Prevalence of human papillomavirus in oral epithelial dysplasia: Systematic review and meta-analysis

Cecilie Dovey de la Cour MD1 | Cecilie Dyg Sperling MSc1
Federica Belmonte MSc, PhD2 | Stina Syrjänen DDS, PhD3,4
Freija Verdoott PhD1 | Susanne Krüger Kjaer MD, DMSc1,5

1Danish Cancer Society Research Center, Unit of Virus, Lifestyle and Genes, Copenhagen, Denmark
2Unit of Statistics and Pharmacoepidemiology, Danish Cancer Society Research Center, Copenhagen, Denmark
3Department of Oral Pathology, Institute of Dentistry, University of Turku, Turku, Finland
4Department of Pathology, University of Turku and Turku University Hospital, Turku, Finland
5Department of Gynecology, University of Copenhagen and Copenhagen University Hospital, Copenhagen, Denmark

Correspondence
Cecilie Dovey de la Cour, MD, Unit of Virus, Lifestyle and Genes, Danish Cancer Society Research Center, Strandboulevarden 49, Copenhagen 2100, Denmark.
Email: cade@cancer.dk

Abstract
The purpose of this systematic review and meta-analysis was to estimate the overall and type-specific prevalence of human papillomavirus (HPV) DNA in oral epithelial dysplasia and assess p16INK4a overexpression in relation to HPV-status. A systematic literature search identified 31 eligible studies (832 cases) evaluating the presence of HPV DNA in oral epithelial dysplasia cases by PCR. Of these, six studies evaluated p16INK4a overexpression in relation to HPV-status. The overall pooled prevalence of HPV DNA in oral epithelial dysplasia was 27.2% (95% CI: 17.6-38.1). We observed substantial interstudy heterogeneity, which could not be explained by differences in continent, tissue type, or severity of epithelial dysplasia. HPV16 was the predominant genotype detected. Moreover, 62.2% of HPV positive and 17.8% of HPV negative oral epithelial dysplasia samples stained intensively positive for p16INK4a. This meta-analysis found that 27% of oral epithelial dysplasia harbor HPV DNA. Whether this represents a transient infection or has a carcinogenic role is unknown.

KEYWORDS
human papillomavirus, meta-analysis, oral dysplasia, p16, prevalence

INTRODUCTION

The role of human papillomavirus (HPV) in the development of oral dysplasia has remained controversial due to conflicting results on reported HPV prevalence. HPV is recognized as a causal agent in virtually all cervical carcinomas and an established risk factor for a proportion of oropharyngeal (around 30%) and several anogenital carcinomas (around 50%-80% depending on site). Based on the five latest meta-analyses published since 2010 on HPV DNA and oral cavity squamous cell carcinoma, the prevalence of HPV DNA varied between 13% and 58%, being highest in China and lowest in North America. However, the mere presence of HPV DNA in a tested oral sample does not provide evidence of HPV as a carcinogenic agent, as HPV DNA may reflect a transient infection. Other markers of viral transcriptional activity should therefore also be considered such as HPV E6/E7 mRNA and/or the overexpression of p16INK4a. When taking these markers into account, the reported attributable fraction of HPV in oral cavity and oropharyngeal carcinoma ranges from 2% to 16% and 19% to 40%, respectively. Oral cavity carcinoma (including lip) is the most common malignancy of the head and neck area with an estimated 355 000 new cases every year worldwide, hence, the crude number of HPV-associated oral cavity carcinomas may in fact be substantial.

Oral epithelial dysplasia is a known risk factor for oral cavity carcinoma development with a malignant
transformation rate of 5% to 36% depending on the severity of dysplasia. A distinct subset of oral epithelial dysplasia with specific histopathological features and high-risk HPV positivity (primarily HPV16) has recently been described. HPV16-specific dysplasias are located most frequently in tongue or floor of the mouth and characterized histologically by epithelial hyperplasia with thick deep invading rete ridges and parakeratosis, karyorrhexis, and apoptotic cells. Although progression characteristics are not fully elucidated, this HPV-associated oral dysplasia could potentially progress into HPV-associated oral cavity cancer.

