# Can immunohistochemical biomarkers distinguish epithelial dysplasia degrees in actinic cheilitis? A systematic review and meta-analysis

**Thalita Santana [1], Bruno Matuck [2], Jefferson R. Tenório [1], Mariana Minatel Braga [3]**

[1] PhD student at School of Dentistry, Department of Oral Pathology, University of São Paulo
[2] MSc at School of Dentistry, Department of Oral Pathology, University of São Paulo
[3] PhD and Professor at Department of Pediatric Dentistry, School of Dentistry, University of São Paulo, São Paulo, Brazil

*Correspondence:*
*School of Dentistry, Department of Oral Pathology*
*University of São Paulo, Av. Lineu Prestes, 2227*
*Cidade Universitária, São Paulo, SP 05508-000, Brazil*
*thalitasc@usp.br*

## Abstract

Background: Actinic cheilitis (AC) is a potentially malignant disorder of the lip, characterized by epithelial and connective tissue alterations caused by chronic exposure to ultraviolet radiation. In the past decades, diverse studies have been conducted in lip carcinogenesis and many biomarkers have been identified in lip lesions, yet there is no scientific evidence that determines its usefulness in the clinical setting or in histopathological routine. Therefore, we conducted the first systematic review in this field to summarize the results of published studies on immunohistochemical biomarkers in lip carcinogenesis, to evaluate if there is a marker than can distinguish the different histological grades of AC.

Material and Methods: Retrospective studies that investigated immunohistochemical biomarkers in AC defined on standardised histological assessment were gathered from five databases and evaluated. Each study was qualitatively evaluated using the Critical Appraisal Tools from SUMARI.

Results: The proliferation marker Ki-67 was the most studied biomarker and we observed, through meta-analysis, that it was differently expressed between AC and lip cancer, but not in AC subgroups. Most articles had a high risk of bias.

Conclusions: In summary, the literature lacks quality follow up studies in actinic cheilitis. Multi-centre cohort studies, with patients stratified by treatment type and the use of image analysis software, could be the solution to further address the issues of investigating potentially malignant lesions and help change clinical practice, in terms of individualizing patients' treatment and prognosis prediction.

*Key words: Lip carcinogenesis, actinic cheilitis, lip cancer, biomarkers.*

## Introduction

Lip squamous cell carcinoma (LSCC) represents 20-30% of all oral cavity tumors and it deserves a specific attention, especially in its pathogeny, that differs from oral squamous cell carcinoma (OSCC) (1). OSCC is related to chronic consumption of alcohol and tobacco, while LSCC is closely related to chronic exposure to the ultraviolet (UV) radiation of sun (1). The establishment of LSCC is preceded by clinical and histological alterations in the lip, which is known as actinic cheilitis (AC). AC is regarded as a potentially malignant lesion and it is characterized as a degenerative disorder, affecting mainly white males over 40-years-old, that usually work outdoors (2). Histologically, this lesion is characterized by cytological and architectural modifications, epithelial dysplasia and solar elastosis (basophilic degeneration of elastic fibres) (3).

To facilitate patient management, grading systems for oral epithelial dysplasia have been proposed. According to the WHO (4), epithelial dysplasia can be characterized as mild, moderate or intense, according to cytological and architectural alterations. However, this system cannot predict patient's prognosis and is regarded by pathologists as subjective (5). In 2006, Kujan et al. (6) proposed a binary graduation system for oral dysplasia, in order to minimize analysis subjectivity. This new system preconizes the division of the lesions in two subgroups, according to the risk of malignant transformation (low risk and high risk).

In the past years, researches have tried to elucidate the mechanisms underlying oral epithelial dysplasia. Thereby, many different immunohistochemical biomarkers have been investigated in oral carcinogenesis and a compilation of these results has been outputted (7); yet, to our best knowledge, there are no systematic reviews on biomarkers of lip carcinogenesis, and researchers and practioners are still not able to determine which AC cases will undergo malignant transformation. For this reason, we conducted a systematic review to examine if there is some immunohistochemical biomarker that could be related to the degree of epithelial dysplasia in AC.

