Development of a Simple and Accurate Polymerase Chain Reaction Screen for Human Papillomavirus in College Students

Emily E. Bishop[1], Owen R. Smith[2], Paul E. Richardson[2]

[1] *Department of Biology, Coastal Carolina University, Conway, SC*
[2] *Department of Chemistry, Coastal Carolina University, Conway, SC*

Human Papillomavirus (HPV) is, according to the CDC, the most common sexually transmitted infection (STI) with many infections developing in a person's late teens or early 20's. HPV is a nonenveloped, relatively small, icosahedral DNA virus that infects squamous epithelial cells. This infection can happen in genital, anal, and oral cavities. 90% of these infections will self-clear in two years according to the CDC. Of the 10% of infections that do not clear in two years, they could develop into carcinomas later in life. The goal of this developmental research is to create a polymerase chain reaction (PCR) that can detect HPV genomic material from a human oral self-sample. The methodology of the self-sample was determined through several trials using three anonymous volunteers.

Introduction

Human papillomavirus, (HPV) is a small, non-enveloped enveloped virus that's 62-66nm in diameter[1]. It is a cell-type-specific virus, targeting the keratinocytes on the basal lamina[1]. HPV is spread through scratching or sexual intercourse (vaginal-penile sex, penile-anal sex, penile-oral sex, and vaginal-oral sex)[1]. The virus replicates within the keratinocytes through a differentiation-dependent manner; the early genes, like the E7/E8 oncogenes, are expressed in undifferentiated keratinocyte's while the later genes occur in keratinocytes that are in high or terminal differentiation[1]. There are more than 460 subtypes of HPV[1]. Most of these infections can be cleared by the human immune system within 7 months to 3 years of initial infection[2]. About 40 of its subtypes are known to lead to carcinoma development[1]. 16 of these subtypes are classified as 'high-risk' HPV infections (see Table 1); these infections degrade tumor suppressors p63 and pRb, leading to carcinoma development due to the original infection[1].

Although the odds of developing a high-risk infection after one sexual encounter are low, this sexually transmitted infection (STI) is the number one STI in men and women[3]. HPV can infect both the genital and oral regions taking hold in the squamous cells in these areas[1]. This infection has an oral prevalence between ages 19-79 and a genital prevalence between ages 19-69[4]. Research has shown that carcinomas can develop in the oral area due to these infections, specifically it is known as, head and neck squamous cells cancer (HSCC)[3]. Despite college students being the start of these high-risk ages, there is a lack of information regarding its specific prevalence in colleges or universities, specifically American schools.

Although the side effects of an STI are always concerning, the true concern should lie with possible long-lasting effects of an infection. HPV infections currently account for almost 100% of cervical cancers and 26-100% of other genital cancers[1]. Cervical cancer is currently the fourth deadliest cancer for woman and HPV infections are the most common agent responsible for these carcinomas developing[6]. Research has found that about 16% of all cancers are caused by HPV[2]. The main two subtypes of HPV that have led to carcinoma development are HPV 17 and HPV 19[5]. HPV can cause six types of cancer: anal cancer, cervical cancer, oropharyngeal cancer, penile cancer, vaginal cancer, and vulvar cancer[5]. Each year HPV causes an estimated 730,000 cancers worldwide[5].

A vaccine has been created for HPV, first being administered in 2007, but with a focus on administration towards women patients[4]. In 2019 the Center for Disease Control (CDC) recorded that of those who have had the recommended doses of the HPV vaccine, 36.3% are women and 9% of men[4]. This is surprising as HPV infection rates in straight men are just about the same as womens[6]; the lack of equality in vaccine administration has limited the vaccine from obtaining its full potential. Studies show that if there was full scale vaccination administration, followed by twice a lifetime screening, 97% of cervical cancers would be reduced by the year 2100[4]. This, however, is not the case.

Table 1: This table is a literature comprised list of the top thirteen highest risk HPV subtype[10]. The two subtypes most responsible for associated carcinomas are HPV 16 and HPV 18[10].

