (16), in which each patient is associated with multiple topics with varying probabilities, based on the probability distribution of the entire dataset. For example, if five topics are created for the HPV-positive group, each patient is associated with the topics with probabilities (for example, 20% to topic 1, 10% to topic 2, 15% to topic 3, 15% to topic 4 and 40% to topic 5). A topic is then associated with five clinical measurements, including age, progress status, genotype, clearance (natural healing) and duration of infection with varying proportions, to show the probability distributions of the terms to constitute the topic (for example, age between 20-30 with a probability of 20%, progress 'a' with 10% and genotype 16 with 40%, clear '0' with 10% and duration with 20%). A large body of works exists that have applied LDA to a number of tasks, including news article analysis (17), cancer prediction (18) and analysis of scientific ideas (19). However, limited studies have been directed towards identifying complex interactions of clinical measurements, as proposed in the present study, in association with HPV-positive patients. In the present study,

Table I. Clinical progression of HPV in each group (n=491).

Group	Normal (n)	Koilocytosis (n)	CIN1 (n)	CIN2 or worse (n)	Total (n)
HPV-negative	281	59	11	1	352
HPV-positive	73	30	17	19	139
Age (years)					
≤50	269	75	19	12	375
>50	85	14	9	8	116

HPV, human papillomavirus; CIN, cervical intraepithelial neoplasia.

Table II. Results of the human papillomavirus test by DNA chip to detect histologically confirmed high-grade squamous intraepithelial lesions or worse (PAP test result worse than CIN2).

Parameter	HPV-16	HPV-18	HPV-51	HPV-52	HPV-56	HPV-58
Sensitivity (%)	60.0	ND	ND	10.0	5.0	15.0
Specificity (%)	94.9	97.5	97.2	96.2	96.8	96.6
PPV (%)	33.3	ND	ND	10.0	6.3	15.8
NPV (%)	98.2	95.8	95.8	96.2	96.0	96.4

HPV, human papillomavirus; PPV, positive predictive value; NPV, negative predictive value; ND, not determined.

LDA analysis was performed only on the HPV-positive group (117 patients positive for HPV with ASCUS).

Results

Clinical progress. All patients were separated into four groups according to their final clinical progress: Normal according to histology or PAP (HPV-positive, n=281; HPV-negative, n=73), as well as biopsy specimens showing koilocytosis (HPV-positive, n=59; HPV-negative, n=30), CIN1 (HPV-positive, n=11; HPV-negative, n=17) or high-risk lesions (CIN2 or worse; HPV-positive, n=1; HPV-negative, n=19). Positive and negative groups and age groups are presented in Table I. In brief, CIN1 or worse was identified in 25.9% (36/139) of the HPV-positive group and 3.4% (12/352) of the HPV-negative group; and CIN2 or worse was identified in 13.7% (19/139) in the HPV-positive and 0.3% (1/352) in the HPV-negative group. This indicated that HPV-positive patients had a significantly poorer prognosis than HPV-negative patients ($P<0.001$). In the HPV-positive group, CIN1 or worse was more prevalent in subjects aged >50 years (45.2%, 14/31), when compared with those aged ≤50 years (20.4%, 22/108; $P=0.005$; Table I).

Genotype and clinical progress association. The most common genotypes of HPV detected were, in decreasing order of frequency, HPV-16 (25.9%, 36/139), HPV-52 (14.4%, 20/139), HPV-58 (13.7%, 19/139), HPV-56 (11.5%, 16/139), HPV-51 (9.4%, 13/139) and HPV-18 (8.6%, 12/139). There was no difference in infection rates among the age groups ($P>0.05$). A total of 33.3% (12/36) HPV-16-positive patients were classified as CIN2 or worse, which is significantly higher than the 1.8% (8/455) of HPV-16-negative patients with CIN2 or worse

($P<0.001$). For other genotypes, there was no significant association between genotype and clinical progression ($P>0.05$). The sensitivity, specificity and positive and negative predictive values of the six most frequently occurring HPV genotypes with a worse result than CIN2 are presented in Table II. These results show the association of these genotypes with high-risk lesion.

