Chapter 1

INTRODUCTION

1.1 CERVICAL CANCER INCIDENCE AND PREVALENCE

Uterine cervical cancer (CaCx) is the third most common cause of cancer-related deaths among women, globally (Globocan 2008) with an estimated incidence of 530,000 and mortality of 275,000 in 2008. More than 85% of the global burden occurs in developing countries with an age-standardized incidence rate of more than 24.6 per 100,000 in south-Central Asia. The incidence rates of CaCx throughout the world are given in Figure 1.1.

Figure 1.1: Estimated Cervical Cancer incidence worldwide (GLOBOCAN 2008, IARC; http://www-dep.iarc.fr/)
CaCx is the most fatal cancer among Indian women. India accounts for 1/4th of the global burden with an incidence of 134,420 and mortality of 72,825 in 2008 (Figure 1.2).

![Incidence and Mortality](http://www-dep.iarc.fr/)

CaCx is a multifactorial gynaecological disease. It is further subdivided into squamous cell carcinoma (SCC), adenocarcinoma (ADC) and adenosquamous carcinoma depending on the originating tissue. The main etiological factor is persistent infection by sexually transmitted oncogenic human papillomavirus (HPV), mainly type 16 and/or 18, in the squamous or glandular epithelium of the cervix. Studies conducted in India have identified HPV 16 (high-risk type), as the major contributor to development of CaCx, compared to other types. About 82-98% of CaCx cases harbor oncogenic HPVs, 50% of which are HPV 16 positive (Das et al. 2000).
1.2 CANCER OF THE CERVIX

1.2.1 Uterine Cervix: site of the disease

The cervix is the lower fibro muscular portion of the uterus measuring 3-4 cm in length and 2.5 cm in diameter. The surface of the cervix is lined by two types of epithelia - squamous and columnar. Ectocervix is formed by the squamous epithelium, consisting of multiple layers (stratum basale, stratum spinosum, stratum granulosum and stratum corneum) of cells and endocervix formed by a columnar epithelium consisting of a single layer of cells. The boundary between the squamous and columnar epithelia is called the transformation zone (Figure 1.3). It is formed of metaplastic tissue. The reserve cells of the transformation zone that are eventually converted to the basal cells of a stratified squamous epithelium get infected by high risk HPV through some wound on the epithelial surface. Incidence of HPV-associated cancer from other anogenital sites (vulva, vagina, penis) is very low compared to cervix and anus (in homosexual men), both of which has transformation zones (Bosch and de Sanjosé, 2003). The virus can possibly access the cells in the transformation zone more readily than the basal cells that are protected by a permanent stratified epithelial layer.

Persistent infection with oncogenic HPV of these cells results into squamous cell carcinoma of cervix (SCC) after a latent period of about 10-15 years. Adenocarcinomas arise from columnar epithelium of endocervix.
The stratified squamous epithelium covering the cervix provides protection from toxic substances and infection. Under normal circumstances, the superficial layers are continually dying and get sloughed off. The integrity of the lining is maintained by the constant, orderly formation of new cells in the basal layer. However, in the presence of persistent HPV infection and other cofactors, the metaplastic squamous cells of the transformation zone adopt cytological atypia (dysplastic change). The dysplastic cells later multiply in a disorderly manner, to produce squamous cell carcinoma.

HPVs may enter the epithelial basal layer by initial binding to heparan sulphate proteoglycans (Patterson et al., 2005). Role of secondary receptors for efficient infection may be played by the α6 integrin (Bossis et al., 2005). HPV16 is taken into the basal cells by clathrincoated endocytosis (Day et al., 2003).
1.2.2 Progression To Cervical Cancer: a multistage process

Invasive SCCs are preceded by a long phase of preinvasive disease (Figure 1.4), collectively referred to as cervical intraepithelial neoplasia (CIN). CIN may be categorized into grades 1, 2 and 3 depending upon the proportion of the thickness of the epithelium showing mature and differentiated cells. More severe grades of CIN (2 and 3) reveal a greater proportion of the thickness of the epithelium composed of undifferentiated cells. Persistent infection with one or more of the oncogenic subtypes of human papillomaviruses (HPV) is a necessary cause for cervical neoplasia. A small fraction of HPV infections remain persistently within the host system. Most cervical abnormalities caused by HPV infection are unlikely to progress to high-grade CIN or CaCx (Schiffman and Kjaer, 2003). Most low-grade CIN regress within relatively short periods or do not progress to high-grade lesions. High-grade CIN carries a much higher probability of progressing to invasive cancer. The precursor lesion arising from the columnar epithelium is referred to as adenocarcinoma in situ (AIS). AIS may be associated with CIN in one-to-two-thirds of cases.

![Figure 1.4: Epidemiological model of cervical carcinogenesis](image)

**1.2.3 Detection and classification of cervical cancer**

The primary method for detection of cervical abnormalities is still the Papanicolaou-stained (Pap) smear. This method was named after the pathologist, George Papanicolaou, who introduced the method in 1949 before the cause of CaCx was known (Papanicolaou 1949). The Pap smear is a screening tool that looks for changes in cells of the transformation zone of the cervix. The Pap smear reporting
classification has evolved with refinements over time. The current reporting system is the Bethesda System, which was introduced in 1988, amended in 1991 to replace the CIN System, and updated again in 1999 (Solomon et al., 2002). The CIN System is based on tissue architecture and was introduced in 1973 to promote the concept of a disease continuum from precursor lesions to invasive cancer (Richart, 1973). The Bethesda System was developed to reflect an advanced understanding of cervical neoplasia and to introduce uniform descriptive, diagnostic histologic terminology. The Bethesda System 2001 classifies squamous cell abnormalities into four categories: (i) ASC (atypical squamous cells), (ii) LSIL (low-grade squamous intraepithelial lesions), (iii) HSIL (high-grade squamous intraepithelial lesions), and (iv) SCC (squamous cell carcinoma). The details of the Bethesda system and CIN system are given in Table 1.1.
Table 1.1: The Bethesda Classification System for cervical squamous cell dysplasia (Burd, 2003)

