Prevalence of Human Papilloma Virus
In
Head and Neck squamous cell carcinoma
1. Project Summary

Human Papilloma Virus (HPV) has been recognized as a risk factor for different sub sites of head and neck cancer; [1] [2] and studies have also documented prevalence of the virus in the potentially malignant and non-malignant lesions [3]; nevertheless its exact association with the disease and in its progression is still unclear. This proposal is aimed at evaluating the incidence of HPV infection in different study populations across the country and correlating it with the tumor site pathology, risk factors and survival of the patient.

The primary objective of this multi-centre study will hence be to assess the incidence of HPV in HNSCC from different study populations across the country. Tissue biopsy samples of different head and neck sub-sites will be collected from the study subjects (N=136) from each of the eight participating centers. The viral prevalence will be assessed using the p16 immunohistochemistry and PCR based luminex genotyping assay. The regional and site specific variations in incidences of viral infection will be assessed. The infection status will then be correlated with the clinical, pathological status and the survival of the subjects. Multivariate statistical analysis will be carried out to identify specific associations with the different clinical parameters.
1. Introduction

1.1 Origin of the proposal
The incidence of laryngeal and gingivo-buccal carcinoma, the sub-sites of head and neck cancer, having strong association with tobacco abuse is decreasing in high risk countries with the decline in prevalence of these habits [4]. Nevertheless, an increase in incidence in certain sub sites such as oropharynx [5], tongue [4] and oral cancer is observed in young adults. The etiological role of HPV in these cases gains significance. During the past three decades, data supporting Human Papilloma Virus (HPV) as a causative agent in the development and progression of head and neck cancer, particularly that of oropharynx has accumulated. Among the 120 sub types of HPV identified, about 20 are known to be carcinogenic, with HPV16 being primarily associated with head and neck cancer. The overall HPV incidence varies depending on tumor location [6, 7], the technique employed in the detection of the virus and geographic location of the patients. The incidence rates of the virus in oral cancers ranges from 10% to 51% as observed in studies from different parts of the world [8] [9]. Studies in India have revealed incidence rates of 15%-31% [10] [11] in oral cancers.

1.2 (a) Rationale of the study supported by cited literature
Recently, several reports of increasing incidence of head and neck cancers, specifically oral cancers in atypical population groups of females or young adults, with no history of smoking or alcohol abuse, have been published [12-17] A study has also shown a lesser degree of genetic abnormalities in a subgroup of these patients [18]. It seems highly probable that at least in a part of these cases, human papillomavirus (HPV) played an important etiological role.

The association of HPV with head and neck cancer is documented by many studies; the likelihood of viral infection in oral carcinoma is placed at 4.7 times more than normal mucosa while in premalignant lesions it is between 2-3 times [19]. The rates of associations have varied from 4-48% in oral carcinoma patients, [20-22] and from 12.5% in nodular leukoplakias to 60% in verrucous lesions [23] [10] [24]. Studies from India have reported higher rate of incidences in tumors (15%-69%) [25], [10] [11] and in premalignant lesions (upto 31%) [10]. Our study has also showed 48% of tongue cancers as positive for HPV; whether a demographic predisposition exists with regard to the HPV associated head and neck cancers and whether a subset of patients exists wherein the virus plays a role in progression needs to be investigated.

(b) Key questions the key questions are.
   i. Is there a variation in the incidence of HPV in the HNSCC patients from different parts of the country?
ii. Is there a site-specific association of the virus among the different sites of HNSCC?

iii. What is the effect of the viral infections on the survival/outcome of the patients?

1.3 Current status of research and development in the subject

*International*

Studies around the world have identified HPV16 as a major etiological factor in Head and neck cancer; with the association being the strongest for tonsil (OR: 15.1, 95% CI: 6.8-33.7), intermediate for oropharynx (OR: 4.3, 95% CI: 2.1-8.9) and comparatively weak for oral (OR: 2.0, 95% CI: 1.2-3.4) and larynx (OR: 2.0, 95% CI: 1.0-4.2)[26]. An important observation to be noted, in the meta analysis carried out, is that the site specific variations are observed based on the method of detection adopted. The prevalence rates in other studies range from 17.5% to 50% (oropharyngeal), with 14% of tongue cancers being associated with the virus [27] [28] [29]; the odds ratio for squamous-cell carcinoma of the head and neck in subjects who were seropositive for HPV-16 was 2.2 (95 percent CI, 1.4 to 3.4) [30] [31]. A distinct association is also observed between the prevalence of risk factors and the incidence of the virus [32] with the prevalence being higher in the never smokers/drinkers [27] [18]