## Material and Methods

This systematic review was reported according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses PRISMA Checklist (available at: http://www.prisma-statement.org). The review protocol was registered at the International Prospective Register of Systematic Reviews (PROSPERO) under number CRD 42017055294.

- Study design

We conducted a systematic review of human studies to summarize the results of published studies on immuno-histochemical biomarkers in lip carcinogenesis, in order to evaluate if there is a marker than can distinguish the different histological grades of actinic cheilitis (AC).

- Search strategy

We searched and identified articles of the following bibliographic databases: PubMed, Scopus, Web of Science, ScienceDirect and Scielo. The search included all articles published up to April 25th, 2017, with no time restrictions. Duplicated references were excluded by a reference manager software (Mendeley Desktop version 1.17.9).

The search strategy used for PubMed was: ((((((actinic cheilitis[Title/Abstract]) OR actinic cheilosis[Title/Abstract]) OR solar cheilosis[Title/Abstract]) OR solar cheilitis[Title/Abstract]) OR lip carcinogenesis[Title/Abstract]) OR lip photocarcinogenesis[Title/Abstract]) AND immunohistochem*[Title/Abstract]. For the other databases we conducted independent searches using the first block: (actinic cheilitis OR actinic cheilosis OR solar cheilosis OR solar cheilitis OR lip carcinogenesis OR lip photocarcinogenesis), in order to maximize the inclusion of relevant studies.

- Inclusion criteria

Prospective and retrospective studies that investigated immunohistochemical biomarkers in AC defined on standardised histological assessment as outlined by the WHO (4) and/or Kujan (6).

- Exclusion criteria

The following exclusion criteria were applied: (a) Scientific papers that did not report AC histological grading; (b) Lack of comparison between biomarkers among AC groups or between AC and control (normal lip mucosa or lip squamous cell carcinoma); (c) Studies that investigated immunohistochemical biomarkers in samples other than paraffinized material; (d) Reviews, single case reports, clinical trials, letters, personal opinions, book chapters, and conference abstracts.

- Study selection

The study selection was conducted by two authors (BM and TS), who independently reviewed the titles and abstracts of all the papers, and selected the studies that met the inclusion criteria. A kappa test was performed to verify agreement between authors and we obtained a reliable result of 0.87. Afterwards, both authors independently evaluated all full articles to determine if they reported the expression of immunohistochemical biomarkers in the subgroups of AC, based on histological grading (kappa score = 1). If there were any disagreements between the authors, they were resolved by mutual consensus. Final selection was always based on the full-text of the publication.

- Data collection

Two authors (BM and TS) collected the information from the included papers. The following information was gathered and presented in tables: study characteristics (author, year of publication, country); population (sample size, cases of AC, LSCC and normal lip con-

trols); type of histological grading performed; immuno-histochemical biomarkers that were analysed; expression of biomarkers in each subgroup (AC, LSCC and control); statistical tests performed; main conclusions. A partial grey literature search was performed using Google Scholar in order to investigate detailed results from PhD and Master's degree thesis and dissertations, and perform the meta-analysis.

- Meta-analysis

Due to high heterogeneity in immunopositive cells counting/scoring for the studied biomarkers and contrasting results presentation in the articles, we included only the protein Ki-67 for the meta-analysis. This was one of the most studied proteins and the only one that had a standardized analysis among the studies. To be included for meta-analysis, the articles (or their respective thesis/dissertations) had to report the mean and standard deviation of Ki-67 immunopositive cells in each of the following groups: control, mild dysplasia AC, moderate dysplasia AC, severe dysplasia AC, low grade LSCC, moderate/high grade LSCC. We used a Mixed-effects

Model to estimate the amount of residual heterogeneity (tau$^2$) and unaccounted variability (I$^2$) among groups. The analysis was performed with the R software, package metafor 1.9-8.