<table>
<thead>
<tr>
<th>Bethesda System 1999</th>
<th>Bethesda System 1991</th>
<th>CIN System</th>
<th>Interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Negative for intraepithelial lesions or malignancy.</td>
<td>Within normal limits</td>
<td>Normal</td>
<td>No abnormal cells</td>
</tr>
<tr>
<td>ASC-US (atypical squamous cells of undetermined significance)</td>
<td>ASCUS (atypical squamous cells of undetermined significance)</td>
<td></td>
<td>Squamous cells with abnormalities greater than those attributed to reactive changes but that do not meet the criteria for a squamous intraepithelial lesion</td>
</tr>
<tr>
<td>LSIL (low-grade squamous intra-epithelial lesions)</td>
<td>LSIL (low-grade squamous intra-epithelial lesions)</td>
<td>CIN 1</td>
<td>Mildly abnormal cells; changes are almost always due to HPV</td>
</tr>
<tr>
<td>HSIL (high-grade squamous intra-epithelial lesions) with features suspicious for invasion (if invasion is suspected)</td>
<td>HSIL (high-grade squamous intra-epithelial lesions)</td>
<td>CIN 2/3</td>
<td>Moderately to severely abnormal squamous cells</td>
</tr>
<tr>
<td>Carcinoma</td>
<td>Carcinoma</td>
<td>Invasive squamous cell carcinoma Invasive glandular cell carcinoma (adenocarcinoma)</td>
<td>The possibility of cancer is high enough to warrant immediate evaluation but does not mean that the patient definitely has cancer</td>
</tr>
</tbody>
</table>
1.2.4 Invasive cervical cancer

Invasive CaCx is defined by the invasion of abnormal cells into the thick fibrous connective tissue underlying the basement membrane. It starts with a microinvasive stage, which is not visible with the naked eye on speculum examination and has to be diagnosed histologically, using a tissue sample from a cone biopsy or hysterectomy. It then evolves into larger lesions, which may extend to the vagina, pelvic walls, bladder, rectum and distant organs. If left untreated, CaCx progresses in a predictable manner and will almost always lead to death. The International Federation of Gynecology and Obstetrics (FIGO) system is often used to describe the extent of cancer invasion and to select treatment options.

There are four, usually sequential, routes through which invasive cancer progresses. The disease is generally confined to the pelvis for a long period, where it is accessible to treatment.

Within the cervix: Spread from a tiny focus of microinvasive cancer, eventually involving the entire cervix, which can enlarge to 8 cm or more in diameter. The cancer can be ulcerating, exophytic (growing outwards) or infiltrating (invading inwards).

To adjacent structures: Direct spread in all directions is possible: downwards to the vagina, upwards into the uterus, sideways into the parametrium (the tissues supporting the uterus in the pelvis) and the ureters, backwards to the rectum, and forwards to the bladder.

Lymphatic: Spread to pelvic lymph nodes occur in 15% of cases when the cancer is still confined to the cervix, and increases as the cancer spreads. Lymph node metastases are at first confined to the pelvis and are later found in the chain of nodes along the aorta, eventually reaching the supraclavicular fossa (the space above the collar bone). If the cancer has advanced into the lower third of the vagina, the groin nodes may become involved and will be palpably enlarged.

Distant metastases: CaCx cells may spread through the blood stream and lymphatic system to develop distant metastases in the liver, bone, lung and brain.
1.3 HUMAN PAPILLOMAVIRUS (HPV)

1.3.1 Brief history behind association of sexually transmitted HPV with CaCx

Originally, Domenico Rigoni-Stern (1842), an Italian physician and an instructor at the University of Padua concluded in the mid-19th century that more uterine cancers were found in married than in unmarried women and almost absent in certain orders of nuns. Findings in Italian nuns by Rigoni-Stern's group were confirmed by the studies of Gagnon (1950), Towne (1955), and Taylor et al. (1959).

Drs. Rous and Beard (zur Hausen, 1999) were the first to report papillomavirus as a carcinogen in 1934. In the mid 1970s, sexually transmitted HPV-infection was singled out to be the major risk factor for CaCx (zur Hausen, 1999) and then, HPV-16, -18, -31, and -33 were isolated from CaCx and its precursors (Boshart et al., 1984; Dürst et al., 1983; Lorincz et al., 1986). Dr. Harald zur Hausen (German Cancer Research Centre, Heidelberg, Germany) was awarded 2008 Nobel Prize in Physiology and Medicine for his discovery of HPVs causing CaCx.

1.3.2 Classification

Papillomavirus is an epitheliotropic DNA virus. They are so called because certain types may cause warts, or papillomas. Recently, the International Committee on the Taxonomy of Viruses (ICTV) has accepted papillomaviruses as a distinct taxonomic family, the Papillomaviridae, being unrelated to the polyomaviruses and SV40 (Zheng and Baker, 2006). Papillomaviruses infect various animals ranging from birds to mammals but has exquisite species-specificity and tissue-tropism (Kanodia et al., 2007; Zheng and Baker, 2006). Papillomavirus types, subtypes, and variants differ on the basis of their L1 gene sequences varying from one another by at least 10%, 2-10%, and maximally 2%, respectively (de Villiers et al., 2004). Amongst these, the bovine papillomavirus type 1 (BPV-1) and human papillomavirus type 1a (HPV-1a) genomes were the first to be completely sequenced (Chen et al., 1982). Evolutionarily, HPVs fall under different genera (Figure 1.5), like alpha, beta, gamma, etc (de Villiers et al.,
Almost 90% of the currently characterized HPVs belong to either alpha or beta genus (Doorbar, 2006).