Association between clearance (natural healing), persistent infection or duration of infection and prognosis. During follow-up, a significant association between clearance and good prognosis was identified in the HPV-positive group via PAP test and biopsy results ($P<0.001$). Persistent infection was significantly higher in patients aged >50 years (38.7%) compared with those aged ≤50 years (20.4%; $P=0.036$). Progression from persistent infection to CIN2 or worse was higher (19/34, 55.9%) than clearance (0/105, 0.0%; $P<0.001$). Significant differences were observed between the duration of infection and prognosis. The average infection period of subjects with CIN1 or worse was 568.8 days, which was longer than in those naturally healed or in patients with koilocytosis, which was 393.3 days ($P<0.001$). However, there was no association between genotype and clearance ($P>0.05$).

Variables affecting clinical progress. Logistic regression analysis demonstrated that clinical progress was significantly associated with genotype and age of the patients. Specifically, in cases positive for HPV-16, the risk of progression to CIN1 or worse was higher than in cases negative for HPV-16 or infection with other genotypes with an odds ratio of 13.040 (95% confidence interval, 8.725-21.701). Patients aged >50 years had a higher chance of progression to CIN1 or worse than

patients aged ≤ 50 years with an odds ratio of 2.169 (95% confidence interval, 1.459-3.150).

LDA results. Tables III and IV present the results of the symmetric Dirichlet priors in the LDA estimation where $\alpha=0.1$, $\beta=0.01$ and clusters (k)=5 or 10. In brief, the 5A cluster was a group that primarily represented HPV-16-positive patients, continuous infection and CIN1 or worse. By contrast, the four other topics represented normal prognosis and koilocytosis under natural regression/healing of HPV infection. Furthermore, the 10C cluster exhibited a grouping of HPV-16-positive status, continuous infection and CIN2 or worse, and two 'meaningless' topics, which did not appear to be relevant to any clusters, were identified when $k=10$.

Discussion

Cervical cancer is the second leading cause of cancer worldwide (7). In Korea, it is the fourth most frequently occurring cancer type in women, following breast, stomach and colorectal cancer, and ~5,000 patients are diagnosed annually with cervical cancer in Korea (20). The PAP test and colposcopy provide information on existing lesions at the expense of a high false-negative rate and are not suitable as an indicator if patients have no lesions but exhibit other potential risks. Sasieni *et al* (21) demonstrated that 47% of patients with invasive cervical cancer received screening tests within five years prior to the date of diagnosis of the patient's cancer.

Since the role of HPV infection with regard to cervical cancer has been established, >120 HPV genotypes have been identified, 15 of which are classified as oncogenically high-risk types (HPV-16, -18, -31, -33, -35, -39, -45, -51, -52, -56, -58, -59, -66, -68 and -69) (7). HPV-16 is considered the most infectious type, followed by HPV-18 (7). However, in Asia, HPV-58 and HPV-52 are at times reported as the second most common genotype (22-24). In the present study, HPV-16, -52, -58, -56 and -51 were the most prevalent types, with HPV-18 being ranked sixth in frequency, which is in line with results of previous studies (22,23). The specificity and negative predictive values of the six most frequently occurring HPV genotypes regarding a PAP test result worse than CIN2 were in a clinically applicable range.

The present study demonstrated that HPV-16 is associated with high-risk lesions. Out of the 491 cases included in the present study, an insufficient number of cases with HPV-18 were identified and there was a limited number of patients presenting with persistent HPV-18 infection; therefore, the results of statistical analysis were not able to be used.

In addition to sexual activity, age has been reported to be associated with HPV infection. Higher rates of HPV infection are detected among women aged ≤ 35 years due to increased sexual activity. However, metaplastic activity in the squamocolumnar junction, which provides fertile ground for the infection of HPV, also increases with age (25). This metaplastic activity in the squamocolumnar junction provides an ideal area for growth and infection of pathogenic viruses, such as HPV, which is transmitted through sexual activity and may eventually lead to development of cervical cancer (25). In the present study, the prevalence and type of HPV infection did not vary between age groups (≤ 50 and >50); however, patients

aged >50 years exhibited an increased rate of progression to high-risk lesions.