- Risk of bias in individual studies

Each selected study was qualitatively evaluated using the Critical Appraisal Tools from SUMARI (System for the Unified Management, Assessment and Review of Information), proposed by the Joana Briggs Institute (available at: http://joannabriggs.org/research/critical-appraisal-tools.html). Since almost all cases comprised retrospective studies, with samples chosen by convenience and lack of follow-up, we used the Critical Appraisal Tool for Case Series. This type of study is described as the kind in which "only patients with the outcome are sampled (either those who have an exposure or those who are selected without regard to exposure), which does not permit calculation of an absolute risk" (8). In our case, the exposure is the lesion actinic cheilitis. The evaluated items were scored "Yes", "No" or "Unclear" for each paper individually (Table 1).

**Table 1:** Risk of bias of selected studies according to the Critical Appraisal Tool for Case Reports (SUMARI).

| Criteria | Chrun et al.(13) | Barbosa et al.(14) | Garcia et al.(10) | Garcia et al.(11) | Gonçalves et al.(15) | Lopes et al.(16) | Ariotti et al.(17) | Araujo et al.(18) | Bianco et al.(12) | Lopes et al.(19) | Gomes et al.(20) | Oliveira et al.(21) | Alves et al.(22) | Salvadori et al.(23) | Sarmento et al.(24) | Araujo et al.(25) | Von Bubnoff et al.(26) | Souza et al.(27) | Freitas et al.(28) | Chou et al.(29) | Costa et al.(30) | Fontes et al.(31) | Xavier et al.(32) | Xavier et al.(33) | Freitas et al.(34) | Cury et al.(35) | Pontes et al.(36) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1. Were there clear criteria for inclusion in the case series? | Unclear | Unclear | Unclear | Yes | Unclear | Unclear | Unclear | Yes | Unclear | Unclear | Unclear | Unclear | Yes | Unclear | Yes | Unclear | Unclear | Unclear | Unclear | Unclear | Unclear | Unclear | Unclear | Yes | Unclear | Unclear | Unclear |
| 2. Was the condition measured in a standard, reliable way for all participants included in the case series? | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| 3. Were valid methods used for identification of the condition for all participants included in the case series? | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| 4. Did the case series have consecutive inclusion of participants? | No | No | No | No | No | No | No | No | No | No | No | No | No | No | No | No | No | No | No | No | No | No | No | Yes | No | No | No |
| 5. Did the case series have complete inclusion of participants? | No | No | No | No | No | No | No | No | No | No | No | No | No | No | No | No | No | No | No | No | No | No | No | No | No | No | No |
| 6. Was there clear reporting of the demographics of the participants in the study? | No | Yes | No | Yes | No | No | No | No | No | No | Yes | No | No | Yes | No | Yes | No | No | Yes | No | Yes | No | No | Yes | No | Yes | No |
| 7. Was there clear reporting of clinical information of the participants? | No | Yes | No | No | No | No | No | No | No | No | No | No | No | No | No | No | No | No | No | No | No | No | No | No | No | No | No |
| 8. Were the outcomes or follow up results of cases clearly reported? | No | No | No | Unclear | No | No | Yes | No | No | No | No | No | No | No | No | No | No | No | No | No | No | No | No | No | No | No | No |
| 9. Was there clear reporting of the presenting site(s)/clinic(s) demographic information? | No | No | No | No | No | No | No | No | No | No | No | No | No | No | No | No | No | No | No | No | No | No | No | No | No | No | No |
| 10. Was statistical analysis appropriate? | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | No | Yes | Yes | Yes | Yes | Yes |

Legend:    Yes ▮    No ▮    Unclear ▮

## Results

### - Study selection

1088 articles were identified across the five electronic databases. After removing the duplicates, 822 articles remained. A comprehensive evaluation of the titles and abstracts resulted in the exclusion of 746 articles, with the remaining 76 articles being allocated to full text in-depth review. This process led to the exclusion of 49 studies. Finally, 27 articles were retained for qualitative analysis and three articles were selected for meta-analysis. A flow chart detailing the process of identification, inclusion, and exclusion of the studies is shown in Fig. 1.