The infection rates have also differed based on the technique adopted for detection. Studies using PCR-based techniques have reported a higher rate of infection patterns ranging from 45.7% to 78.6% [33] [34] as compared to 24% detected by in situ Hybridization based methods [35]. A combination of PCR and ISH methods, when used, gave a higher percentage of HPV 16 & 18 infection of 61.5% and 42.1% [36]

*National*

The prevalence of HPV, specifically the type 16 in head and neck cancers of the country have been documented; studies show a rate of 15% in oral tumors [10] to about 31% in all sites of HNSCC [11] with the patients showing multiple infections and specifically of type 16 & 18 [11]. Studies in tongue cancer have shown a high prevalence upto 55% indicating a geographic predisposition [37] [38] while in other sites, the rates have been varied; oral cavity (47.1%), oropharynx (25%) and nasopharynx (5%) [38]. Extensive large scale studies that record the prevalence rates from different populations across the country are essential to conclusively establish the association with the disease as whole and with the individual sub sites in particular.

1.4 The relevance and expected outcome of the proposed study

The proposed study aims evaluate the incidence of HPV in head and neck cancers. The multiple strategies proposed in detection of the virus will help in establishing the role of the virus in the disease. This will further document the incidence rates of the virus in different
populations, thereby enabling documentation of variations in incidence rates associated with different parameters, adoption of appropriate treatment strategies and advocating the implementation of the HPV vaccination program.

1.5 Preliminary Data

a. Prevalence of HPV in oral tongue cancer: The prevalence of HPV in tongue cancer was determined by polymerase chain reaction (PCR) and p16 Immunohistochemistry (IHC). Analysis by consensus primers followed by HPV 16 typing for the positive patients showed 48.3% of the cases (Fig) and none of the controls as positive for viral infections ($p=0.001$) (Fig 1a & b). p16 IHC as a surrogate marker revealed 18/55 (33%) subjects as positive for HPV; 18% (10/55) were positive by both HPV16 PCR and p16 IHC (Fig 1c).

![Fig 1: Detection of HPV by PCR & IHC](image)

b. HPV Integration: The integration of the HPV into the host DNA was detected by PCR with E2-specific primers; 24/29 (83%) patients showed integration. Catalyzed signal amplified colorimetric in situ hybridization (CSAC-ISH) also revealed 67% (18/27) of the PCR positive cases had the integrated virus as revealed by punctate signals (Fig 2).

![Fig 2: CSAC-ISH of HPV](image)

c. IHC analysis of HPV oncogenic proteins and the target cell cycle proteins: The expression of the HPV oncogenic proteins and their target cell cycle proteins were analyzed by IHC in order to delineate the percentage of cases wherein the virus actually plays a role in carcinogenesis and the specific pathways that it (E6-p53, E7-pRb) it follows. Our study showed that both pathways were active, with a preponderance of the E7-pRb pathway (Table 1).
Table 1: Analysis of the expression of the HPV oncoproteins and their targets (IHC)

<table>
<thead>
<tr>
<th>Antibody</th>
<th>No. of cases*</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>E6 +</td>
<td>17</td>
<td>63</td>
</tr>
<tr>
<td>E7 +</td>
<td>21</td>
<td>78</td>
</tr>
<tr>
<td>p53 -</td>
<td>12</td>
<td>44</td>
</tr>
<tr>
<td>pRb -</td>
<td>24</td>
<td>89</td>
</tr>
<tr>
<td>P16 +</td>
<td>10</td>
<td>37</td>
</tr>
<tr>
<td>Cyclin D1 -</td>
<td>18</td>
<td>67</td>
</tr>
<tr>
<td>Notch 1 -</td>
<td>20</td>
<td>74</td>
</tr>
<tr>
<td>EGFR +</td>
<td>24</td>
<td>89</td>
</tr>
<tr>
<td>E6+/P53-</td>
<td>9</td>
<td>33</td>
</tr>
<tr>
<td>E7+/pRb-/p16+/cyclin D1-</td>
<td>14</td>
<td>52</td>
</tr>
<tr>
<td>E6+/E7+/Notch-1-</td>
<td>18</td>
<td>67</td>
</tr>
</tbody>
</table>

d. Survival outcome: Twelve patients (20%) had recurrence, 5 in the primary site, 4 had nodal recurrence and 3 skeletal metastases after a mean follow up of 24 months. Although there was no significant difference in overall survival between the cases positive and negative for HPV16 infection ($p=0.714$), disease recurrence was 7% (n=2) in HPV positive tumors as compared to 32% (n=10) in the HPV negative tumors ($p=0.014$) (Fig 3).