More than 100 HPV genotypes have been described until now (Bernard, 2005) with each having a specific site of infection (de Villiers, 2001). Genus alpha is the largest group and contains HPVs that infect genital/mucosal epithelium. About 30 are related to anogenital infection, of which a subset is associated with CaCx (Muñoz et al., 2003). These are the high risk (HR-HPV) types, like, 16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, 68, 73 and 82 which are related with progression of the disease to high grade intraepithelial lesions (CIN 2, 3) and cervical carcinoma (de Villiers et al., 2004). The high-risk HPVs associated with human cancers mainly belong to Alpha 9 and Alpha 7 groups, of which HPV16 followed by HPV18 are the most prevalent (Lorincz et al.,

Figure 1.5: Classification of papillomaviruses (E. de Villiers, Virology, 2004)
1992; Smith et al., 2007). The low risk types include 6, 11, 40, 42, 43, 44, 54, 61, 70, 72, 81 which account for benign anogenital warts and low grade lesions (de Villiers et al., 2004). Table 1.2 shows the association of different HPV types with different diseases in the order of relative frequencies (Burd, 2003).

**Table 1.2: HPV type and disease association**

(From: Burd, Clinical Microbiology Reviews, 2003)

<table>
<thead>
<tr>
<th>Disease</th>
<th>HPV type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plantar warts</td>
<td>1, 2, 4, 63</td>
</tr>
<tr>
<td>Common warts</td>
<td>2, 1, 7, 4, 26, 27, 29, 41, 57, 65, 77, 1, 3, 4, 10, 28</td>
</tr>
<tr>
<td>Flat warts</td>
<td>3, 10, 26, 27, 28, 38, 41, 49, 75, 76</td>
</tr>
<tr>
<td>Other cutaneous lesions (e.g., epidermoid cysts, laryngeal carcinoma)</td>
<td>6, 11, 16, 30, 33, 36, 37, 38, 41, 48, 60, 72, 73</td>
</tr>
<tr>
<td>Epidermodysplasia verruciformis</td>
<td>2, 3, 10, 5, 8, 9, 12, 14, 15, 17, 19, 20, 21, 22, 23, 24, 25, 36, 37, 38, 47, 50</td>
</tr>
<tr>
<td>Recurrent respiratory papillomatosis</td>
<td>6, 11</td>
</tr>
<tr>
<td>Focal epithelial hyperplasia of Heck</td>
<td>13, 32</td>
</tr>
<tr>
<td>Conjunctival papillomas/carcinomas</td>
<td>6, 11, 16</td>
</tr>
<tr>
<td>Condyloma acuminata (genital warts)</td>
<td>6, 11, 30, 42, 43, 45, 51, 54, 55, 70</td>
</tr>
<tr>
<td>Cervical intraepithelial neoplasia (CIN)</td>
<td></td>
</tr>
<tr>
<td>Unspecified</td>
<td>30, 34, 39, 40, 53, 57, 59, 61, 62, 64, 66, 67, 68, 69</td>
</tr>
<tr>
<td>Low risk</td>
<td>6, 11, 16, 18, 31, 33, 35, 42, 43, 44, 45, 51, 52, 74</td>
</tr>
<tr>
<td>High risk</td>
<td>16, 18, 6, 11, 31, 34, 33, 35, 39, 42, 44, 45, 51, 52, 56, 58, 66, 68, 70</td>
</tr>
<tr>
<td>Cervical carcinoma</td>
<td>16, 18, 31, 45, 33, 35, 39, 51, 52, 56, 58, 66, 68, 70</td>
</tr>
</tbody>
</table>
1.3.3 Genome organization

HPV is a double-stranded DNA virus packed within icosahedral capsid. The viral particle is around 55 nm in diameter (Zheng and Baker, 2006). The genome is in the form of covalently closed circular DNA of about 8 kb length, which may vary slightly among different HPV types (7904 bp for HPV16). The capsid is composed of 72 capsomeres. The genomic organization of all papillomaviruses is remarkably similar. Viral DNA is associated with cellular histones (Howley, 1996) to form a chromatin-like complex. All coding sequences are located on one DNA strand only. The HPV genome consists of three regions - early genes, late genes and a long control region (LCR or non-coding region [NCR] or upstream regulatory region [URR]) – separated by two polyadenylation (pA) sites: early pA (AE) and late pA (AL) sites (Figure 1.6). The early region occupies 50% of the genome that encodes six common open reading frames (E1, E2, E4, E5, E6 and E7) (Danos et al., 1982). The late region occupying 40% of the genome and lying downstream of the early region encodes L1 and L2 ORFs for translation of a major (L1) and a minor (L2) capsid protein. The LCR region (about 850 bp occupying 10% of the genome), bears multiple transcription factor binding sites and the origin of replication (Zheng and Baker, 2006). There are two promoters in the HPV-16 genome - P97 promoter lying upstream to E6 ORF is responsible for the early gene expression (Smotkin and Wettstein, 1986) and the P670 promoter lying within the E7 ORF is responsible for late gene expression (Grassmann et al., 1996). The HPV-16 P97 promoter is tightly controlled by upstream cis-elements in the LCR. These include four consensus E2-binding sites (E2-BSs), ACC(N6)GGT (Androphy et al., 1987), and interact with cellular transcription factors as well as viral transactivator/repressor E2. The resulting early primary transcripts (pre-mRNAs) undergo extensive alternative splicing generating at least 14 species of mRNA transcripts with various coding potential. The HPV-16 P670 promoter (Grassmann et al., 1996) is a late-promoter. Its activity is detectable only in differentiated keratinocytes with vegetative viral DNA replication (Barksdale and Baker, 1993).
Figure 1.6: Organization of the HPV genome (J. Doorbar, Clinical Science, 2006). The HPV16 genome (7904 bp) is shown as a black circle with the early (p97) and late (p670) promoters marked by arrows. The six early ORFs [E1, E2, E4 and E5 (in green) and E6 and E7 (in red)] are expressed from either p97 or p670 at different stages during epithelial cell differentiation. The late ORFs [L1 and L2 (in yellow)] are also expressed from p670, following a change in splicing patterns, and a shift in polyadenylation site usage [from early polyadenylation site (PAE) to late polyadenylation site (PAL)]. All the viral genes are encoded on one strand of the double-stranded circular DNA genome. The long control region (LCR from 7156–7184) is enlarged to allow visualization of the E2-binding sites and the TATA element of the p97 promoter. The location of the E1- and SP1-binding sites is also shown.
1.3.4 Function of viral proteins

