Potential biomarkers for cervical cancer screening and treatment

Mari Uyeda¹

¹ AC Camargo Cancer Center, Department of Radiotherapy of the Hospital, Tamandaré St, 753-Liberdade, São Paulo, Brazil, 01509-020

Abstract: Cervical cancer (CC) is the most common cancer in women worldwide and is almost always associated with repeated infections by human papillomavirus (HPV). Screening by traditional tests combined with biomarker identification techniques for low- or high-grade injuries, are becoming increasingly important in diagnosis and prognosis, avoiding countless deaths. This article explores existing literature on the main serum biomarkers and the identification of biomarkers associated with the oncogenesis of HPV expression in the identification of pre-cancerous lesions and of CC for an effective treatment, with reduction of recurrence, as well as we explain strategies for the development of biomarkers based on DNA, proteins and other markers. We also describe how markers of chromosomal instability host DNA, promoting hypo- or hypermethylation of DNA, as well as polymorphisms and epigenetic events in the p53 gene. Finally, we discuss changes in gene expression using cDNA microarray techniques and changes in the expression of proteins and markers identified through mass spectrometry.

Keywords: Cervical cancer; Biomarkers; Human papillomavirus; Pap smear; cervical intraepithelial neoplasia

1. Introduction

Cervical cancer (CC) is considered the fourth type of cancer that mostly affects women worldwide, with an estimated incidence of 569,847 cases and 311,365 deaths, according to the latest Globocan report [1], that is, every 2 minutes a life is lost to this disease. CC is almost caused always by repeated infections with human papillomavirus (HPV). Although it has started to show a decline in developed countries, this reality does not yet occur in developing countries, being high among women in about 43 countries [2]. About 85% of all deaths from CC originate from low-income countries and are 18 times higher than in developed countries [3]. This inconsistency in the mortality rate between developed and developing / underdeveloped countries is due to a lack of awareness, screening programmes, inaccessibility to adequate diagnoses and effective treatment procedures, in addition to greater exposure to the risk factors that lead to the CC. In addition to HPV infection, poor hygiene, smoking, use of oral contraceptives, to exposure to diethylstilbestrol and genetic predisposition are particularly common among women of low socioeconomic status [4]. Although HPV infection cannot induce cervical carcinogenesis alone, it is a primary requirement for most cases of CC [5]. On November 2020, a Global Health Strategy was launched by the World Health Organization (WHO) to accelerate the elimination of CC worldwide.

There are more than 200 different types of HPV that can infect epithelial cells [6], 40 of which may present tropism in mucous tissue, which may be of low or high risk (L-SIL and H-SIL) [7]. Anogenital warts associated maybe with L-SIL, while cervical intraepithelial neoplasia (CIN) associated with H-SIL, where about 70% of globally identified CC cases are associated with HPV
16 and 18. H-SIL leads to the integration of viral DNA in the hosts and consequent expression of viral proteins E6 and E7, which are responsible for causing the inactivation of suppressors tumours relevant, thus leading to the escape of apoptosis, one of the milestones for the development of cancer.

1.1. Pathophysiology

Cervical adenocarcinoma differs from squamous cell carcinoma in that it begins in the glandular epithelium of the endocervical canal and begins immediately with adenocarcinoma in situ. The time difference between HPV infection and the development of CC is generally 20 years; therefore, rapid progression of CC rarely occurs [8]. Figure 1 shows how the formation of cervical intraepithelial neoplasia occurs.

The CIN is different, depending on its degree. CIN1 is an L-SIL, where 70 to 80% of lesions regress spontaneously without treatment or become undetectable [9, 10], reflecting a state of infection and not a stage in the development of the disease. Detection of CIN1 after HPV infection, therefore, does not automatically represent disease progression. Also, clearance attributed can be to the inability to detect the infection [11].

CIN2 and CIN3 are considered H-SIL, but they develop in different ways. NIC2 has a lower incidence of cancer, with an annual regression of 15 to 23%, reaching 55% in 4-6 years [8]. About 2% of NIC2 develops for NIC3 in the same period. CIN3 is pre-cancer with the potential to progress to invasive cancer at a rate of 0.2% to 4% in 12 months [10]. Untreated CIN3 is 30% likely to become invasive cancer over 30 years, although only about 1% of correctly treated CIN3s become invasive [10]. CIN3 is pre-cancer with the potential to progress to invasive cancer at a rate of 0.2% to 4% in 12 months [10]. Untreated CIN3 is 30% likely to become invasive cancer over 30 years, although only about 1% of correctly treated CIN3s become invasive [8].

Most women infected with HPV do not develop CC because the immune response controls the infection, preventing the development of cervical lesions and their progression to cancer [12]. Thus, only a small fraction of infected women cannot control the infection and develop CC. This fact suggests that additional factors may influence the progression of the CIN to CC or its regression.