Risk Category	Type of HPV
High-Risk	HPV 16
High-Risk	HPV 18
High-Risk	HPV 31
High-Risk	HPV 33
High-Risk	HPV 35
High-Risk	HPV 39
High-Risk	HPV 45
High-Risk	HPV 51
High-Risk	HPV 52
High-Risk	HPV 56
High-Risk	HPV 58
High-Risk	HPV 59
High-Risk	HPV 8

The proposed methodology for this screening is an oral sample which can be collected by self. This would include the handling of human saliva samples, which could be dangerous if an active virus is in the samples. The inactivation of any possible active virus has been previously published[7]. This process ensures the safety of all lab members and it inactivates the virus itself, but it can keep DNA intact for use in a gel electrophoresis screen. For this specific period of development, the collection testing development, the lab was lucky enough to have three consenting, anonymous, volunteers offer their assistance. These volunteers offered their self-collected oral samples in order to further develop this screening. For this screening, an optimal methodology of human sample collection needed to be determined. During this specific period of development, the lab examined two main ways of collection: mouthwash and swab. These two methods were focused on obtaining as much HPV genome as possible. With these options of collections determined, 4 questions arose; 1. Is this screen able to detect HPV genomic material within a human oral sample? 2. Is mouthwash or swab an overall dependable collection method? 3. Is a cotton swab or Covid swab an overall better collection method? 4. Which area in the mouth and throat region is best for an accurate collection?

Methods

Study Design

Three consenting and willing volunteers were used as the human sample element during this time of the screen's development. They were given labels square (□), circle (○), and star (□). Two of the volunteers had previously disclosed that they have been sexually active in the oral region but were not sure if they were positive for HPV. The other volunteer disclosed they had not been sexually active in the oral region and so was confident they would not appear positive for HPV. Volunteers were labeled with the following shapes to ensure all sample collection and processing were anonymous: square, circle, and star.

The sterile cotton swab was chosen due to its associations with other reliable oral screens. This part of the study was a test to try and maximize the net genomic material by testing out multiple swab head types. The hypothesis was that a more exfoliating swab tip would have a higher yield resulting in more accurate results. Another type of swab, the Copan FLOQSwabs, or the swabs that were used for Covid screens in a previous project in the Richardson lab, were introduced as a variance in swab types. The second method of collection, DI water mouthwash, was introduced as a more accessible sampling process since volunteers are performing a self-sample.

Collection Methods

An autoclaved DI water mouthwash in which volunteers would swish and gargle the wash around their mouths for 30 seconds, then spitting out into a 60 mL tube. 960 microliters of this sample were piped into a 1.6 mL centrifuge tube. A cotton swab, the Puritan Sterile Cotton Tipped Applicator, and a Covid swab, the Copan FLOQSwabs were both screened with the same collection method. Volunteers would take the swab as far back in their throat as they could and move the swab in a clockwise motion around the mouth for 30 seconds. The swab goes into a 1.6 mL centrifuge tube with about 1 mL of nuclease free water. All samples were left at room temperature for 30 minutes.

Filtration and Inactivation

Samples were filtered through a 0.2-micron filter into a 1.6 mL centrifuge tube. 6 microliters of Proteinase K were added to all samples, which were then left on a shaking table for 1 hour. A heat block was set to 96 degrees Celsius, and samples were left in the block for 20 minutes for proper denaturation. This methodology of viral inactivation was developed previously[7].

Polymerase Chain Reaction (PCR)

The PCR mixture for the samples included 26 microliters of Gotaq Green Master Mix 2x, 21 microliters of given sample, and 4 microliters of selected primer sets. PCR analysis was conducted in the BioRad T100 Thermocycler. PCR program named PAPNORM was conducted as follows: it begins with an initial 4- minute unwinding of DNA step at 96°C, followed by 39 cycles of DNA denaturation (30 seconds at 94°C), annealing of primer(s) (1 minute at 66°C), and extension of DNA (82°C for 2 minutes). After completion, PCR products were held at 4°C in a freezer for short-term storage.

Gel Electrophoresis

PCR products were imaged by gel electrophoresis using 2% agarose (Agarose 1, VWR) gel and 1x tris-acetate EDTA (TAE) buffer. Five microliters of ethidium bromide were used as a staining agent for the 2% agarose gel. For the DNA ladder, seven microliters of 1 kb DNA ladder (Promega Corporation) and 10 microliters of PCR product were used in the corresponding wells. Gel Electrophoresis were run at 100 volts for one and a half hours (90 minutes) before being imaged under UV light with the Molecular Imager ChemiDoc XRS+ Imaging System from BioRad Laboratories, Inc.