### - Study characteristics

All reviewed articles comprised retrospective studies (10-36) (Table 2). The studies presented a great geographic polarization; from the 27 analyzed articles, 25 where originated from Brazil, one from Germany and one from the USA. The included studies were published between 2003 and 2017. The number of AC cases in each study varied from 10 to 70, while LSCC cases went from 0 to 65 cases. Forty-one [41] different proteins were researched in these articles, and the relation between these proteins expressions among LSCCs and ACs was investigated in 20 papers. The mean number of studied ACs was 34.5 cases per article, while the mean of LSCC was 32 cases. Only five articles used the binary grading system proposed by Kujan *et al*. (Table 1). Among the studied biomarkers, the ones that were most investigated were DNA repair proteins, with 12 antibodies assessed. The inflammatory markers were the second most assessed group. Other groups of proteins were also analyzed, including apoptosis markers, metalloproteins, cell cycle markers, growth factors, neural and muscle markers (Fig. 2). This variety of biomarkers hampered any analyses between the articles.
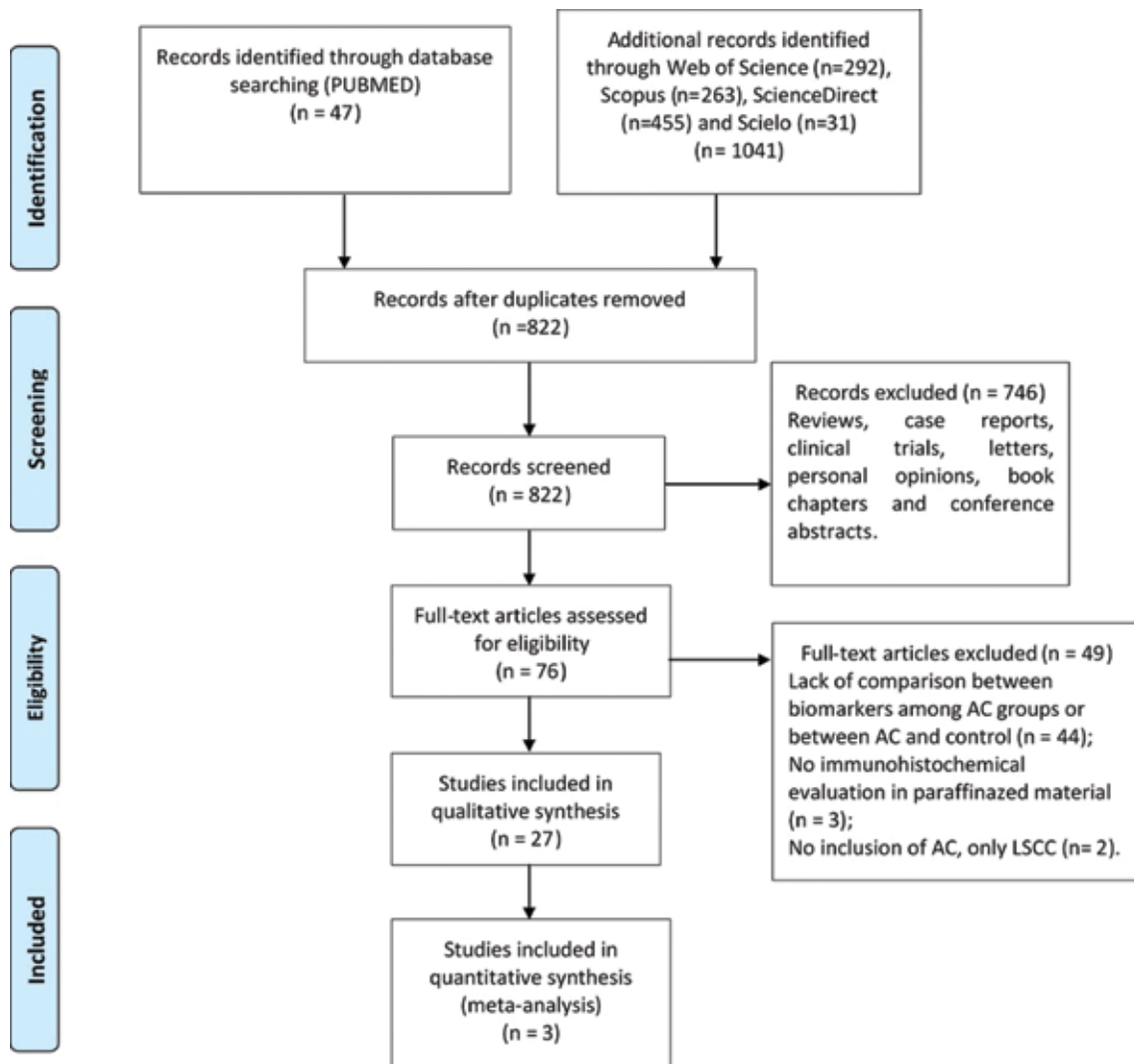