**Figure 1** Schematic representation of the events resulting from HPV infection

Source: adapted from Chan et al, 2019

HPV needs to infect the cells of the basal epithelium located in the cervical transformation zone, which actively replicates and differentiates, to establish the infection. The HPV genome (Figure 2) is composed of early proteins (E1, E2, E4, E5, E6 and E7) and late structural proteins (L1 and L2) [13, 14]. The E6 and E7 proteins are two transcriptional units that play a causal role in the development of CC. E6 promotes the degradation of protein 53 (p53) (usually through ubiquitination) and E7 which inactivates the retinoblastoma protein (pRb). In the hypophosphorylated state, proteins in the pRb family can bind to transcription factors such as E2F and repress the expression of genes involved in DNA synthesis and cell cycle progression. E7 can bind to pRb, which leads cells to enter the S phase prematurely by breaking down the pRb-E2F complex [15].

**Figure 2** - Organization of the HPV genome and its integration in the host cells

Source: adapted from ZUR HAUSEN et al., 2002.

Regarding oncoprotein E6, it can increase the induction of hypoxia inducing factor 1 alpha, which consequently leads to increased expression of vascular endothelial growth factor (VEGF),
which makes it a therapeutic target valid for cervical cancer. In addition to the action of E6 and E7 in malignancy, the oncoprotein E5 stands out for its role in cervical carcinogenesis.

E5 can contribute to cervical carcinogenesis by several mechanisms, including activation of the epidermal growth factor receptor pathway, modulation of inflammatory signalling pathways, induction of angiogenesis through VEGF and inhibition of apoptosis [15, 16].

Regardless of the type, infectious viral particles reach germ cells in the basal layer likely through micro-trauma to the mucosa [17]. Penetrating sexual intercourse is not strictly necessary for the transmission of the virus, and HPV types can be transferred to the cervix from the initial infection at the vaginal entrance [18]. About a third of women who have levels of HPV infection detected by DNA testing have cytopathological abnormalities detected since cytological changes are less sensitive for HPV detection than molecular testing. High-grade squamous intraepithelial lesions are more likely to be caused by HPV-16 infection and related types; on the other hand, HPV-18 associates the contamination with a low level of this type of lesions [19]. This fact may explain, at least partially, the low efficiency of screening for endocervical lesions and the increase in the proportion of adenocarcinoma, which is associated with HPV-18 infection [20].

**Table 1** – Description of the specific functions of each HPV gene

Source Adapted from Dasari et al., 2015

<table>
<thead>
<tr>
<th>HPVs</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, and 59</td>
<td>Defined by the International Agency for Cancer Research as causing high-risk types, with types additional being known as 'possibly' cancer-causing (68 and 73) [21]. Persistent high-risk HPVs are the etiologic agent in the pathogenesis of cervical cancer [22]. After high-risk-HPV infection, several cellular changes associated with the epithelial-mesenchymal transition identified [23]. Transforms epithelial cells into mesenchymal cells, which can invade and migrate. Besides, it contributes to the metastatic progression of human cancer cells [24].</td>
</tr>
</tbody>
</table>

The link between high risk-HPV and the development of CC has contributed to the introduction of new screening programmes. The test to identify high-risk HPV recommended in the screening recommended by the WHO and European Guidelines to ensure quality in the identification of CC [25, 26]. The HPV test is efficient in detecting pre-cancerous cervical lesions, particularly in population-based cervical screening programmes [8].

Determining the risk and development prognosis, as well as the success of treatment in response to a medication or procedure is the main reason for identifying biomarkers. Tumour biomarkers are indicators of the physiological status of changes that occur during the neoplastic process. The expression of these markers may reflect several processes in progress in tumour cells, such as hyperproliferation, alteration of gene expression patterns, hyperplasia, genotoxicity, inflammation, and enzymatic changes related to the development of the tumour, among others. An ideal tumour biomarker is related to the malignant process, correlates with the tumour mass, and allows the characterisation of the tumour type, location, tumour staging, in addition to providing a prognostic assessment of the tumour tissue [8]. There are many biomarkers for CC when compared to other gynaecological cancers.

**1.2. Existing tests for the examination of cervical lesion Papanicolaou smear (oncological cytology)**

The Papanicolaou exam is still the method used to detect CC and its precursor lesions [27] and is a cytological screening test that detects abnormal epithelial cells scraped from the transformation zone of the cervix [28]. Several criteria exist to assess the degree of change, including the nucleus/cytoplasm ratio, nuclear shape and intensity of the marks and chromatin. Currently, the Bethesda classification is the most widely used for cytological diagnosis of cervical samples and categorises abnormal cells in L-SIL and H-SIL [29]. The vast majority of L-SIL
represents morphological changes associated with active HPV replication (for example, koilocytosis), while H-SIL indicates cell transformation, characterised mainly by nuclear changes. Epidemiological studies showed an impact on the incidence of CC and its precursor lesions when this methodology was used [30]. The p16 / Ki-67 tests cannot recommend screening women with ASCUS or L-SIL cytology due to insufficient high-quality evidence [31].

Despite its specificity, this methodology has limited sensitivity in detecting precursor lesions of the cervix, a fact that be attributed to the variation in the interpretation of this method has a variability of 34 to 94% in the detection of H-SIL and not decisive in the classification of atypical samples defined as “squamous cells atypical meaning of uncertainty” [29, 32].