The results of the LDA analysis may be provided to patients as supplementary information, as it demonstrates that HPV-16-positive patients with persistent infection are more likely to have a poor prognosis. Odds ratios determine the strength of the association between the presence or absence of certain features, but should not be used as explanatory evidence in patient-doctor consultations. The choice of the number of topics is a very important issue in LDA. The data collected in the present study demonstrated the most pertinent outcome was obtained when $k=5$ rather than $k=2$ or 10, in terms of explaining the association between the disease and its etiological factors.

In the present study, statistical and LDA analyses produced noteworthy results; the statistical results showed a strong correlation between persistent infection with HPV-16 and the presence of high-risk lesions in clinical progress in follow-up examinations. In addition, patients aged >50 years exhibited higher levels of persistent infection than younger patients and there was a correlation between disease clearance and good prognosis. These results were ascertained by the LDA results with a cluster generated from HPV-16, old age and persistent infection as factors, which, in combination, revealed an association with CIN2. This was not detected in the logistic regression due to an insufficient number of subjects with CIN2 or worse in the present study.

LDA as an analytical tool and data model may be beneficial when used in conjunction with statistical analysis that provides a definitive view on the data.

In conclusion, LDA and statistical analysis revealed significant associations between persistent infection of HPV-16, old age and a prolonged infection period with the clinical progression of CIN2 or worse. LDA was considered to be a good supplementary means of communication to deliver necessary information regarding a patient's status during a consultation between doctor and patient with a limited time.

References

1. Bosch FX, Manos MM, Muñoz N, Sherman M, Jansen AM, Peto J, Schiffman MH, Moreno V, Kurman R and Shah KV: Prevalence of human papillomavirus in cervical cancer: A worldwide perspective. International biological study on cervical cancer (IBSCC) Study Group. *J Natl Cancer Inst* 87: 796-802, 1995.
2. Khan MJ, Castle PE, Lorincz AT, Wacholder S, Sherman M, Scott DR, Rush BB, Glass AG and Schiffman M: The elevated 10-year risk of cervical precancer and cancer in women with human papillomavirus (HPV) type 16 or 18 and the possible utility of type-specific HPV testing in clinical practice. *J Natl Cancer Inst* 97: 1072-1079, 2005.
3. Wright TC Jr, Stoler MH, Sharma A, Zhang G, Behrens C and Wright TL: ATHENA (Addressing The Need for Advanced HPV Diagnostics) Study Group: Evaluation of HPV-16 and HPV-18 genotyping for the triage of women with high-risk HPV⁺ cytology-negative results. *Am J Clin Pathol* 136: 578-586, 2011.
4. Thomsen LT, Frederiksen K, Munk C, Junge J, Iftner T and Kjaer SK: Long-term risk of cervical intraepithelial neoplasia grade 3 or worse according to high-risk human papillomavirus genotype and semi-quantitative viral load among 33,288 women with normal cervical cytology. *Int J Cancer* 137: 193-203, 2015.
5. Solomon D, Schiffman M and Tarone B; ALTS Study Group: Comparison of three management strategies for patients with atypical squamous cells of undetermined significance: Baseline results from a randomized trial. *J Natl Cancer Inst* 93: 293-299, 2001.