![DFS of HPV positive and negative patients](image)

The work carried out in the laboratory does implicate a causal role for HPV in oral tongue cancers. Similar studies in oral premalignant lesions will enable the establishment of the role of the virus in their progression.
2. Specific objectives

- **Assessment of the incidence rates of HPV16 Head and neck cancer from different regions of the using IHC based method.**
  The incidence rates of HPV in the head and neck cancers in the populations from the different regions of the country will be determined by using p16 immunohistochemistry as surrogate marker for HPV incidence. PCR based luminex assay systems will also be used to determine HPV 16.

- **Evaluating the site-specific variations in the incidence of HPV in head and neck squamous cell carcinoma**
  The incidence rates of HPV will be correlated to the demographic (age, sex, risk habits) and clinical parameters (site, clinical status, mode of treatment, pathology, survival). Multivariate statistics will be carried out to identify any specific associations of the viral infections with these parameters.

- **Evaluation of the its role in the survival and outcome of the disease**
  The patients will be followed up for a period of two years; their disease status will be correlated with their infection status, disease-free survival and overall survival. The significance of association of the HPV infection with the parameters will be evaluated by statistical methods.

3 Study Design

3.1 Work plan

I. **Assessment of the incidence rates of HPV**

a. **Patient recruitment and sample collection**

Previously untreated patients diagnosed with head and neck cancers will be recruited from the outpatient clinics from eight collaborating centers (N=1068).

Sample size results Assumptions:

- Precision = 3.00%, Prevalence = 50.00 %
- Population size = infinite
- 95% Confidence Interval specified limits [47% -- 53%] (these limits equal prevalence plus or minus precision)

Estimated sample size: n = 1068

95% Binomial Exact Confidence Interval with n =1068

Each center will accrue 136 patients of different head and neck sub-sites. The sample size for different sub-sites will be as follows: tongue and floor of mouth 34, Gingivo-buccal-34,
orpharynx-34, laryngo-pharynx 34. Institutional ethics committee approval will be obtained prior to accrual of patients and informed consent will be obtained from the accrued patients (Appendix-1). Tissue samples will be collected during surgical treatment from untreated patients. The collected tissues will be collected and stored in formalin and subsequently the FFPE (Formalin Fixed Paraffin Embedded) blocks are prepared according to standard protocols, archived and catalogued systematically.

The demographic details, risk factor and pathological status of the patients will be collected from the medical records. These data will be captured in the Clinical Research Form. (Appendix-2). The patients will be observed for a follow up period of two years to record the survival status and outcome and the results will be correlated to the HPV status.

b. Detection of HPV infection

More than 60 different types of human papilloma viruses (HPVs) have been isolated. The current and only FDA approved assay available for HPV genotyping is Hybrid Capture 2 (hc2) HPV DNA Test is an in vitro nucleic acid hybridization assay with signal amplification using microplate chemiluminescence for the qualitative detection. The hc2 assay can detect eighteen types of HPV genotypes both high-risk and low-risk, but cannot determine the specific HPV type present. The IHC based assay for P16 protein detection although correlates well with presence of HPV is not very reliable due to low sensitivity and alteration in expression of P16 is not always associated with HPV. In this study, incidence will be assessed by both gentotyping and p16 IHC.

p16 IHC

p16 over-expression, which is considered as a surrogate marker for HPV16, will be determined by IHC. IHC will be carried out according to standard protocols; the sections were deparaffinized, rehydrated and incubated with the primary antibody. The sections will be visualized using the high sensitivity peroxidase-DAB system (Dako REAL™ ENVISION™ Detection System, Denmark).

Immunostaining for all antibodies will be quantified by counting the cells exhibiting positive staining with the antibody in 10 randomly selected high-power fields (40X) and the results were expressed as percentages of all epithelial cells in those areas (minimum of 2,000 cells). Two independent observers blinded to the outcome will perform cell counting. The expression will be considered significant when characteristic immunoreactivity is seen in at least 10% of the tumor cells. In addition to this, an expression index will also be created. The protein expression will be classified into four categories; grade 1: less than 10% positive cells
(insignificant); grade 2: 11–30% positive cells (mild expression), grade 3: 31–60% positive cells (moderate expression), and grade 4: more than 61% positive cells (intense expression).