A previous meta-analysis by Jayaprakash et al., including 12 observational studies, examined the prevalence of HPV16 and 18 in oral dysplastic lesions and reported an HPV16/18 DNA pooled prevalence of 25%. In another meta-analysis of observational studies, Syrjanen et al. (n = 6 studies) found that HPV DNA was more than five times as likely to be detected in oral dysplasia than in the controls. These previous meta-analyses did not examine the heterogeneity among studies on oral dysplasia further and a large number of studies assessing HPV DNA prevalence in oral dysplasia have emerged since the publication of the previous meta-analyses.

This prompted us to conduct a meta-analysis of the overall and type-specific HPV DNA prevalence in oral epithelial dysplasia. Furthermore, this is the first meta-analysis examining overexpression of p16INK4a in relation to HPV status in oral dysplasia.

2 MATERIALS AND METHODS

2.1 Literature search strategy and eligibility criteria

The study was reported in accordance with the PRISMA guidelines. We performed a systematic literature search of the Pubmed, Embase, and Cochrane Library databases up to 1 June 2018. The following MESH or EMTREE terms and relevant keywords were used in the search: precancerous conditions, carcinoma in situ, potentially malignant disorder, oral, lichen planus, leukoplakia, erythroplakia, submucous fibrosis, and human papillomavirus (the search string for Pubmed is provided in the Supporting Information). The searches were restricted to publications in English language. Reference lists of key publications were also reviewed to identify additional relevant studies.

Studies were eligible for inclusion, if a polymerase chain reaction (PCR)-based test was used to detect HPV DNA in tissue or cell samples from histologically verified cases of oral dysplasia. Studies based on saliva, gargoyle, or serum samples were thus excluded. PCR is a highly sensitive HPV DNA detection method and restricting our meta-analysis to PCR-based studies increased comparability between studies. Studies including a selected group of immunocompromised patients (eg, HIV) or less than five samples were excluded. In case of duplication of a study population in several papers, the paper with the largest sample size, the most detailed description or newest publication was included. Two authors (C.D.C. and C.D.S.) reviewed the titles, abstracts, and full text papers independently and any inconsistencies concerning identification of eligible papers were discussed to reach consensus.

2.2 Data extraction

Two authors (C.D.C., C.D.S.) independently extracted data from the included papers and any inconsistencies concerning data extraction were discussed by C.D.C., C.D.S. and F.V. to reach consensus.

For each study, information was extracted (if available) on first author, publication year, country; sample size; type of tissue; type of dysplasia; type of primer (PCR); detectable HPV genotypes; overall and type-specific HPV DNA prevalence. In studies with control subjects (ie, saliva, tissue or cell samples of normal oral mucosa), the abovementioned information was also extracted (if available) on controls. Furthermore, in studies with information on p16INK4a overexpression in relation to HPV status, information was extracted (if available) on first author; p16INK4a detection method and antibody; definition of p16 INK4a positivity; number of cases positive for p16INK4a overexpression among HPV DNA positive and negative cases.

2.3 Statistical analysis

The pooled overall HPV DNA prevalence in oral dysplasia was estimated based on a random effect model by applying an arcsine transformation to the raw proportions. The HPV DNA prevalence from the individual studies was presented with exact binomial 95% confidence intervals (CIs). The interstudy variance was determined with the DerSimonian-Laird estimator. The heterogeneity across the studies was evaluated by the I² statistic and the significance of the heterogeneity was determined with the Cochrane’s Q test. Potential sources of heterogeneity were explored in stratified analyses by continent, tissue type, PCR primer type and severity of dysplasia. The severity of dysplasia was divided into epithelial dysplasia, mixed dysplasia/carcinoma in situ (CIS), and CIS.
To validate our results in relation to the HPV DNA prevalence among normal oral mucosa samples, we performed a sensitivity analysis restricted to studies that included a control group.