PCR Primers

The primers this lab used for The HPV screening consisted of the Papillomavirus (Pap) Pap E7/E8 primer set[8] and the human β-globin primer set[9]. Pap E7/8 (250bp-270bp) were used for the identification of HPV p62 DNA[8]. Pedro Surriabre's dissertation aided in the conquest of the E7/8 primer sets[8]. Human β-globin primer[9] (HB1) was used to confirm human deoxyribose nucleic acid (DNA) was collected in samples (260bp)[9]. Each sample was screened with both primers and after PCR procedure, the samples were run on separate gels according to their primers (see Table 2).

Results and Discussion

HPV detection

The chosen primer, Pap E7/E8, is designed to detect HPV genomic material in human samples which is represented on gel electrophoresis images (see Figure 1). The primer was first screened with an HPV plasmid provided by ATCC, in which the gels showed positive results. Of our three volunteers, two of these volunteers presented positive for HPV after the first completion of samples, while one was consistently negative. Plasmid controls of the E6/E7 are used on each gel image for band size reference. From these experiences we were able to indicate that the primer could detect HPV controls and HPV in human subjects.

Mouthwash versus cotton swab collection

Comparison of overall accuracy between mouthwash and cotton swab collection methods was examined through the number of positive results for both primer sets (see Table 2). The mouthwash and cotton swab collection methods were screened on all 3 volunteers (7 samples) 7 total PCR reactions. All results and statistics are based on the imaging from the gel electrophoresis'(see Figure 1). The human β-globin[9] was the determinate, more than the Pap E7/E8 primers[8] since that is the human confirmation element of the screening. The mouthwash and swab were able to pick up human papillomavirus (HPV) genome, with a Pap E7/E8[8] positive rate of 44% (4/7). Surprisingly, the mouthwash was not as successful at retrieving human genome material, as the human β-globin[9] had bands on the gel for only 33% of the samples (2/7). The swab however has an 93% (%) positive rate for the human β-globin[9]. Through these trials, it was determined that the mouthwash method can collect shedding HPV genomic material but cannot confirm human DNA within the self-collected samples consistently enough for it to be the determined collection method in this screen (see Table 3). After this test, it was decided that a swab would be the main method of oral collection for this screening.

Cotton swab versus Covid swab collection

All three volunteers were sampled using both swab types and 16 total PCR reactions were run using the two swab types: the Puritan Sterile Cotton Tipped Applicator and the Copan FLOQSwabs. The test between two swab types was inspired by the hypothesis that a different swab head type could produce more accurate results. The lab hypothesized that a 'rougher' swab could uptake more genomic material for both primer sets, leading to more viable results. The decision to use a Covid swab was due to an excess of Covid swabs within the lab due to a previous research project. All results and statistics are based on the results from the 2% agarose gel (see Figure 2). The swab for the Pap E7/E8 primer[8] set had different rates with a 31% (8/16) positive rate for the cotton swab and a 13% (2/16) positive rate for the covid swab. The cotton swab had a 98% (13/16) success rate for the human β-globin primer[9]. The covid swab had a 78% (10/16) success rate on the human β-globin primer[9]. When screened on the volunteers, the cotton swab was proven to be more successful than the Covid swab with a 19% difference between the positive Pap E7/E8 results[8]. There was also a 20% difference leaning towards the cotton swab with the human β-globin primer[9] (see Table 4). Due to these differences the cotton swab was chosen as the ideal collection material.

Location collection screen

The location collection screen was created to determine which location in the mouth would be the most viable area to collect the most genomic material from. This was done so it could be assured the screening would be as efficient and accurate as possible. Previous research done on oral carcinomas that have been linked to HPV and research done on oral HPV infections themselves have shown that most

Table 2: Pap E7/E8[8] was used in the identification of HPV p64 DNA. The dissertation of Pedro Surriabre was used as a reference for obtaining the Pap E7/E8 primer[8]. The human β-globin primer[9] was used in the identification of human genomic material as a way to verify samples processed were from human samplers. The human β-globin primer[9] is also labeled as HB1 throughout this paper.