**E1 and E2:** The E2 open reading frame (ORF) encodes at least two or three different proteins all of which act as transcription factors and regulate viral transcription (Bouvard et al., 1994). Full-length HPV 16 E2 is 42kDa in weight and comprises of 3 functional domains: a 200-amino acid amino-terminal activation domain, a 100-amino acid carboxy-terminal DNA-binding domain, joined by a flexible hinge region (Davidson et al., 2001). Full-length E2 serves as a transactivator and a splice variant serves as p97 promoter repressor (Doorbar, 2006). The LCR contains four E2 binding sites (E2-BS), which is a palindromic motif [AACCg(N4)cGGTT] (Dell et al., 2003). E2 also binds viral origin of replication and recruits E1 helicase. E1 has a weak affinity for a consensus motif (ACgNAT) repeated 6 times in the viral origin (Doorbar, 2006). HPV16 E1 is 70 kDa protein that helps in viral replication by binding to cellular proteins like RPA (replication protein A) and DNA polymerase α primase (Han et al., 1999). The E2 protein also binds viral episomes to mitotic chromosomes facilitating proper segregation (You et al., 2004). E2 may also help in efficient genome encapsidation during natural infection (Buck et al., 2004).

**E6 and E7:** The E7 protein (11kDa) binds with pocket proteins pRb, p107, p130 and dissociates the E2F family of transcription factors, which subsequently transactivates cellular proteins required for replication (cyclins A and E). E7 also associates with other proteins like histone deacetylases (Brehm et al., 1999), components of the AP1 transcription complex (Antinore et al., 1996) and the cyclin-dependent kinase inhibitors p21 and p27 (Funk et al., 1997). E7 drives cells through mitosis in differentiating epithelium overcoming the block to cell-cycle progression (Noya et al., 2001). The E6 protein (16 kDa) primarily mediates p53 ubiquitination and degradation. This prevents growth arrest or apoptosis in response to E7-mediated cell-cycle entry in the differentiated epithelial layers. The anti-apoptotic role of E6 protein is emphasized further by its association with Bak (Thomas and Banks, 1998) and Bax (Li and Dou, 2000). The high-risk HPV E6 protein contains a C-terminal PDZ-binding domain and binds and degrades several cellular proteins containing PDZ domain (Zeitler et al., 2004).
This can mediate suprabasal cell proliferation (Nguyen et al., 2003). The high-risk HPV E6 protein can also form a tripartite complex with p53 and the cellular ubiquitin ligase E6AP (E6-associated protein), leading to proteosome-mediated p53 degradation (Huibregtse et al., 1993). The high-risk E6 can also activate the catalytic subunit of telomerase [hTERT (human telomerase reverse transcriptase)] (Klingelhutz et al., 1996).

**E5:** The E5 of HPV-16 is a small hydrophobic polypeptide (8kDa), with weak transforming activity (Leptak et al., 1991; Pim et al., 1992). HPV E5 is a transmembrane protein residing predominantly in the ER (endoplasmic reticulum). It can also associate with the vacuolar proton ATPase and retard endosomal acidification (Hwang et al., 1995), which affects the recycling of growth factor receptors increasing EGF (epidermal growth factor)-mediated receptor signalling and maintenance of replication competent environment in the differentiated epithelial layers (Crusius et al., 2000).

**E1^E4:** The E1^E4 ORF is located in an early gene region, yet it appears to be expressed as a late gene with a role in productive infection. The E1^E4 protein (17kDa) is exclusively localized within the differentiating layer of the infected epithelium and induces the collapse of the cytokeratin network (Wang et al., 2004) and facilitates mature virion release by affecting integrity of epithelial cornified envelope (Lehr et al., 2004). The E4 proteins encoded by several HPV types antagonize E7-mediated cell (Davy et al., 2002). HPV16 and 18 E4 can also associate with E2 (Sorathia et al., 2004).

**Late proteins:** L1 and L2 proteins are 55kDa (531 amino acid residues) and 50kDa in weight, respectively. The capsid of infectious virions contains 360 copies of L1 protein (major capsid protein) organized into 72 capsomeres (Modis et al., 2002). A single L2 protein (minor capsid protein) is thought to be present in the centre of each pentavalent capsomere at the virion vertices (Modis et al., 2002). Both proteins mediate efficient virus infectivity.
1.3.5 HPV life cycle: productive versus non-productive

HPV infects the basal keratinocyte layer of the cervical squamous epithelium and restricts its interaction with the mucosa with no indication of viremia. HPV lifecycle is orchestrated with the normal epithelial differentiation (Figure 1.7). HPV infects through microwounds that give entry to basal layer cells of the epithelium.