2. Biomarkers

In 1998, the National Institutes of Health Biomarkers Definitions Working Group defined a biomarker as being (a characteristic that can be objectively measured and evaluated as an indicator of normal, pathogenic biological processes or pharmacological responses to a therapeutic intervention) [33]. Its main objective is not only to update therapies but also to improve methods to prevent an individual's risk assessment of developing cancer. Biomarkers are generally found in blood or tissues or even other body fluids, providing a signal for normal or abnormal processes. They can measure by genetic, proteomic, cellular, or molecular substances found in greater than ideal amounts in body fluids (blood, urine) of a cancer patient. A perfect biomarker test would have 100% sensitivity and specificity, but none of the currently available biomarkers achieves this [32].

Biomarkers can be DNA, RNA, and protein. DNA biomarkers can be single-base polymorphisms, chromosomal malformations, changes in the number of DNA copies and methylation of the promoter region of the gene.

RNA markers may demonstrate an increase or decrease in gene expression and microRNAs (miR). On protein biomarkers, they include surface receptors, tumour antigens, phosphorylation states [34].

2.1. Serum Biomarkers

Serum biomarkers such as SCC-Ag, CA-125, CEA and CYFRA used to predict prognosis in CC [35]. Evidence indicates that increased serum levels of CA 125 in pre-treated patients during follow-up may precede the clinical diagnosis of recurrent cervical adenocarcinoma [36]. Serum levels of vascular endothelial growth factor and C-reactive protein are often elevated in patients with CC, but decrease significantly after treatment when successful, and can consider as additional independent prognostic parameters in patients with CC [37,38]. Besides, the ratio of serum Angiopoietin-1 to Angiopoietin-2 can also be a valuable diagnostic and prognostic biomarker for CC [39].

According to the bioinformatics study, five genes are considered potential independent prognostic biomarkers for patients with CC [40]. The GPER is detected in various subcellular staining patterns and serves as a prognostic marker in the early stage of CC [41]. The use of new biomarkers can facilitate personalized treatment and improve treatment results for CC [42]. SEL1L, Notch3, SOCS3, c-Met and Ki-67 / MIB-1 also have potential values as diagnostic and prognostic indicators for CC [43, 44]. Although studies have already shown diagnostic biomarkers for patients with CC, there are few reports on potential prognostic serum markers for CC after surgery.

The recurrence rate of patients with CC is high, and early diagnosis cannot determine recovery after surgical treatments. Therefore, the discovery of prognostic markers for CC will improve the effectiveness of clinical processing. Proteomic approaches have used to understand the correlation between the HPV virus and CC pathology, and to discover putative biomarkers for the early diagnosis of cervical cancer and the drug is a mode of action [45, 46]. As a tool for
Biomarker research, mass spectrometry has led to the discovery of diagnostic tests readily available in gynaecological oncology, with an emphasis on clinical proteomics. In the technologies used to identify diagnostic markers in the early stages of CC, significant advances have achieved in three main areas: protein profile, multidimensional liquid chromatography combined with cutting-edge mass spectrometry and high-performance validation techniques. These new technologies have promise in identifying more robust, sensitive, and specific markers for CC.

The advent of high productivity processes that support ‘omic’ technologies is now contributing to the addition of new biomarkers. Thus, known biomarkers (i.e., BRCA1, ESR1, PCNA, FGFR2, CD86, EGFR, P2RX4, ETS1 and E2F4), as well as new biomolecules (i.e., KAT2B, PARP1, CDK1, GSK3B, WNK1, CRYAB, acid metabolism arachidonic, CCR6, EPBH2, NR2C1, NR2C2, CUTL1, miR-192-5p and miR-215-5p), have been reported using a multi-omics approach on transcription data associated with CC. The proteomics-based analysis shows great promise for the discovery of new and more useful biomarkers.

2.2. Identification of biomarkers associated with HPV expression oncoproteins

The detection of cellular changes caused by unregulated expression of viral oncoproteins can mark tumour progression markers and, thus, contribute to the identification of H-SIL cell populations of progression. The recognition and establishment of the pattern of alteration of these factors can define markers with high positive predictive power. The use of these markers will complement the results of other tests to identify lesions with a higher risk of malignant progression.

HPV infections that persist and express viral oncoproteins that inactivate p53 and retinoblastoma (Rb) promote increased genomic instability, accumulation of somatic mutations, and, in some cases, integration of HPV into the host genome. The association between cancer risk and histological subtypes varies substantially between HPV types, but the reasons for these differences are little known.

The identifications are based on cellular/tissue or circulating proteins that demonstrate altered expression in response to viral oncoproteins; studying changes in the methylation pattern of cellular genes that could predict the onset of neoplasia, changes in chromosomes or viral genomes in different regions and identify as modified by the viral integration event; the identification of polymorphisms associated with a better prognosis.