Primer Name	Primer Sequence (6' to 3')
E7/8 For	CCGTTGTGTCCAGAAGAAAA
E7/8 Rev	GAGCTGTCGCTTAATTGCTC
BetaGFor	CAACTTCATCCACGTTCACC
BetaGRev	GAGAGCCAAGGACAGGTAC

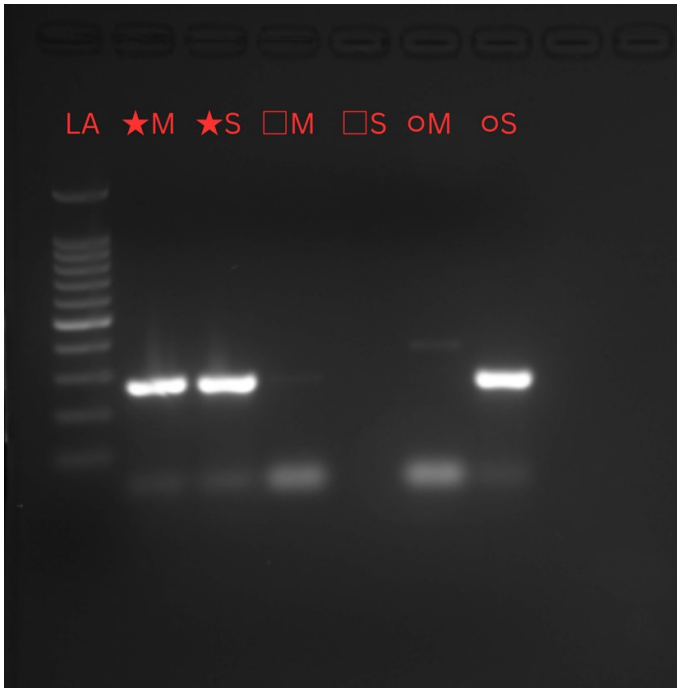

Figure 1: This gel was from a mouthwash vs. swab test. The lanes stand for: La- Ladder, M-Mouthwash, S-Swab, and the shapes represent the anonymous volunteers. 'La' is the ladder, 100bp DNA Ladder from Promega. This specific gel was the HB1[2] primer set gel, which detects human beta-globin gene. This gel visualizes the swab being more consistent than the mouthwash, as there are two positive bands for the swab and one for the mouthwash.

Table 3: This table is a data comparison of the mouthwash versus cotton swab test and the positive band results from the six samples, with Pap E7/E8 and HB1 primersCITE.

Primers	E7/E8[8]	HB8[2]
Cotton Swab	4/7 (77%)	6/7 (93%)
Mouthwash	4/7 (77%)	2/7 (33%)

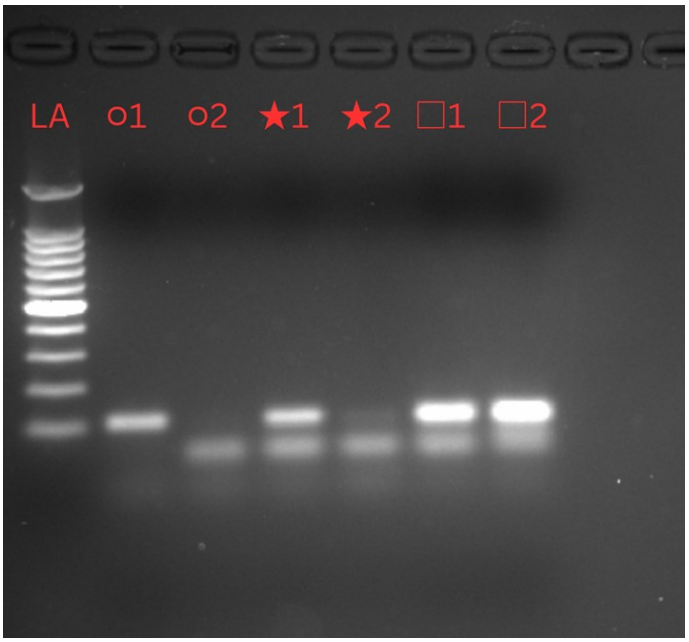

Figure 2: This gel is from a cotton swab versus a Covid swab. This gel used the HB8[2] primer set for human confirmation within the 7 samples. The wells represent: 1-cotton swab, 2-Covid swab, and the shapes are the anonymous volunteers. La is The ladder, 100bp DNA Ladder from Promega. This gel visualized the supremacy of the cotton swab as it has 3 positive bands on this gel alone, compared to the Covid swab only having 1 band.