In studies with information on type-specific HPV DNA prevalence, the percentage of HPV16, 18, 6, and 11 positive cases were estimated among HPV DNA positive cases. Some of the included studies in our meta-analysis detected other HPV genotypes present in oral dysplastic samples (HPV31, 33, 45, 58 and 66), however, pooled analyses of these were not possible due to small numbers. In studies with p16\textsuperscript{INK4a} protein detection in relation to HPV status, the percentage of cases positive for p16\textsuperscript{INK4a} overexpression according to HPV status was estimated. The statistical software R with the package “meta” was used to perform the analyses.\textsuperscript{17,18}

All data included in this meta-analysis have previously been published in the original studies. The authors take full responsibility for the integrity and accuracy of the data.

3 | RESULTS

3.1 | Search results

The systematic literature searches of the Pubmed, Embase, and Cochrane Library databases yielded a total of 2113

**FIGURE 1** Flow diagram of study selection process. HPV, human papillomavirus; PCR, polymerase chain reaction
records. After removal of 572 duplicates, 1541 papers were reviewed for potential relevance. Of these, we excluded 1105 papers based on title, 338 papers based on abstract, and 67 papers after review of the full-text article (Figure 1). A total of 31 studies assessing HPV DNA prevalence in oral epithelial dysplasia were thus included in the meta-analysis. The percentage of HPV16 positivity was reported in 15 studies, HPV18 in 9, HPV6 in 5, and HPV11 in 4 studies. The percentage of p16\(^{INK4a}\) overexpression according to HPV status in oral epithelial dysplasia was examined in six studies.

### 3.2 Study characteristics

The included studies were published between 1991 and 2017 with six studies between 1991 and 2000, nine studies

---

### FIGURE 2

Study-specific and pooled prevalence of HPV in oral epithelial dysplasia cases, A, and controls, B. CI, confidence interval; HPV, human papillomavirus; n, number of HPV positive subjects; N, total number of subjects; * denotes studies with control subjects.
between 2001 and 2010, and 16 studies between 2011 and 2017 (Table S1). In eight studies, less than 10 cases were included, and only five studies included more than 50 cases. Most studies originated from Asia (n = 13) and Europe (n = 8), followed by North America (n = 6), South America (n = 3), and Africa (n = 1). The vast majority of studies extracted HPV DNA from biopsies of oral dysplastic lesions (n = 29), while the remaining two studies used exfoliated cells sampled directly from the dysplastic lesions. A third of the studies used a combination of consensus and type-specific PCR primers (n = 10), while seven studies used a combination of consensus primers, eight studies used a single consensus primer, and the remaining six studies used type-specific primers. In most studies (n = 28), the prevalence of HPV DNA was reported among samples with epithelial dysplasia, while three studies reported on a mixed group of dysplastic/CIS samples, and another five studies among CIS samples only. In 14 studies, a control group was included for comparison of HPV DNA prevalence among healthy oral mucosa. All studies assessing p16\(^{INK4a}\) expression (n = 6) used immunohistochemistry to detect overexpression of the p16\(^{INK4a}\) protein and applied various test-positivity thresholds (Table S2).

### 3.3 Overall HPV DNA prevalence

The overall pooled prevalence of HPV DNA in oral epithelial dysplasia based on the 31 studies (832 cases) was 27.2% (95% CI: 17.6-38.1) with substantial interstudy heterogeneity (I\(^2\) = 91%, \(P < .01\)) (Figure 2).