The integration of HPV into the host genome is a critical step in cervical carcinogenesis and is found in almost all invasive CCs and often interrupts the expression of the HPV E2 gene (regulatory protein E2), leading to increases in viral oncoproteins E6 and E7, which in turn promote immortalisation and cellular transformation. Overexpression of the E6 and E7 gene (E7 transforming protein) contributes to marked genomic instability, the accumulation of secondary mutations and malignant transformation. Also, the virus integrates with host genes and regulatory elements, which can cause structural changes in the host’s genome and transcriptional deregulation of gene expression. Integration sites often occur in fragile regions common in the genome, but these sites may be less random than initially estimated. Further evaluation of HPV integration sites by the various types of HPV could shed light on other cancer-causing genes.

The CC has 8kb HPV genomes with CpGs scattered in its genes and, as the disease progresses, the CpGs are more methylated by the host’s DNA methyltransferases, and there is a change in the expression patterns of viral genes. Hypermethylation is greater in invasive CC than in cervical dysplasias. Snellenburg et al. (2012) demonstrated that the methylation of several HPV-16, E2 binding sites is main in invasive CC than in cervical lesions in situ.

The immune response in immunosuppressed women may increase the progression of CC. The immune system detects damage, activates the secretion of interferons and promotes the secretion of cytokines, activating Langerhans cells, which have antigens in the cervix, causing the...
signals of the innate immune system to make the activation of the adaptive immune system, generating T cells specific auxiliaries 1 CD4+, which support the development of effector and cytotoxic TCDB8+ cells of memory. HPV-induced diseases are associated with a lack of particular CD4+ and CD8+. The proportion of CD8+ T cells infiltrated by the tumour in Foxp3+ regulatory T cells (Treg) is a significant prognostic factor in CC [54]. Low rates of CD8+/Treg cells and decreased HLA gene expression is associated with reduced survival in invasive CC [54].

Hereditary susceptibility to invasive CC is also of great clinical interest. The pattern of decreased relative family risk with a lower degree of genetic relationship indicates that there is a strong relationship between genetics and family aggregation [50]. Variants inherited from the germline have identified in several genes in virus-host interactions and immune function [50, 55, 56]. The genes under long-term investigation include HLA class I genes, HLA class II genes and p53 [50]. Wang et al. (2010) [56] using single nucleotide polymorphism analysis, identified IFNG (interferon, gamma), TMC6 genes (similar to transmembrane channel 6; formerly known as EVER1) and TMC8 (similar to transmembrane channel 8; formerly known as EVER2) associated with progression to CC [56].

Abnormal hypermethylation of the promoter leads to silencing or decreased expression of tumour suppressor genes in the CC [57]. These epigenetic changes can be reversible and occur at the beginning of cervical carcinogenesis. The genes considered promising are associated with functions such as cell cycle control, apoptosis, cell signalling and DNA repair. Relevant genes include DAPK, CDH1, RASSF1, CDKN2A, FHIT, MGMT and RARB. The cancer cell has a genome with global hypomethylation, in addition to hypermethylation of the regional promoter. Progressive hypomethylation occurs with an increased progression from dysplasia to invasive CC [57]. Hypomethylation can play a relevant role in the development of invasive diseases. Table 2 shows each biomarker and its expression in the CC.

Table 2. Expression of Biomarkers in the CC.

2.3. Assessment of protein levels involved in cells cycle control

HPV infection promotes unregulated replication of the host genome and, thus, several proteins involved in the control of the cell cycle have their levels increased or decreased. Studies aim to analyse the levels of proteins related to this event, such as p53, p16INK4a, MCM2-7, EGFR, cyclin D, cyclin E, p21WAF1, p27KIP1, among others [56].

The largest number of studies found in the literature focuses on the assessment of p16INK4a protein levels in CC samples, both in histological sections and in derived from cervical stains. The HPV E6 and E7 proteins are known to promote the degradation of tumour proteins p53 and Rb, respectively. This interference activates a minus feedback process that results in the exacerbated expression of the inhibitor protein of the cyclin-dependent kinase complexes, p16INK4a [5].

Studies describe an increase in H-SIL expression in about 100% of the samples evaluated, when compared to the normal epithelium and where p16INK4a is practically undetected. Approximately 60% of L-SIL are strongly positive for p16INK4a (proliferating basal or parabasal cells), while the remaining 40% have no detectable levels of this protein, although they are positive for HPV DNA and have morphological characteristics associated with infection (koilocytosis, among others). This observation suggests that only a part of the lesions infected with L-HPV disrupted the expression of viral oncoproteins in basal and parabasal cells, whereas, in negative p16INK4a lesions, the look of oncoproteins would be very low or even null in this type of cells. Thus, L-SIL that did not have high levels of p16INK4a in the cells of the basal or suprabasal layer would be less likely to progress to H-SIL or cancer. It is worth mentioning that the H-SIL of p16INK4a is associated with lesions infected by H-HPV. The same effect is not identified in positive L-HPV lesions [57].
The analysis of this protein in samples originating from CC showed increased expression of p16INK4a in 98% of patterns derived from H-SIL. Thus, in addition to the research of this protein in CC biopsies, its research patches could be an option in the primary screening of cervical lesions whose classification has not been determined by oncotic cytology [58, 59].