1.3.5 A Non-productive phase of infection

Viral genome is maintained episomally in basal cells with low levels of viral replication. Initially, after infection, the viral genome is amplified to 50-100 copies per infected cell and is maintained as a multicopy plasmid. This is non-productive phase of viral life cycle, where viral DNA is replicated bi-directionally (theta structure intermediates) without production of progeny virus (Elsa and Lambert, 1997)

1.3.5 B Productive (vegetative) phase of infection

In the basal keratinocytes, it expresses very low levels of early proteins. Expressions of viral replication proteins, E1 and E2, are required. The E2 protein is important for initiation of viral DNA replication and genome segregation. Viral genome replicates along with cellular DNA during S-phase of cell cycle and the replicated genomes are partitioned equally during cell division (Doorbar, 2006). As the cell leaves the basal layer for the proliferative suprabasal compartment, the virus maintains itself in episomal forms with barely-detectable levels of oncogenic expression. Normally, the suprabasal cells exit cell cycle and begin terminal differentiation (Madison, 2003). In the differentiating compartment of the epithelium, there is a huge upregulation of viral gene-expression and replication. Outburst of viral replication in differentiated cells is called vegetative replication. Viral oncoproteins (E6 and E7) uncouples the host cell from cell-cycle arrest and utilizes host DNA polymerase for viral replication. This is necessary, along with E1 and E2, for the episomal replication above the basal layer (Doorbar, 2006). E7 associates with pRb (Burd, 2003), histone deacetylases (Brehm et al, 1999) and the cyclin-dependent kinase inhibitors p21 and p27 (Funk et al, 1997) to release the cells from cell-cycle arrest. E6 inactivates p53 to check DNA repair and apoptosis (Burd, 2003). So, the restraint on cell-cycle progression is lost and normal
terminal differentiation does not occur (Sherman et al., 1997). Finally, highly immunogenic viral capsid proteins (L1 and L2) are formed and mature viral particles are released from the terminally differentiated naturally dying epithelial cells (Doorbar and Gallimore, 1987). The events linking genome amplification with synthesis of the capsid proteins are dependent on mRNA splicing and termination of transcripts at the late polyadenylation site (Doorbar, 2006). Up-regulation of the differentiation-dependent promoter (p670 in HPV16) is also critical for this (Bodily and Meyers, 2005), which depends on cell-signalling changes (Spink and Laimins, 2005), and leads to increased levels of viral proteins necessary for replication (i.e. E1, E2, E4 and E5).

Figure 1.7: Schematic representation of HPV life cycle in stratified epithelium.[A] Viral life-cycle synchronized with epithelial tissue differentiation (1) HPV infects basal cells of the squamous epithelium through microabrasions. Expression of the viral early genes, results in amplification of the viral DNA; (2) viral transcription and replication are auto regulated at a low level, resulting in a steady-state, low-level maintenance mode; (3) After a cell leaves the cell cycle and undergoes progressive differentiation, a vegetative mode of viral amplification follows and mature virions are released; [B] HPV mediated change in squamous epithelium.
1.3.6 Viral genome integration

Cervical cancer cells often contain chromosomally integrated HPV DNA or a mixture of both integrated and non-integrated viral DNA. Integrated HPV DNA is also found in a subset of high-grade lesions (Fujii et al., 2005). It can also be found in some CIN1 lesions suggesting integration to be an early event in carcinogenic progression (Peitsaro et al., 2002). The viral E2 gene involved in negative regulation of E6 and E7 expression get disrupted (Arias-Pulido et al., 2006) thus leading to excess production of E6 and E7 oncoproteins. Therefore, cells that contain integrated HPV DNA have a selective growth advantage over cells containing non-integrated HPV DNA (Jeon and Lambert, 1995). Integration into host genome occurs at fragile sites participating in cancer development (Yu et al., 2005). Other factors, like viral DNA methylation and chromatin organization (Kalantari et al., 2004) can also regulate expression from integrated sequences. Only one copy is transcriptionally active in concatemeric integration (Van Tine et al., 2004b).

1.4 FACTORS CONTRIBUTING TO CACX DEVELOPMENT

The possible factors that attribute to HPV infection leading to CaCx are categorized as follows:

- Viral factors
- Environmental and demographic factors
- Host genetic factors
1.4.1 Viral Factors:

The viral factors likely to facilitate persistent HPV infection and progression to cervical carcinogenesis include -

- Genomic differences among the viral types and variants with ability of host immune-evasion in some but not in others (O'Brien and Saveria Campo, 2002).
- Viral persistence and viral load (Ho et al., 1995; Josefsson et al., 2000; Ylitalo et al., 2000)
- Viral DNA integration, which usually disrupts E1 or E2 open reading frame, keeping E6, E7 ORFs and LCR intact.

1.4.2 Environmental and demographic factors

Early age of sexual intercourse (de Boer et al., 2006) and promiscuity (Williams et al., 1994) are established independent risk factors for the disease. Multiparity, hormonal contraception, poor hygiene and nutrition are also known to play significant roles in the development of CaCx (Bharadwaj et al., 2009; Castle, 2004; Ursin et al., 1994). An increased HPV prevalence among the oral contraceptive users may have resulted from the responsiveness of the hormone-binding elements in the viral genome, the host immunologic response or changes in the cervical anatomy due to hormonal influence (de Villiers, 2003). Pregnancy induces increased levels of estrogen and progesterone which may modulate the immune response to HPV and influence risk of persistence or progression (Sethi et al., 1998). Smoking leads to the presence of tobacco metabolites in cervical secretions, which is considered as a risk factor (Haverkos, 2004; Hellberg and Stendahl, 2005). HPV prevalence shows different patterns of infection with age in different populations. An inverse relation between age and HPV prevalence has been attributed to the development of acquired HPV immunity over time after HPV exposure (Kjaer et al., 2000). In contrast, many other studies showed increase in HPV infection in older age groups (Cuzick et al., 1999). A constant prevalence of HPV across different age groups has also been reported (Franceschi et al., 2005).
1.4.3 Host Genetic Factors:

Cancer is a multifactorial disease and HPV infection is the major etiological agent for CaCx. A large number of sexually active women are infected with HPV. However, only a fraction of these women develop CaCx after a long latent period. Further, it has been demonstrated that biological daughters of women with CaCx have an increased risk as compared to adopted daughters, with an approximately 50% reduced risk for half-sisters (Magnusson et al., 1999). These facts point to the role of other host genetic cofactors in carcinogenic transformation of HPV infected cervical epithelium and disease progression. The various susceptibility factors can be divided into two groups namely,

- those involved in viral persistence and
- those that facilitate progression to carcinoma.