When we stratified on continent, we found the highest pooled HPV DNA prevalence in samples from South America (37 cases, 73.5%, 95% CI: 58.3-86.2), while a single study from Africa reported the lowest HPV DNA prevalence (31 cases, 0.0%, 95% CI: 0.0-3.1) (Table 1). Regarding tissue type, the pooled prevalence of HPV DNA was 25.0% (95% CI: 15.3-36.2) in biopsies (734 cases) and 61.0% (95% CI: 22.0-93.1) in exfoliated cells (98 cases). Stratification by type of PCR primer revealed a pooled prevalence of HPV DNA in studies using a combination of consensus and type-specific primers of 47.3% (95% CI: 33.2-61.6), while studies using a combination of consensus primers reported 17.9% (95% CI: 7.9-30.8), studies using a single consensus primer reported 21.6% (95% CI: 2.9-50.9) and finally, studies using only type-specific primers found 11.8% (95% CI: 0.0-44.7). When stratified according to the severity of dysplasia, lesions with epithelial dysplasia had a pooled HPV DNA prevalence of 24.9% (95% CI: 14.5-36.3), the mixed group of dysplastic/CIS lesions a prevalence of 29.8% (95% CI: 7.7-58.7), and lesions with CIS a prevalence of 55.1% (95% CI: 25.8-82.5).

In a sensitivity analysis restricted to 14 studies that also included a control group, the overall pooled HPV DNA prevalence in oral dysplasia was 32.6% (95% CI: 18.1-49.0), while the pooled HPV DNA prevalence among control subjects was 11.1% (95% CI: 3.5-22.2) (Table 1).

### 3.4 Percentage of type-specific HPV positivity

The pooled percent positivity of the specific HPV genotypes HPV16, 18, 6, and 11 was estimated among HPV DNA positive cases of oral dysplasia (Table 2 and Figure S1A-D). The most common genotype detected was HPV16 with a pooled percentage of 69.2% (95% CI: 36.4-93.6) among the HPV DNA positive cases. HPV18 accounted for 13.7% (95% CI: 0.0-45.9) of all HPV infections, while 12.2% (95% CI: 0.0-43.2) and 20.1% (95% CI: 0.8-55.3) of HPV DNA positive cases were positive for HPV6 and HPV11, respectively.

### 3.5 p16\(^{INK4a}\) overexpression in relation to HPV status

The pooled percentage of cases positive for p16\(^{INK4a}\) overexpression among HPV DNA positive oral epithelial dysplasia (51 cases) was 62.2% (95% CI: 17.4-96.6), whereas the pooled percentage of cases positive for p16\(^{INK4a}\) overexpression among HPV DNA negative samples (122 cases) was 17.8% (95% CI: 1.5-46.5) (Table 3).

### 4 DISCUSSION

This meta-analysis is to our knowledge the largest meta-analysis conducted on HPV DNA prevalence in oral epithelial dysplasia with inclusion of 832 cases from 31 studies. We found a pooled overall HPV DNA prevalence in oral dysplasia of 27%, and a pooled percentage of HPV16 positivity of 69% among HPV DNA positive cases. Furthermore, overexpression of p16\(^{INK4a}\) was present in 62% of HPV DNA positive and 18% of HPV DNA negative cases.

Similar results were reported in the meta-analysis by Jayaprakash et al\(^{12}\) that included 186 cases from 12 observational studies and found a prevalence of 25% HPV16/18 positive. In the meta-analysis by Syrjänen et al\(^{4}\) including four studies on oral epithelial dysplasia with type-specific HPV data, HPV16 was present in 42/83 cases. Miller and White\(^{19}\) reviewed 19 studies (286 cases)
assessing HPV in oral intraepithelial neoplasia and detected HPV in 18.5% of cases. In our study, we also tested the robustness of the results in a sensitivity analysis restricted to studies including a control group, which showed an HPV DNA prevalence among cases comparable with the main analysis. The prevalence

### TABLE 1
Pooled prevalence of HPV DNA in oral epithelial dysplasia according to continent, tissue type, primer type, severity of dysplasia, and within a subgroup of studies with control samples