2.4. Chromosomal instability markers host aneuploid DNA

CIN results from chromosomal errors during mitosis, leading to structural and numerical chromosomal abnormalities, besides generating genomic heterogeneity, it promotes inflammatory signalling by introducing double-stranded DNA into the cytosol, involving the cGAS-STING antiviral pathway. These multifaceted effects distinguish CIN as a central driver of tumour evolution and as a genomic source of interference between the tumour and its microenvironment during immunological editing and evasion [60, 61].

Aneuploidy denotes an abnormal state in the number of chromosomes; there is a change in the number of chromosomes present in the cells, characterising HPV-positive lesions, including precursor lesions and cancer [61]. Some studies indicate that the presence of aneuploidy precedes viral integration into the host genome in advanced dysplastic lesions, implying that viral integration would be a consequence and not the cause of chromosomal instability. CIN refers to errors in chromosomal segregation in progress over consecutive cell divisions [62, 63]. Both aneuploidy and CIN frequently occur in human cancer, in tumours with a high level of abnormality in the number of chromosomal copies and exhibit evidence of sustained chromosomal dysregulation [60].

Although aneuploidy can evaluate using widely available experimental techniques, such as mass DNA sequencing, fluorescent in situ hybridisation or conventional karyotype, CIN can only infer indirectly using these methods. The experimental evaluation of CIN must identify the continuous rate of chromosomal segregation and can be identified through the frequency of cells in anaphase (phase of mitosis characterised by the migration of chromatids towards opposite poles of the cell), clonal assays, sequencing of a single cell with the reconstruction of a phylogenetic tree or sequencing of a multiregional tumour considering the information of the specific copy number of the allele [62, 63, 64, 65].

It estimated that 60-80% of human tumours exhibit chromosomal abnormalities suggestive of CIN [66, 67]. CIN correlates positively with the tumour stage and enriched in recurrent and metastatic tumour specimens [60, 68, 69]. Also, complex aneuploidies and the duplication of polyploidies resulting from the entire genome are characteristic of tumour types with a predilection for metastasis, resistance to treatment and decreased overall survival, such as CC [67, 70, 71].

Oncogenic signalling, pre-mitotic replication stress and defects in centrosome replication, sister chromatid cohesion, signalling of the spindle assembly checkpoint or microtubule connections to chromosomes have established to induce CIN. Also, a wide variety of antineoplastic therapies used in the initial treatment or metastatic configurations can disrupt the fidelity of chromosomal segregation during anaphase [72, 73].

2.5. HPV Integration

Based on DNA sequence data, more than 200 types of HPV have identified that have genomic differences. Eighty-five HPV genotypes are well-characterised. Another 120 isolates are parcel characterised as potential new genotypes [74].

HPV DNA believed to randomly integrate into the host's genome during the repair process that is triggered after ruptures in the double-stranded cell genome begin. Viral integration would be an indicator of genomic instability during the cell transformation process since 80 to 90% of CC samples have integrated HPV DNA [59].
The HPV genome has a single circular double-stranded DNA molecule containing approximately 7,900 bp associated with histones [75]. All protein-coding sequences in the open reading frame presented on tape. The genome divided into three regions:

(i) The first is a non-coding regulatory region, from 400 to 1,000 bp, with long control or upper regulatory part, containing the p97 nucleus together with the sequence enhancers and silencers that regulate DNA replication through the transcription of open reading, showing a greater degree of variation in the viral genome 53

(ii) The second region prove the open reading frame E1, E2, E4, E5, E6 and E7 involved in viral replication and oncogenesis.

(iii) The third is a late region, which encodes structural proteins L1 and L2 for the viral capsid.

Several methodologies described detecting the integration of HPV DNA in the viral genome, such as the real-time polymerase chain reaction technique that allows obtaining a relationship between E2 levels (often interrupted in viral integration) and the HPV E6 / E7 genes. When not integrated, the viral genome would have a 1:1 ratio between the E2 and E6 / E7 genes and the integration of the viral genome would result in a decrease in the genetic detection of E2 [5].

Currently, the commonly used diagnostic markers include proteins related to HPV L1, E6 and E7. The main protein in capsid viruses is L1, produced in the cytoplasm. E6 and E7 are primary HPV oncoproteins with multiple cell targets, including p53, and the pRB. E6 inhibits p53 to prevent apoptosis, while E7 is the primary transforming protein and inhibits pRB to regulate cell cycle stop [76]. Telomeres are specialised structures, located at the ends of chromosomes, known to be essential for the stability of the genome [76, 77]. Telomere dysfunction and telomerase activation have previously implicated in the progression of human cancer [76, 78]. The expression of hTERT is known to be the rate-limiting factor for human telomerase activity and a more sensitive indicator of telomerase function and action than the expression levels of other telomerase subunits that are expressed constitutively in normal and cancerous cells [76]. Ki67 is a nuclear antigen expressed during all active phases of the cell cycle (G1, S, G2 and M) and the level of expression can demonstrate to determine the state of proliferation and predict tumour development [76].

Screening for these diagnostic markers can also be useful in assessing the progression of CC after the intermediate stage, as demonstrated by a prospective study previously carried out on its expression in histological, clinical formalin-fixed and paraffin-embedded samples. Cytological samples used in clinical screening tests were also performed [79]. However, the pattern of severe pre-cancerous lesions was not sufficient to conduct statistical analysis. It usually takes a long time to collect samples of H-SIL, which are the most pre-cancerous lesions in cancer.