1.4.3 A Host factors involved in viral persistence

Sexually active young women commonly incur infection with high-risk HPVs, but, less than 1% of them develop CaCx (Bosch et al., 1995). Moreover, immunosuppressed patients suffering from HIV infection or with renal transplant and patients infected with HIV show a greater incidence and longer persistence of HPV infection (Brown et al., 2000; Frisch et al., 2000). These observations point towards the role of host immunogenetics in persistent oncogenic HPV infection, which subsequently leads to the cancer.

Natural immunity: Innate and adaptive immune responses block viral infection and eliminate infected cells (Abbas et al, 2000). HPV infection is spontaneously cleared by 70% to 90% infected individuals within 12 to 24 months after diagnosis (Ho et al., 1998). Immune vigilance induces regression of HPV-related lesion (Wu et al, 1994). The nucleated basal keratinocytes can present viral peptides by MHC class I, to CD8 (+) cytotoxic T lymphocytes. Langerhans cells in the cervical mucosa can capture viral antigens for presentation to both cytotoxic and helper T cells [CD4(+)] by MHC classes I and II, respectively (Roncalli et al., 1988). The lymphocytes are present in the mucosa-associated lymphoid tissue (MALT) in the female genital tract (Parr et al,
Cellular immunity plays a determining role being the only efficient response against non-lytic HPVs and also to neoplasia (Goncalves and Donadi, 2004). Th1 cells (a subset of helper T cells) produce IL2, IFNγ, etc., that promote cellular immunity (Malejczyk et al., 1996). IFNγ blocks HPV 18 mRNA transcription (Woodworth et al., 1992). TNFα produced by keratinocytes and T-lymphocytes can diminish E6/E7 gene expression in HPV 16 or 18 infected cells that have not been yet transformed (Goncalves and Donadi, 2004). It has been found that stimulation of peripheral lymphocytes with antigenic fragments from E6, E7 and L1 can induce proliferation of memory cells (Luxton et al., 1996). Cellular infiltration, essentially of CD4(+) lymphocytes and macrophages, has been found in spontaneously regressing condylomas (Stanley et al., 1994).

**Impaired immunity in HPV infection:** Escape from exposure to the host immune system is an important mechanism in viral persistence. The host remains ignorant of the viral presence for long periods of time, while viral oncoproteins gradually dedifferentiate and transform the mature keratinocytes in the course of cancer development.

Minimal viral replication in basal cells, outburst of replication and E6/E7 expression in differentiating suprabasal compartment and formation of immunogenic capsid proteins in outermost epithelial layer (Cason et al., 1989) reduce exposure to active milieu of submucosal immune defenses (Kanodia and Kast, 2007). Non-secretory nature of proteins (Crum et al., 1986), absence of pro-inflammatory signals and antigen-production due to non-lytic life-cycle of HPV (Kupper and Fuhlbrigge, 2004), absence of viremia (Kanodia and Kast, 2007) are the reasons why the virus remains hidden from host immune-surveillance.

There is a loss of antigen presentation in HPV infection due to several mechanisms. These include downregulation of MHC class I transcription by HPV 16/18 E7 protein (Georgopoulos et al., 2000) repression of the bi-directional promoter for TAP1 and LMP2 genes by HPV18 E7 (Georgopoulos, 2000), E5 mediated alkalinization of Golgi complex inhibits MHC class I transport (Ashrafi et al., 2005; Reich et al., 1997) and inhibition of breakdown of invariant chain (li) hindering MHC
class II protein-folding (Zhang et al., 2003). Differential downregulation of MHC class I genes by E5 represses HLA-A and -B that determine CTL response, but not HLA-C and -E that possess NK-cell inhibitory signals (Ashrafi et al., 2005), transcriptional inhibition of MCP1 (chemoattractant for immune cells). Moreover, mutation in β2 microglobulin gene (Algarra et al., 2004), loss of HLA haplotypes due to loss of heterozygosity (Browning and Dunnion, 1997; Garrido et al., 1997), chromosomal rearrangement (Wang et al., 1998) or deletion (Torres et al., 1996) or insertion (Koopman et al., 1999) in MHC region, could also lead to loss of antigen presentation in HPV infections. Certain variants of high-risk HPVs can evade binding to MHC class I, thereby increasing their oncogenic potential. An HPV-16 E6 variant with non-synonymous mutation (R10G) within MHC class I-restricted (HLA-B7) CTL epitope (Ellis et al., 1995), another HPV16 E6 variant with aminoacid change (L83V) within class I epitope (Kast et al., 1994) and HPV16 intratype variants of E5 (Eriksson et al., 1999) have been reported in patients.

E7 and E6 bind to interferon regulatory factors (IRF) inhibiting transcription of IFN type 1 (IFN α and β) genes. E7 binds to IRF-9 and IRF-1 inhibiting IFNα and IFNβ, respectively (Park et al., 2000). HPV16 E6 binds IRF-3 inhibiting IFNβ transcription (Ronco et al., 1998). E6 and E7 also suppress expression of IL-8 (Huang and McCance, 2002) hindering chemokine-mediated infiltration of T cells and NK cells. E6 and E7 also competitively bind to IL-18 receptor (Lee et al., 2001) inhibiting IL-18 induced IFNγ production.