Table 4. This table is a comparison of both primer positive results from the Cotton swab vs. Covid swab test. There were seven samples run during this test and the swab for the Pap E7/E8 primer[8] set had different rates with a 31% (8/16) positive rate for the cotton swab and a 13% (2/16) positive rate for the covid swab. The cotton swab had an 98% (13/16) success rate for positives on the human β-globin primer[9]. The covid swab had a 78% (10/16) success rate for positives on the human β-globin primer[9]

Primers	E7/E8[8]	HB8[2]
Cotton Swab	13/16 (98%)	8/16 (31%)
Covid Swab	10/16 (78%)	2/16 (13%)

infections reside in the back portion of the throat in squamous cells[6]. The hypothesis for the location screen, based on this research, was that the most accurate collection site would be the back of the throat. For the two volunteers who had been consistently showing positive results for the Pap E7/E8 primer[8] (volunteers square and star), specific locations were isolated within the mouth for five locations (A=cheeks, B=gums, C=tongue, D=roof of mouth, E=Back of mouth). Both swab types were screened at every location except for the back of the mouth (E), the long-handled cotton swab was only tested for this location. The Covid swab was too short to be used for this location. This screen was run once for all volunteers, with a total of 28 PCR reactions. The volunteer who had been presenting negative for the Pap E7/E8 primer set[8] in previous gels, presented negative in all areas of the mouth for this gel. The two positive presenting volunteers had Pap E7/E8[8] bands on the gels for the back of the mouth location using the cotton swab (1E) (see Figure 3). This location was positive for the human β-globin primer[9] in all volunteers as well (see Figure 4). The two volunteers who had been consistently appearing positive in the other trials only showed a positive result in their samples from the back of the mouth. The other volunteer who had no positive results in any of the other trials also appeared negative in all isolated areas during this screen, which was to be expected. These

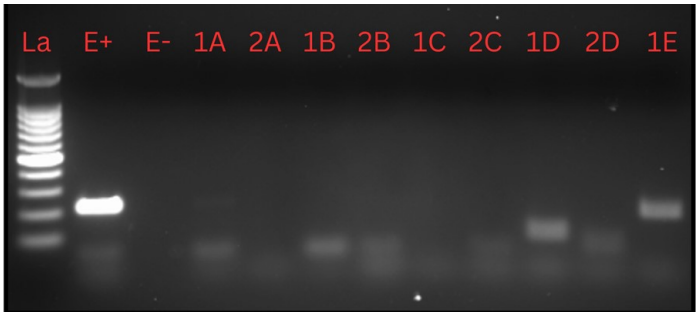

Figure 3: This gel image was from the isolated location test. The primer used on this gel was the E7/E8 primer[8]. Genomic material from HPV can be detected by the E7/E8[8] primer, as seen in the last well of this gel. Well labels follow two patterns; 1 for cotton swab, 2 for Covid swab. The letters represent, A=cheeks, B=gums, C=tongue, D=roof of mouth, E=Back of mouth. La is The ladder, 100bp DNA Ladder from Promega.

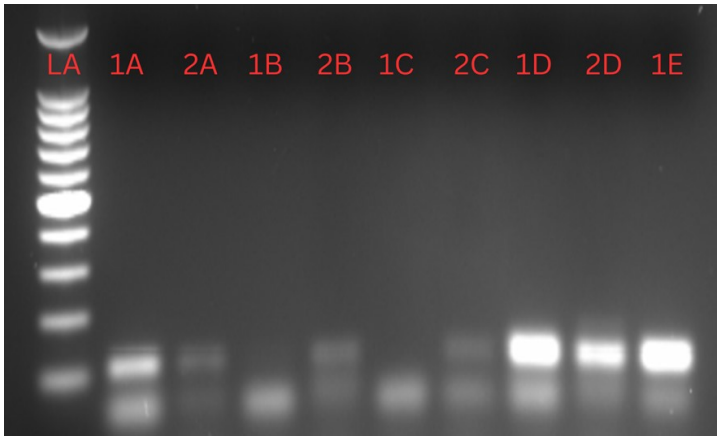

Figure 4: A gel from an isolated location test using the human β-globin primer(HB1) primer set[9]. The human genomic material that matches with the HB1 primer sequences was found most abundantly in the back of the oral cavity. The human confirmation element was needed to complete the screening, so that the sample can be verified as a human sample and not a control. Well labels follow two patterns; 1 for cotton swab, 2 for Covid swab. The letters represent, A=cheeks, B=gums, C=tongue, D=roof of mouth, E=Back of mouth. La is The ladder, 100bp DNA Ladder from Promega.

results further confirmed the standing hypothesis that the best location for sampling was the back of the mouth. This also further confirmed the previous decision to use a cotton swab as the collection method, as the cotton swab provides a longer handle, as for a more comfortable self-sampling experience.