6. Castle PE, Solomon D, Schiffman M and Wheeler CM: Human papillomavirus type 16 infections and 2-year absolute risk of cervical precancer in women with equivocal or mild cytologic abnormalities. *J Natl Cancer Inst* 97: 1066-1071, 2005.
7. de Freitas AC, Gurgel AP, Chagas BS, Coimbra EC and do Amaral CM: Susceptibility to cervical cancer: An overview. *Gynecol Oncol* 126: 304-311, 2012.
8. Gage JC, Schiffman M, Solomon D, Wheeler CM, Gravitt PE, Castle PE and Wentzensen N: Risk of precancer determined by HPV genotype combinations in women with minor cytologic abnormalities. *Cancer Epidemiol Biomarkers Prev* 22: 1095-1101, 2013.
9. Morris M: Management of stage IA cervical carcinoma. *J Natl Cancer Inst Monogr*: 47-52, 1996.
10. Morris M, Mitchell MF, Silva EG, Copeland LJ and Gershenson DM: Cervical conization as definitive therapy for early invasive squamous carcinoma of the cervix. *Gynecol Oncol* 51: 193-196, 1993.
11. Saslow D, Solomon D, Lawson HW, Killackey M, Kulasingam SL, Cain J, Garcia FA, Moriarty AT, Waxman AG, Wilbur DC, *et al*: American cancer society, American society for colposcopy and cervical pathology and American society for clinical pathology screening guidelines for the prevention and early detection of cervical cancer. *CA Cancer J Clin* 62: 147-172, 2012.
12. Kahng J, Kim EH, Kim HG and Lee W: Development of a cervical cancer progress prediction tool for human papillomavirus-positive Koreans: A support vector machine-based approach. *J Int Med Res* 43: 518-525, 2015.
13. Blei DM, Ng AY and Jordan MI: Latent Dirichlet allocation. *J Machine Learning Res* 3: 993-1022, 2003.
14. Solomon D, Davey D, Kurman R, Moriarty A, O'Connor D, Prey M, Raab S, Sherman M, Wilbur D, Wright T Jr and Young N; Forum Group Members; Bethesda 2001 Workshop: The 2001 Bethesda system: Terminology for reporting results of cervical cytology. *JAMA* 287: 2114-2119, 2002.
15. Lee GY, Kim SM, Rim SY, Choi HS, Park CS and Nam JH: Human papillomavirus (HPV) genotyping by HPV DNA chip in cervical cancer and precancerous lesions. *Int J Gynecol Cancer* 15: 81-87, 2005.
16. Blei DM: Probabilistic topic models. *Commun ACM* 55: 77-84, 2012.
17. Newman D, Chemudugunta C, Smyth P and Steyvers M: Analyzing entities and topics in news articles using statistical topic models. In: *Intelligence and Security Informatics*. Springer, pp93-104, 2006.
18. Korthauer K, Dawson J and Kendzierski C: Predicting cancer subtypes using survival-supervised latent Dirichlet allocation models. In: *Advances in Statistical Bioinformatics: Models and Integrative Inference for High-Throughput Data*. Do K-A, Qin ZS and Vannucci M (eds). Cambridge University Press, pp366-381, 2012.
19. Blei DM and Lafferty JD: A correlated topic model of Science. *Ann Applied Statistics* 1: 17-35, 2007.
20. Gynecologic Oncology Committee of the Korean Society of Obstetrics and Gynecology: Annual report of gynecologic cancer registry program in Korean for 2002 (Jan. 1st, 2002-Dec. 31st, 2002). *Korean J Obstet Gynecol* 47: 1029-1070, 2004.
21. Sasieni PD, Cuzick J and Lynch-Farmery E: Estimating the efficacy of screening by auditing smear histories of women with and without cervical cancer. The National Co-ordinating Network for Cervical Screening Working Group. *Br J Cancer* 73: 1001-1005, 1996.
22. Hwang HS, Park M, Lee SY, Kwon KH and Pang MG: Distribution and prevalence of human papillomavirus genotypes in routine pap smear of 2,470 Korean women determined by DNA chip. *Cancer Epidemiol Biomarkers Prev* 13: 2153-2156, 2004.
23. Kahng J and Lee HJ: Clinical efficacy of HPV DNA chip test in the era of HPV vaccination: 1,211 cases, a single institution study. *Korean J Lab Med* 28: 70-78, 2008.
24. Cho NH, An HJ, Jeong JK, Kang S, Kim JW, Kim YT and Park TK: Genotyping of 22 human papillomavirus types by DNA chip in Korean women: Comparison with cytologic diagnosis. *Am J Obstet Gynecol* 188: 56-62, 2003.
25. Burd EM: Human papillomavirus and cervical cancer. *Clin Microbiol Rev* 16: 1-17, 2003.