A significant percentage of L-SIL and H-SIL have copies of HPV in episomal form, in addition to integrated patterns. This aspect interferes with the direct quantification of the number of integrated viral copies. Some studies assess the location of HPV integration in the viral genome. This criterion may be valid in women after previous treatment, since the integration site identified may be a tumour marker in a possible recurrence [59].

2.6. Evaluation of polymorphisms and epigenetic events in the p53 gene

Regarding the regulation of cell proliferation, DNA repair, apoptosis, genomic stability, senescence, and metabolic homeostasis, p53 plays an important role [80]. This protein activated by several signals and acts as a transcription factor and, when DNA is damaged, it induces the expression of cyclin-dependent p2 (kinase inhibitor) (CDK) that suppresses the cyclin-CDK complexes, resulting in the inactivation of the cycle cell in G1 phase, allowing DNA repair before S1 replication [34]. When cells cannot repair this damage, p53 induces apoptosis, activating genes such as BAX, PUMA, Noxa and PERP. The loss of p53 function allows abnormal cell proliferation, is associated with carcinogenesis, which is observed in many malignant tumours [59].
This dysfunction occurs due to the inactivation of the p53 protein by binding proteins or p53 mutations. MDM2 plays the role of a negative regulator by direct connection. The MDM2 protein acts as a ubiquitin E3 ligase recognising the N-terminal transactivation domain of the tumour suppressor p53 and as an inhibitor of transcriptional p53 activation. In response to DNA damage, MDM2 releases p53, resulting in its activation [81]. In the CC, p53 inactivated by the HPV E6 oncoprotein [59]. The results on the prognostic impact of p53 overexpression in granulosa cell tumours are conflicting. Mutations in p53 have seen in about half of patients with malignant tumours [59, 82].

p53 inactivation is associated with CC carcinogenesis. Notch1 acts as a tumour suppressor gene [59, 83] and induces cell differentiation in keratinocytes, [59, 84] by reducing cell proliferation in CC cell lines through oncogenes E6 and E7. Yugata et al. (2007) [85] demonstrated that E6 suppressed the expression of Notch1 by inactivating p53 and binds directly to the Notch1 promoter, regulating the expression of Notch1 at the transcriptional level. In CC carcinogenesis, the promotion of p53 degradation by E6 reduces Notch1 expression [86].

E6 and E7 can inactivate cervical cells independently and synergistically. INK4A inhibits E6, while E7 suppresses this inhibition. On the other hand, E6 inhibits apoptosis induced by E7, degrading the apoptosis-inducing proteins p53 [87]. This synchrony is relevant in CC [59].

2.7. Evaluation of epigenetic events

Additional genetic and epigenetic changes in the host cell genome are necessary for progression to CC. Methylation of cytosines at CpG sites in promoter regions can lead to gene silencing. DNA methyltransferases (DNMTs) responsible for CpG methylation triggered by E6 and E7, where E7 can bind directly and activate DNMT1, while E6 can positively regulate DNMT1 through p53. On the other hand, silencing E6 and E7 has been shown to reduce DNA methylation of tumour suppressor genes and restore the phenotype transformed into CC cells. Increased levels of DNA methylation of several tumour suppressor genes (candidates) are associated with CC and a subset of its H-SIL CIN 2 and 3 [88].

Methylation is the epigenetic phenomenon by which a gene silenced, being a means of regulating gene expression [89]. Particularly susceptible to the effect of methylation are CpG dinucleotides, which contain cytosine and guanine bases. Most of these dinucleotides are in small regions, called CpG islands, which demethylated in normal cells. That zone founding in the regulatory sequences present at the 5’ end of each gene, where the genetic promoters located. Methylation of these regions interrupts the transcription of genes by silencing their originator. This event can also be observed in the histone structure [62].

Epigenetics is a well-studied phenomenon, involved in several biological processes, including development, cell differentiation, immune function, and various stages of carcinogenesis [63]. Reference studies using exome sequencing have identified only a small proportion of mutations in classic tumour suppressor genes, such as p53 and Rb1 in CC, because of these proteins inactivated by oncogenic H-HPV proteins, such as E6 and E7 [5, 63].

The hypermethylation of the specific gene and the role of global DNA hypomethylation has been well-recognised during the progression from premalignant conditions to malignant diseases fully developed in the CC. The analysis of global hypomethylation explains the underlying mechanisms of carcinogenesis that may occur due to crude chromosomal instability and reactivation of transposable elements and, subsequently, oncogenic expression [35, 63].

The acetylation, methylation and phosphorylation of histones are fundamental in gene expression in normal cells. Any modification can facilitate the inadequate transcription and transformation of normal cells, interrupting cell proliferation and cell death [64]. Immunohistochemical analysis of histone H3 phosphorylation and acetylation shows an association with CC progression [65].
Viral oncoproteins derived from HPV E6 and E7 bind histones acetylases and histones acetyltransferases, which are the main enzymes involved in the modification of histone and the regulation of gene expression [5, 65].