<table>
<thead>
<tr>
<th></th>
<th>No. of studies</th>
<th>No. of cases</th>
<th>Prevalence % (95% CI)</th>
<th>$I^2$ (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>31</td>
<td>832</td>
<td>27 (18-38)</td>
<td>91</td>
</tr>
<tr>
<td>Continent</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Europe</td>
<td>8</td>
<td>128</td>
<td>37 (16-62)</td>
<td>86</td>
</tr>
<tr>
<td>North America</td>
<td>6</td>
<td>175</td>
<td>34 (18-52)</td>
<td>83</td>
</tr>
<tr>
<td>South America</td>
<td>3</td>
<td>37</td>
<td>74 (58-86)</td>
<td>0</td>
</tr>
<tr>
<td>Asia</td>
<td>13</td>
<td>461</td>
<td>15 (5-30)</td>
<td>93</td>
</tr>
<tr>
<td>Africa</td>
<td>1</td>
<td>31</td>
<td>0 (0-3)</td>
<td>—</td>
</tr>
<tr>
<td>Tissue type</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Biopsy</td>
<td>29</td>
<td>734</td>
<td>25 (15-36)</td>
<td>91</td>
</tr>
<tr>
<td>Exfoliated cells</td>
<td>2</td>
<td>98</td>
<td>61 (22-93)</td>
<td>86</td>
</tr>
<tr>
<td>Primer type</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Comb. of consensus and type-specific</td>
<td>10</td>
<td>322</td>
<td>47 (33-62)</td>
<td>83</td>
</tr>
<tr>
<td>Comb. of consensus</td>
<td>7</td>
<td>212</td>
<td>18 (8-31)</td>
<td>74</td>
</tr>
<tr>
<td>Consensus</td>
<td>8</td>
<td>152</td>
<td>22 (3-51)</td>
<td>93</td>
</tr>
<tr>
<td>Type-specific</td>
<td>6</td>
<td>146</td>
<td>12 (0-45)</td>
<td>95</td>
</tr>
<tr>
<td>Severity of dysplasia</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Epithelial dysplasia</td>
<td>28</td>
<td>675</td>
<td>25 (15-36)</td>
<td>90</td>
</tr>
<tr>
<td>Mixed dysplasia/CIS</td>
<td>3</td>
<td>108</td>
<td>30 (8-59)</td>
<td>90</td>
</tr>
<tr>
<td>CIS</td>
<td>5</td>
<td>49</td>
<td>55 (26-83)</td>
<td>76</td>
</tr>
<tr>
<td>Among studies with controls</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cases</td>
<td>14</td>
<td>388</td>
<td>33 (18-49)</td>
<td>90</td>
</tr>
<tr>
<td>Controls</td>
<td>14</td>
<td>445</td>
<td>11 (4-22)</td>
<td>91</td>
</tr>
</tbody>
</table>

Abbreviations: CI, confidence interval; CIS, carcinoma in situ; comb., combination; HPV, human papillomavirus.

### TABLE 2
Pooled prevalence of specific HPV genotypes in HPV DNA positive oral epithelial dysplasia

<table>
<thead>
<tr>
<th>HPV genotype</th>
<th>No. of studies</th>
<th>No. of HPV positive cases</th>
<th>Prevalence % (95% CI)</th>
<th>$I^2$ (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HPV16</td>
<td>15</td>
<td>149</td>
<td>69 (36-94)</td>
<td>94</td>
</tr>
<tr>
<td>HPV18</td>
<td>9</td>
<td>99</td>
<td>14 (0-46)</td>
<td>92</td>
</tr>
<tr>
<td>HPV6</td>
<td>5</td>
<td>43</td>
<td>12 (0-43)</td>
<td>77</td>
</tr>
<tr>
<td>HPV11</td>
<td>4</td>
<td>28</td>
<td>20 (1-55)</td>
<td>73</td>
</tr>
</tbody>
</table>

Abbreviations: CI, confidence interval; HPV, human papillomavirus.