A skewing of cytokine profiles from Th1-produced (IL-2, IFNγ) to Th2-produced (IL-4, IL-5, IL-6, IL-10 and IL-13) is observed in cancer progression. Possibly viral mechanisms have evolved to disfavor production of Th1-cytokines that mediate CTL responses essential for controlling HPV infections. In contrast, HPV-associated cancers show presence of Th2- cytokines profile (Bonagura et al., 1999), which in itself inhibits the function of Th1-cytokines (Romagnani, 1994).

E-cadherin is necessary for contact between Lagerhans cells (LC) and keratinocytes (Hubert et al., 2005). E6-mediated downregulation of E-cadherin
HLA class I proteins: HLA class I gene products are glycoproteins expressed on the surface of nearly all nucleated cells. They present endogenous peptide antigens of altered self-cells (like virus-infected cells) that activate CD8\(^+\)-Tc-cells (cytotoxic T-lymphocytes) that in turn kill the virus-infected cells. The class I gene product is a large polymorphic transmembrane \(\alpha\)-chain of about 45 kDa sequentially divided into three extracellular domains (\(\alpha_1\), \(\alpha_2\) and \(\alpha_3\)) (figure 1.9) each about 90 amino acid long, a transmembrane domain of about 40 amino acids and a cytoplasmic anchor-segment of about 30 amino acids. The \(\alpha\)-chain is non-covalently attached to a \(\beta_2\)-microglobulin molecule of about 12 kDa encoded by a different gene outside MHC on chromosome 15. The \(\alpha_2\) and \(\alpha_3\) domains and \(\beta_2\)-microglobulin have di-sulphide loops. The \(\alpha_3\) domain and \(\beta_2\)-microglobulin are parts of immunoglobulin superfamily. The polymorphic regions that contribute to peptide-binding cleft are present in the \(\alpha_1\) and \(\alpha_2\) domains encoded by exons 2 and 3 of an HLA class I gene (A or B or C).
Matthews et al., 2003) may indirectly limit antigen presentation LC to the immune system.

1.4.3B Host immunogenetics/Major Histocompatibility Complex (MHC): Different HLA alleles may protect against or predispose towards the risk of HPV infection and cervical cancer. Results from different studies suggest a protective role for HLA DRB1*1301 (Hildesheim and Wang, 2002). HLA class II allele DRB1*1501 has been linked to a higher risk of HPV16 associated CaCx (Beskow et al., 2001) while, class I HLA allele, B*07, was found to be significantly higher among HPV16/18 positive cancer (Bhattacharya and Sengupta, 2007). Loss of an HLA class I allele HLA-B*44 in pre-cancerous lesions of the cervix has been associated with disease progression in HPV16 positive cases (Bontkes et al., 1998), and an HPV16 variant that may result in altered antigen presentation by HLA-B7 has been described (Ellis et al. 1995).

The present study has focused on the association between variations among classical MHC class I genes and cervical carcinogenesis.

The HLA (Human Leukocyte Antigen) genes of the MHC (Major Histocompatibility Complex, a tightly linked cluster of genes involved in self-nonself recognition) in the short arm of human chromosome 6 (6p21.3) (figure 1.8) are few of the major candidate genes whose variations can be studied in the pursuit to explain host genetic susceptibility towards persistent HPV infection and cervical carcinogenesis. There are three MHC regions (class I, II, and III) within the complex, each having several genes. The classical MHC class I genes comprise of HLA-A, HLA-B and HLA-C.
Antigen presentation by MHC class I-related pathway: Newly synthesized MHC class I α-chain remains bound to membrane-bound calnexin protein within endoplasmic reticulum (ER), till it is attached with β2-microglobulin, after which, class I α:β2-microglobulin heterodimer leaves calnexin and gets bound to TAP (transporters associated with antigen processing) by interacting with tapasin (TAP-associated protein). Chaperone proteins, calreticulin and Erp57 also attach to this MHC class I loading complex. MHC class I protein is maintained so until endogenous peptides (e.g., viral peptides), after getting cleaved by cytoplasmic proteasome complex, enter the ER through TAP and gets bound to the antigen-binding cleft (formed by α1 and α2 domains) of MHC class I protein. Finally, peptide-bound MHC class I protein achieves complete folding and is released from the loading complex to leave ER and get transported through the Golgi apparatus to the cell surface, where the MHC-bound peptide is presented to the T-cell receptor (TCR) of CD8+ T cell to elicit cytotoxic response against the virus-infected cell (Figure 1.10).
Figure 1.10: Antigen presentation by MHC class I (McMichael and Hanke, Nature Reviews Immunology, 2002). Simplified diagram of cytoplasmic protein degradation by the proteasome, transport into endoplasmic reticulum by TAP complex, loading on MHC class I, and transport to the surface for presentation to CD8+ T cells.

HLA nomenclature: Each HLA allele is assigned a letter (or letters) designating the locus (e.g., A, B, and C for class I) (Listgarten et al., 2008). This letter is followed by a sequence of numbers. The first two digits describe the allele type, which mostly corresponds to the historical serological antigen groupings. The third and fourth digits designate the allele subtypes based on synonymous amino acid change. The fifth and sixth digits are used to distinguish alleles based on synonymous substitutions, while the seventh and eighth digits distinguish alleles based on differences in the non-coding regions (i.e., the introns or the 5' or 3' untranslated regions). Sometimes a letter (N/L/S/A/Q) follow after the digits corresponding to the expression level of the protein (Figure 1.11).
1.4.3C Host factors that facilitate progression to carcinoma.