The synergistic effect of DNA hypermethylation and histone acetylation is necessary for the modulation of expression throughout the genome, as well as at the individual genetic level. miR-mediated post-transcriptional regulation of gene expression is an epigenetic event implicated in several human diseases, including cancer. Studies have highlighted the role of miRs during the progression of CC based on their look in clinical samples and their usefulness as a diagnostic marker. Also, DNA methylation and histone modifications, and miR regulation of gene expression provide an additional layer of complexity to understand the role of epigenetics during CC progression [61, 63]. Overexpression of miR-9 can lead to epithelial-mesenchymal transition, as regulated by c-Myc and PROX1 [90, 91]. It also significantly decreases E-cadherin and increases vimentin by increasing cellular motility and invasiveness [92]. Besides, miR-9 promotes EMT and metastasis by directly regulating KLF17 expression [93].

Dysregulation of miR-205 in CC occurs readily, serving as a tumour suppressor or oncogene depending on the cell base [94]. As an oncogene, miR-205 is significantly overexpressed in human cervical cancer tissue and increases cervical cancer cell proliferation and migration, targeting CYR61 and CTGF [95].

Unlike genetic changes, epigenetic changes are reversible, tissue-specific and governed by gene-environment interactions. The specificity and reversible capacity of epigenetic changes favour preventive, diagnostic, prognostic, and therapeutic interventions [35]. Methylation can contribute to pathogenic processes, such as cancer, and the hypermethylation of some parts of DNA can consider a factor for tumour formation, interfering in carcinogenesis, such as the methylation of tumour-suppressor gene promoting regions, where the originated proteins control the cell proliferation and the loss of this regulation is associated with neoplastic processes. Hypermethylation of the 5th region of the gene encoding pRb is an example of silencing a tumour suppressor gene and the oncogenic process. Senescent cells or that have been affected by environmental factors, such as radiation, smoke, exposure to certain viruses, and other aspects, have a higher incidence of hypermethylation in genes associated with the neoplastic process [96].

In CC, methylation of several genes described and the examples are DcR1 / DcR2, hTERT, p73, p16, PTEN, E-cadherin, APC, MGMT, FANCF, BRAC1, hMLH1, RASSF1A, DAPK, TSIC1, FHIT, HIC1, RARβ, TIMP2 / TIMP3, CAV-1. As methylation markers, 5-hmC and 5-mC together with 5-foc and 5-caC draw the reversible cycle outline, and 6-mA participates in RNA methylation, especially miRNA [97]. The silencing of the epigenetic gene through dense methylation of DNA within the CpG islands occurs in the CC associated with HPV. Tumour suppressor genes are common targets for gene silencing in this disease. The identification of a tumour suppressor genes panel has great promise to provide a set of DNA methylation biomarkers for use in the diagnosis or prognosis of the disease [98].

The genes that encode several regulators of the Wnt / β-catenin oncogenic pathway, such as CDH1 (E-cadherin), APC and WIF1, are often silenced by dense methylation of their promoter regions in CC [98]. Other genes supposedly hypermethylated in CC with little or no methylation in normal or low-grade CINs include DAPK1, RARB, TIMP, CCA1 and FHIT [98].

Pyrosequencing bisulfite has recently emerged as a quantitative method for measuring DNA methylation at individual CpG sites within a population of DNA molecules that can be used with a variety of biological specimens, making this approach amenable to use in the clinical environment [98].

3. Perspectives in the identification of biomarkers for CC

3.1. Changes in gene expression
Studies identify differentially expressed genes using the microarray cDNA technique, such as tumorigenic and non-tumorigenic strains compared to the positive points of HPV16, with the identification of 49 differentially expressed genes, such as the C4.8 gene, which is associated with cell proliferation and its expression is increased in H-SIL [54, 60].

Comparing the gene expression between CC and normal keratinocytes, more than 500 differentially expressed genes found; among them, CDKN2A/p16INK4a, topoisomerase 2A, among others [5]. One study identified transcriptions of kallikrein seven and superoxide dismutase 2 when compared to strains immortalized with HPV16 and 18 for primary cultures of normal keratinocytes [51].

Study of intraepithelial lesions of H-SIL in 40 patients caused by infection by HPV16 showed low levels of E-cadherin after the expression of E6/E7 from HPV16 [60]. This decrease led to the proliferation and increased cell migration and invasion due to weaker adhesion properties [60]. Also, decreased of E-cadherin and b-catenin were observed in 135 CC samples, which were associated with histological differentiation, metastasis, and recurrence, suggesting that e-cadherin could be a prognostic biomarker.

3.2. Change in protein expression and markers identified using the mass spectrometry (MS) technique

Some circulating markers detected by an enzyme immunoassay have indicated to CC, including IGF-2, VEGF-C and CIFRA proteins [68]. Methylation of CDH1 and CDH2 measured in serum samples [56]. Although these standard bioanalytical methods, such as ELISA and electrochemical immunoassay, are still usual in medicine for protein evaluation, several studies have focused on technology-based MS.