### TABLE 3
Pooled prevalence of p16$^{INK4a}$ overexpression in HPV DNA positive and HPV DNA negative oral epithelial dysplasia

<table>
<thead>
<tr>
<th>HPV status</th>
<th>No. of studies</th>
<th>No. of cases</th>
<th>Prevalence % (95% CI)</th>
<th>$I^2$ (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HPV DNA positive cases</td>
<td>5</td>
<td>51</td>
<td>62 (17-97)</td>
<td>91</td>
</tr>
<tr>
<td>HPV DNA negative cases</td>
<td>6</td>
<td>122</td>
<td>18 (2-47)</td>
<td>91</td>
</tr>
</tbody>
</table>

Abbreviations: CI, confidence interval; HPV, human papillomavirus.
among the control subjects was considerably lower (11%) and equivalent to other studies reporting on oral HPV DNA prevalence in healthy subjects.\textsuperscript{4,20}

We examined the percentage of type-specific HPV positivity in a subset of studies that evaluated HPV genotypes and found that HPV16 appeared to be the predominant HPV type among HPV DNA positive oral epithelial dysplastic cases. A considerably lower proportion of HPV DNA positive samples were positive for HPV18 and the low-risk types HPV6 and 11. A similar pattern for type-specific HPV percent positivity has been found in oral cavity carcinoma, where more than 50% of HPV positive samples have been reported to be attributable to HPV16.\textsuperscript{2,21} However, our results of the type-specific HPV percent positivity should be interpreted with caution as HPV genotypes were only evaluated in half of the included studies and the exact genotypes investigated varied between studies. Furthermore, we were not able to evaluate single vs multiple HPV infections due to missing information and small numbers.

We found substantial variation in the HPV DNA prevalence across the included studies, and potential sources of heterogeneity were evaluated by stratification (ie, continent, tissue type, severity of dysplasia). When stratifying by geographical area, we observed considerable variation in HPV DNA prevalence across continents. The geographical variation in HPV DNA prevalence may be due to differences in the underlying HPV DNA prevalence of the study population, sexual habits, ethnicity, or prevalence of other etiologic factors for oral epithelial dysplasia, such as use of alcohol, tobacco products, and nutrition.

We found a higher pooled prevalence of HPV DNA among studies using exfoliated cells for HPV DNA detection compared to studies using biopsied tissue; however, only two studies used exfoliated cells. Previous studies assessing oral HPV DNA prevalence by exfoliated cells vs biopsies in subjects with normal, premalignant, and malignant tissue report equivocal results.\textsuperscript{4,22,23} The higher prevalence found among exfoliated cells samples in our study could be due to contamination of the scrape during sampling with HPV DNA from other parts of the oral cavity, especially the oropharynx via saliva.

Furthermore, we observed a tendency of increasing HPV DNA prevalence with increasing severity of dysplasia, ranging from 25% in epithelial dysplasia to 55% in CIS, albeit confidence intervals were wide and overlapping. Further evaluation by grade of dysplasia was not possible due to limited information provided by the individual studies. Studies assessing HPV DNA prevalence in other precancerous lesions (ie, vaginal and penile) have found varying association between HPV DNA and severity or grade of the lesions.\textsuperscript{24,25}