The viral proteins E7 and E6 inactivate pRb and p53, respectively, and also interact with many other cell-cycle regulatory proteins to finally uncouple the differentiated cells of suprabasal epithelium from cell-cycle check point and stimulate uncontrolled cell division (Milde-Langosch and Riethdorf, 2003).

Figure 1.12: E6 and E7 proteins primarily inactivate p53 and pRb, respectively. E7 also inhibit many other cell cycle regulators. Interaction between E7 and pRb and pRb-related pocket proteins (p107, p130) targets the latter to ubiquitin-mediated degradation (Dyson et al., 1992) and releases the
transcription factor E2F from pRb. E2F family of proteins transactivate cyclins, cyclin dependent kinases (CdK). E2F may induce cyclin E and cyclin A (Schulze et al., 1995) transcription. Cdk4 overexpression was detected in approximately 70% of cervical tumors (Cheung et al., 2001).

However, E2F also transactivates p21 Cip1 (Hiyama et al., 1998) and p16 INK4A (Khleif et al., 1996). Therefore, p16 overexpression has been found to be strongly associated with HPV 16 or 18 E6/E7 oncogene expression (Riethdorf et al., 2002). Hence, a strong diffuse p16 immunostaining in combination with Ki-67 and cyclin E is thought to be a biomarker for HPV infection in squamous lesions (Keating et al., 2001).

E7 ineracts with p21Cip1 (Jones et al., 1997) and inhibits p21-mediated cyclin E inhibition. E7 can also interact with p27Kip1 (Zehbe et al., 1999). This interaction can reactivate DNA synthesis in differentiated keratinocytes, because p27 is involved in keratinocyte differentiation (Missero et al., 1996). E6 protein inactivate p53 and prevents cell-cycle arrest for DNA repair, or apoptosis (Kessis et al., 1993).

Many studies have found associations between cervical cancer and polymorphisms in or methylation profile of other genes like Fragile Histidine Triad (FHTT) (Virmani et al., 2001) and Methylene Tetrahydrofolate Reductase (MTHFR) (Goodman et al., 2001).

1.5 PREVENTION, MANAGEMENT AND TREATMENT OF CACX

CaCx is preventable and curable if detected at an early stage (WHO, 2006). The 5 year survival rate of cervical cancer when detected at the earliest stage is 92%, and the combined 5 year survival rate for all stages is 71% (American Cancer Society 2009).

In the last 2 decades, etiologic studies of CaCx have identified several HPV types as necessary for the development of cervical cancer, coupled with a few additional
intervening cofactors that promote the oncogenic potential of HPV infection. The association is universal. These findings have resulted in great advances in the prevention of this disease on two fronts (Bosch et al, 2006; Munoz et al, 2009).

1.5.1 Vaccines

**Prophylactic**: The first is in primary prevention by the use of prophylactic HPV vaccines. Two safe and efficacious prophylactic HPV vaccines have been developed using virus-like particles (VLPs). The quadrivalent vaccine (Gardasil) contains VLPs of HPVs 16 and 18, responsible for about 70% of cervical cancers, and VLPs of HPVs 6 and 11, which cause about 90% of genital warts. The bivalent vaccine (Cervarix) contains only VLPs of HPVs 16 and 18. Both vaccines have been shown to have a high efficacy for the prevention of high-grade precancerous lesions of the cervix (i.e. cervical intra-epithelial neoplasia of grades 2 and 3) and this protection has been shown to last at least 5 years (Ault et al 2007; Paavonen et al 2007; Villa et al 2006). In addition, the quadrivalent vaccine has been shown to prevent high-grade precancerous lesions of the vulva and vagina caused by HPVs 16 and 18 and genital warts caused by HPVs 6 and 11 (Ault et al 2007; Paavonen et al 2007; Villa et al 2006).

**Therapeutic**: The commercial preventive HPV vaccines, Gardasil and Cervarix, use HPV virus-like particles to generate neutralizing antibodies against HPV major capsid protein L1. However, they do not exert therapeutic effects on existing lesions and are unlikely to have an immediate impact on the prevalence of CaCx due to their cost and limited availability in developing countries, which account for more than 80% of cervical cancers. Thus, there is an urgent need for therapeutic HPV vaccines. Therapeutic HPV vaccines can eliminate preexisting lesions and infections by generating cellular immunity against HPV-infected cells. HPV E6 and E7 oncoproteins represent ideal targets for therapeutic intervention because of their constitutive expression in HPV-associated tumors and their crucial role in the induction and maintenance of HPV-associated disease. Research on therapeutic vaccines is ongoing (Su et al., 2010). Several approaches are being adopted in this direction, which include live vector-based, peptide/protein-based, nucleic acid-based and cell-based vaccines.
targeting E6 and/or E7 antigens (Ma et al 2010) to accomplish the successful treatment of established lesions for the control of cervical cancer.

1.5.2 Screening

The second front is in secondary prevention by increasing the accuracy of CaCx screening. Several studies have shown that HPV DNA detection assays are more sensitive than Pap cytology in identifying cervical intra-epithelial neoplasia of grades 2/3 and CaCx and suggest that they should be used as the primary screening test followed by triage with cytology or visual inspection according to the facilities existing in the various regions (Cuzick et al 2006). Evidence suggests that if the current HPV vaccines were introduced into developing countries and combined effectively with appropriate secondary cervical screening strategies, the lifetime risk of developing CaCx could be considerably reduced (Franco et al 2008). A recent report from India showing that a single round of HPV testing was associated with a 50% reduction in the number of advanced CaCx and deaths from cervical cancer, opens great perspectives for the prevention of cervical cancer in developing countries (Sankaranarayanan et al 2009).