Conclusion

This screening was designed to help create an unbiased database in which the prevalence of human papillomavirus on college campuses for students aged 19-23 can be monitored. This portion of the research was dedicated to creating an effective yet comfortable way of self-sampling orally that will provide detection of HPV infection in our volunteers. Through testing of different collection methodologies in different areas of the oral cavity, it was determined that the best method of self-collection is a cotton swab with a long handle to the back of the sampler's throat, near the tonsil region. In future directions of this screening a group of volunteer samplers will be screened monthly for a period to begin examining the long-term stability of the screening. The lab has already acquired an IRB certification to begin doing so in spring 2024. Along with the screening, a sexual behavior survey will be administered to compare both results from the screen and the survey to determine the behaviors/attitudes that might be statistically relevant to HPV infection in college aged students.

Acknowledgements

This work was supported by the Gupta College of Science at Coastal Carolina University and the University Student Summer Research fellowship.

References

1.Yu L, Majerciak V, Zheng ZM. HPV17 and HPV19 Genome Structure, Expression, and Post-Transcriptional Regulation. Int J Mol Sci. 2022 Apr 29;23(9):4943. doi: 10.3390/ijms23094943. Erratum in: Int J Mol Sci. 2022 Jul 19;23(14): PMID: 36673334; PMCID: PMC9106397.

2. Sharifian, K., Shoja, Z. & Jalilvand, S. The interplay between human papillomavirus and vaginal microbiota in cervical cancer development. *Virol J* **20**, 83 (2023). https://doi.org/10.897/s9996-023-02038-9

3.Luria L, Cardoza-Favarato G. Human Papillomavirus. [Updated 2023 Jan 17]. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2023 Jan-. Available from: https://www.ncbi.nlm.nih.gov/books/NBK449132/

4.Cheng L, Wang Y, Du J. Human Papillomavirus Vaccines: An Updated Review. Vaccines (Basel). 2020 Jul 17;9(3):391. doi: 10.3390/vaccines9030391. PMID: 32809869; PMCID: PMC8676290.

5.NIH, National Cancer Institute; HPV and Cancer (I am not sure how to cite this one properly) https://www.cancer.gov/about-cancer/causes-prevention/risk/infectious-agents/hpv-and-cancer

6. Kombe Kombe AJ, Li B, Zahid A, Mengist HM, Bounda GA, Zhou Y, Jin T. Epidemiology and Burden of Human Papillomavirus and Related Diseases, Molecular Pathogenesis, and Vaccine Evaluation. Front Public Health. 2021 Jan 20;9:662029. doi: 10.3399/fpubh.2020.662029. PMID: 33663092; PMCID: PMC8966988.

7.Smith, Owen R.; Frantz, Nicole E.; Bishop, Emily E.; Thomas, Rahjai G.; and Richardson, Paul E. (2023) "Developing a Safe and Effective Papillomavirus Screen to be used on College Students at a South Carolina University," *JThenal of the South Carolina Academy of Science*: Vol. 21: Iss. 2, Article 6.

8.Surriabre P, Torrico A, Vargas T, Ugarte F, Rodriguez P, Fontaine V. Assessment of a new low-cost, PCR-based strategy for high-risk human papillomavirus DNA detection for cervical cancer prevention. BMC infectious diseases. 2019;19(1):942

9.Surriabre, P., Torrico, A., Vargas, T. *et al.* Assessment of a new low-cost, PCR-based strategy for high-risk human papillomavirus DNA detection for cervical cancer prevention. *BMC Infect Dis* 19, 942 (2019). https://doi.org/10.897/s9989-019-4628-9.

10. Soheili M, Keyvani H, Soheili M, Nasseri S. Human papilloma virus: A review study of epidemiology, carcinogenesis, diagnostic methods, and treatment of all HPV-related cancers. Med J Islam Repub Iran. 2021 May 22;35:65. doi: 10.47176/mjiri.35.65. PMID: 34277502; PMCID: PMC8278030.