When examining potential sources of heterogeneity, substantial heterogeneity ($I^2 > 80\%$) remained within the different strata. This suggests that these factors only partly explain the interstudy heterogeneity and our results should be interpreted with this in mind. The extensive variance in HPV DNA prevalence across studies may also be due to the differences in study population sizes, case selection (archived material vs material from prospective recruitment of patients), sample storage, and PCR primer type. Most studies used the oral dysplasia diagnosis defined by the World Health Organization, which is based on a combination of distinct architectural and cytological features.\textsuperscript{26} However, it is broadly acknowledged that the intra- and interobserver reproducibility is low,\textsuperscript{27} and the selection of the oral epithelial dysplastic cases included in our meta-analysis may depend on the histopathological assessment conducted in each study. We excluded studies with only a selected sample of oral epithelial dysplasia cases (eg, immunocompromised patients). We did, however, include a study by Khanal et al\textsuperscript{28} in which the database-derived study population consisted of two-thirds randomly retrieved cases of high-grade oral dysplasia along with one-third of cases selected specifically based on HPV-associated phrases in the medical file. This may explain their finding of a relatively high HPV DNA prevalence of 58%, however, as the study only contributed with 5% of the total number of cases, it is unlikely that it will have influenced our results substantially. Finally, the majority of the studies did not report study population characteristics such as age, sex, tobacco and alcohol history, sexual habits, previous oral cancer or year of diagnosis separately for HPV positive and negative cases, therefore we were not able to explore these potentially confounding factors further.

In this meta-analysis, we focused on detection of HPV DNA by PCR. In oral cavity cancer, the mere presence of HPV DNA may not translate into transcriptional HPV activity but could represent a transient HPV infection. Despite an estimated overall HPV DNA prevalence of 20% in oral cavity carcinomas, only around 60% of HPV DNA positive oral cavity carcinomas show mRNA expression for HPV oncogenes E6 and E7.\textsuperscript{2,5} In contrast to oropharyngeal squamous cell carcinoma, where HPV positive and negative tumors represent distinct biological diseases both regarding genomic sequencing and clinical response to treatment, no such evidence exists within oral cavity squamous cell carcinoma. Moreover, our result of an HPV DNA prevalence of 11% among control subjects also underlines that asymptomatic HPV infection can be detected in the oral cavity of healthy subjects without dysplastic or malignant lesions, similarly as in the genital tract.
The discrepancy between the level of HPV DNA and mRNA prevalence could presumably also apply to oral epithelial dysplasia, and our finding of a 27% pooled HPV DNA prevalence may overestimate the actual proportion of HPV-infected oral dysplasia. However, as none of the included studies performed testing of HPV mRNA, this could not be examined.

Since this meta-analysis was based on cross-sectional studies, the result cannot be regarded as evidence of a causative role for HPV in oral epithelial dysplasia. Whether HPV is involved in the pathogenesis or malignant progression of oral epithelial carcinoma thus remains unsolved. Interestingly, a recent study was the first to demonstrate that 18.2% (6/33) of severe oral epithelial dysplasia cases were HPV16/18 positive by RNA in-situ hybridisation, of which 66% (4/6) cases showed immunohistochemical expression of HPV E4 and HPV E1, indicative of productive and biologically relevant HPV infection. Furthermore, a preclinical spontaneous HPV16 positive buccal tumor model has been developed using submucosal injection of oncogenic plasmids expressing HPV16 E6/E7 and mutant Ras. The mice developed squamous carcinomas at the injection sites and this development was preventable by injection with a therapeutic HPV DNA vaccine. Any potential causal association needs to be assessed in follow-up studies detecting HPV E6/E7 mRNA in oral epithelial dysplasia.