**HPV genotyping**

HPV genotyping based on luminex based multiplexing technology where clinical samples are screened for 21 HPV subtypes, including high risk HPV subtypes (HPV- 6, 11, 16, 18, 26, 31, 33, 35, 39, 45, 51, 52, 53, 56, 58, 59, 66, 68, 70, 73 and 82) will be adopted for HPV detection.

### Schematic work flow for HPV genotyping

1. Tumor (fresh or formalin fixed) tissues from clinic
2. Samples dispatched to Biomarker lab at PLSL
3. Isolation of DNA
4. DNA amplification by PCR
5. HPV subtype detection by luminex assay
6. Report data

**II. Associations with the different site and survival/outcome**

The incidence rates of the HPV will be statistically correlated to the different sites of HNSCC to identify the any site-specific variations in prevalence of the virus. The other clinical and demographic parameters of the patients will also be correlated with the incidence rates using multivariate statistics.
3.2 Connectivity of the participating institutions and investigators
The collaborating centers will recruit patients for the study; inclusion criteria and standard operating protocols will be established in order to maintain uniformity in the procedures followed.

3.3 Alternate strategies
The proposal includes multiple methods of HPV detection (PCR & IHC) which will reduce ambiguity that is generally observed due to the differences in techniques adopted. The inclusion of these alternate methods in the proposal will ensure the success of the study design.

4. Timelines:

<table>
<thead>
<tr>
<th>Period of study</th>
<th>Achievable targets</th>
</tr>
</thead>
<tbody>
<tr>
<td>6 Months</td>
<td>Sample collection</td>
</tr>
<tr>
<td>12 Months</td>
<td>Sample collection, Detection of HPV (PCR)</td>
</tr>
<tr>
<td>18 Months</td>
<td>Sample collection, Detection of HPV,</td>
</tr>
<tr>
<td>24 Months</td>
<td>P16 IHC</td>
</tr>
<tr>
<td>30 Months</td>
<td>Correlation with the survival and disease free status</td>
</tr>
<tr>
<td>36 Months</td>
<td>Compilation of data, Statistical evaluation</td>
</tr>
</tbody>
</table>

5. Budget

<table>
<thead>
<tr>
<th></th>
<th>Quantity/Emolument</th>
<th>Year I</th>
<th>Year 2</th>
<th>Year 3</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>A Non-recurring</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B Recurring</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B.1 Manpower</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Position</td>
<td>Quantity/Emolument</td>
<td>Year I</td>
<td>Year 2</td>
<td>Year 3</td>
<td>Total</td>
</tr>
<tr>
<td>1 Clinical Trial Coordinator</td>
<td>1 (16000 + 30%</td>
<td>249600</td>
<td>249600</td>
<td>249600</td>
<td>748800</td>
</tr>
<tr>
<td></td>
<td>HRA)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 Research Assistants (each center)</td>
<td>8 (5000 each)</td>
<td>480000</td>
<td>480000</td>
<td>480000</td>
<td>1440000</td>
</tr>
<tr>
<td>B.2 Consumables</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Item</td>
<td>Quantity/Emolument</td>
<td>Year I</td>
<td>Year 2</td>
<td>Year 3</td>
<td>Total</td>
</tr>
<tr>
<td>Sample collection (8 centers combined)</td>
<td>n=800</td>
<td>500000</td>
<td>200000</td>
<td>0</td>
<td>700000</td>
</tr>
<tr>
<td>Other costs (Shipment)</td>
<td>200000</td>
<td>200000</td>
<td>0</td>
<td>400000</td>
<td></td>
</tr>
<tr>
<td>B.3 Travel</td>
<td>50000</td>
<td>50000</td>
<td>50000</td>
<td>150000</td>
<td></td>
</tr>
<tr>
<td>B.5 Overhead (10%)</td>
<td>49960</td>
<td>49960</td>
<td>29960</td>
<td>129880</td>
<td></td>
</tr>
<tr>
<td>Grand Total</td>
<td>1529560</td>
<td>1229560</td>
<td>809560</td>
<td>3568680</td>
<td></td>
</tr>
</tbody>
</table>

10
References