The first studies for proteomic analysis, called SELDI MS (enhanced surface laser desorption and ionization mass spectrometry) were to identify circulating protein markers characteristic of situ CC. This technique allows the tracking of hundreds of proteins differentially expressed in body fluids, such as plasma, saliva, urine, and others. Some studies have described plasma protein profiles capable of discriminating patients with or without CC, with sensitivity and specificity ranging from 87 to 92% and 97 to 100%, respectively [56, 68].

Gu et al. (2015) [67] reported proteomic analysis of high-grade dysplastic cervical cells obtained from ThinPrep slides using a linear ion trap along with MS with Fourier transform. They identified more than 1000 protein biomarkers for high-grade dysplastic cervical cells, and laser capture microdissection used to isolate normal and high-grade dysplastic cells. In a study, two-dimensional gel electrophoresis used to separate proteins from 10 pairs of scaly human CC tissues and combined samples from the unaffected adjacent uterine cervix. [61]. Differentially expressed proteins were identified by matrix-assisted. Overexpression of Tyk2, S100A9 and zinc finger protein 217 in scaly CC observed.

Two-dimensional differential gel electrophoresis accompanied by MALDI-ToF MS was performed by Guo et al. (2015) [99] in quantitative analysis of plasma proteins of women with early-stage CC. Ten plasma proteins that can be as biomarkers identified, such as proteins related to lipid metabolism: APOA4, APOA1, APOE; metabolic enzymes: ceruloplasmin, endogloboidase F2, MASP2, glycoprotein CLU and proteins related to immune function. The same method used by Guo et al. (2015) [99] to research predictive markers of response to neoadjuvant chemotherapy in patients with CC. The proteins evaluated in 10 cases of the advanced CC of patients who received cisplatin-based neoadjuvant chemotherapy. In the study, overexpression of Hsp70 inhibited cisplatin’s efficacy, so Hsp70 is a potential biomarker of sensitivity to chemotherapy [99].

Yin et al. (2016) [100] identified two metabolites, phosphatidylcholine and lymphosoftidylcholine, which are significantly reduced and overloaded in CC plasma compared to patients with uterine fibroids, using MS by ultra-performance liquid chromatography. The authors suggest that these lipids may serve as new biomarkers to facilitate the diagnosis of CC. Gas chromatography-mass spectrometry, Godoy-Vitorino et al. (2018) [101] identified three
urinary metabolites, 5-oxoprolinate, erythronic acid and N-acetylaspartic acid, which discriminate positive HPV with simultaneous H-SIL of controls negative. These metabolites are known to be involved in a variety of biochemical processes related to energy and metabolism and are likely to be biomarkers for H-HPV infection.

Recently, Longuespée et al. (2019) [102] used data from a biomarker discovery study to correlate m/z of MALDI images with masses of peptides identified by LC-MS/MS in the high-grade squamous intraepithelial lesion. This concept combines LC-MS/MS-based on quantitative proteomics with MALDI images and allows for the reliable identification of peptides. The authors assigned identifications to three m/z of interest. May represent an option for further confirmation of the peptide identity.

Further analysis of protein expression is needed, as new biomarkers designed using MS-based proteomics tools can improve the diagnosis and treatment of CC.

4. Conclusion

Currently, several biomarkers associated with CC are analysed. The most convincing data are associated with the persistent detection of the H-HPV genome, as well as the study of p16INK4a expression through immunohistochemistry regarding the evaluation of HPV transcripts E6 / E7 type 16, 18, 31, 33 and 45 in CC samples has become quite popular. Serum biomarkers used to predict the prognosis; the detection of cellular changes due to unregulated expression of viral oncoproteins can characterise tumour progression markers, as well as the silencing of the look of tumour suppressor genes. It estimated that 60-80% of human tumours exhibit chromosomal abnormalities and, with the use of the cDNA microarray technique, the identification of differentially expressed genes can happen, as well as the pyrosequencing bisulfite that has emerged as a quantitative method for measuring methylation of DNA within a population of DNA molecules. The discovery of biomarkers, the understanding of the action and the importance in the development of the disease is the solution found to overcome deaths from CC in developing countries.

Abbreviation:

CC       Cervical Cancer
CIN      Cervical intraepithelial neoplasia
DNA     Deoxyribonucleic acid
HPV     Human Papillomavirus
H-SIL   High-grade squamous intraepithelial lesion
L-SIL   Low-grade squamous intraepithelial lesion
miR    microRNA
pRb    Protein retinoblastoma
Rb     Retinoblastoma
p53    Protein p53
WHO   World Health Organization

Competing interests: The authors declare that they have no competing interests.

References


Berrington de González, A.; Green, J. Comparison of risk factors for invasive squamous cell carcinoma and adenocarcinoma of the cervix: Collaborative reanalysis of individual data on 8,097 women with squamous cell carcinoma and 1,374 women with adenocarcinoma from 12 epidemiological studies. International Journal of Cancer, v. 120, n. 4, p. 885–891, 2007


Rader J. Host and viral genetics and risk of cervical cancer. In: Welsch P; ed. The role of genetics in breast and reproductive cancers. New York, Springer; 2009, 263-84