In addition, we aimed to provide estimates of carcinogenic activity of HPV in oral epithelial dysplasia by examining the overexpression of p16INK4a in relation to HPV status. The HPV E7 oncoprotein activates the p16INK4a tumor suppressor protein, hence overexpression of p16INK4a is considered a relatively accurate marker for HPV-infected oropharyngeal carcinoma. However, it may be unsuitable for use in other head and neck carcinoma subsites as high p16INK4a protein expression can be caused by mutation or amplification of CDKN2A, RB1 mutation, or NSD-1 hypermethylation. Furthermore, p16INK4a point mutations are detected in nearly 30% of head and neck carcinoma. Recently, it was shown that only 48.7% and 77.4% of HPV DNA positive oral cavity and oropharyngeal carcinomas, respectively, were also positive for E6*I mRNA and p16INK4a. Among the HPV16 E6*I mRNA positive oral cavity and oropharyngeal carcinomas, respectively, nearly 22% and 16% remained p16INK4a negative. This is in line with the results from the comprehensive genomic characterization of head and neck squamous cell carcinomas by the Cancer Genome Atlas Network. Within oral epithelial dysplasia, a distinct subset of high-grade lesions with characteristic histopathological features positive for both HPV16 and p16INK4a has been identified. To our knowledge, no previous meta-analysis has assessed p16INK4a overexpression in oral epithelial dysplasia in relation to HPV status. In the six studies included in our meta-analysis, the percentage of p16INK4a overexpression positivity ranged from 0% to 100% among HPV DNA positive and 0% to 89% among HPV DNA negative cases. With such a large degree of heterogeneity, cautious interpretation of the results is required. Nonetheless, our pooled results do suggest that a considerable proportion (62%) of HPV DNA positive cases in fact also overexpress the p16INK4a protein compared to only 18% of HPV DNA negative cases. Among the HPV DNA positive cases overexpressing p16INK4a, we were not able to evaluate the HPV genotypes due to limited cases and lacking information. In the two studies providing this information, HPV16 was detected in 5/5 and 5/6 HPV DNA and p16INK4a positive cases, respectively. Most of the studies assessing p16INK4a expression defined overexpression of p16INK4a as staining in >5-10% of cells, which is the lowest test-positivity threshold that has been used and may have resulted in a generally high pooled prevalence of cases positive for p16INK4a overexpression found in this meta-analysis. Currently, the consensus for grading a sample as p16INK4a positive is that 70% or more of the carcinoma cells show p16INK4a expression. However, in our case, the percentage of p16INK4a overexpression should be affected similarly in HPV DNA positive and negative cases within each study. Still, our results indicate a potential correlation between HPV DNA and p16INK4a in oral epithelial dysplasia and more research assessing this relationship is warranted.

The strengths of this meta-analysis included the thorough literature search, with a large number of included studies and thereby large number of cases. We provided both an overall analysis including studies regardless of study design (with and without control groups), and we performed a sensitivity analysis restricted to studies with a control group where it was also possible to compare with healthy individuals. Finally, to our knowledge, this is the first meta-analysis to examine pooled percent positivity of p16INK4a overexpression in oral epithelial dysplasia stratified by HPV status.

Some limitations to our study also exist. We observed large heterogeneity in study-specific HPV DNA prevalence, which could not be explained by continent, tissue type and severity of dysplasia. Unfortunately, we were not able to evaluate this further due to limited information on additional factors in the included studies. Restricting our meta-analysis to PCR-based studies may result in an overestimation of clinically relevant HPV DNA prevalence.

Publication bias could be a potential limitation, however as this is a meta-analysis synthesizing prevalence studies, and we include several studies reporting an HPV
DNA prevalence of 0%, this is not likely to influence our estimates considerably. Finally, inclusion of only English-language studies also constitutes a potential limitation.

In conclusion, this meta-analysis offers data to support that 27% of oral epithelial dysplasia harbor HPV DNA with HPV16 as the predominant genotype, and overexpression of p16INK4a in the majority of HPV DNA positive samples. Whether this HPV DNA represents a transient infection or contributes directly to the pathogenesis of oral cavity cancer is still unknown. Further research is warranted including HPV E6/E7 mRNA prevalence studies, genomic analyses of HPV positive and negative dysplastic lesions as well as follow-up studies evaluating outcomes of HPV positive vs negative oral dysplastic cases.

ORCID
Cecilie Dovey de la Cour https://orcid.org/0000-0002-4211-1134
Federica Belmonte https://orcid.org/0000-0002-2920-5005

REFERENCES


SUPPORTING INFORMATION
Additional supporting information may be found online in the Supporting Information section at the end of this article.