**Fig. 1:** Flow diagram for study selection.

**Table 2:** Details of selected studies.

| Bib-liographic Citation | Country | Grad-ing | Number of AC cases/ LC cases | Immunohistochemistry analysis | Antibody | Mild | Moderate | Severe | LSCC |
|---|---|---|---|---|---|---|---|---|---|
| **Chrun et al., 2017** | Brazil (South) | WHO and binary | 30/30 | Quantitative analysis – 500 cells counted in 5 fields at 400x | HDAC1 | + | + | + | + |
| | | | | | HDAC2 | +++ | +++ | +++ | + |
| | | | | | HAT1 | + | + | + | + |
| **Barbosa et al., 2016** | Brazil (North-east) | WHO | 40/ x | Quantitative in five fields – represented by positivity index | VEGF- c | + | + | + | X |
| | | | | | HIF1a | + | + | + | X |
| | | | | | D240 | + | + | + | X |
| **Garcia et al., 2016** | Brazil (South-east) | WHO | 45/ 15 | Quantitative in four ran-domized fields presented as means by epithelial layers | Cyclin d1 | + | ++ | +++ | ++++ |
| | | | | | Ki-67 | + | + | + | ++++ |
| | | | | | B-catenin | + | + | + | ++++ |
| **Garcia et al., 2016** | Brazil (South-east) | WHO | 45/20 | Cytokeratins were observed in epithelial layers in 10 fields and for Ki-67 a cell proliferation index was calculated (epithelial posi-tive cells x the total number of epithelial cells) | CK10 | + | + | + | ++++ |
| | | | | | CK13 | + | + | + | ++++ |
| | | | | | Ki-67 | + | + | + | ++++ |
| **Gonçalves et al., 2016** | Brazil (Midwest) | WHO and binary | 30/20 | Semi-quantitative - 0, no tumor cells stained; 1, <25% of cells stained; and 2,>25% of cells stained. Staining intensity was scored as: 0, no staining; 1, weak staining; 2, moder-ate staining; and 3, strong staining. | HLAG | + | + | + | ++++ |
| | | | | | HLAE | + | + | + | ++++ |
| | | | | | IL10 | + | + | + | ++++ |
| **Lopes et al., 2016** | Brazil (North-east) | WHO | 40/30 | Quantitative by mean - ran-domized fields until a num-ber of 1000 cells | p53 | + | + | + | + |
| | | | | | p21 | ++++ | ++++ | ++++ | + |
| | | | | | hMSH2 | There is no association between hMSH2 and AC degrees or LSCC | | | |
| **Ariotti et al., 2016** | Brazil (South) | WHO | 27/46 | Semi-quantitative score - percentage of positive cells in a 200× magnification field (0=0 % stained cells 1=1–50 % cells, 2=51–100 % of cells). CD31 was used to calculate microvascular density in hotspots. | VGRF 1* | ++++ | ++++ | ++++ | + |
| | | | | | VGRF2 | ++++ | ++++ | ++++ | |
| | | | | | CD31 | + | + | + | ++++ |
| **Araujo et al., 2010** | Brazil (North-east) | WHO | 35/x | Quantitative by mean – positive cells counted in 10 fields at 200x magnification | CD1a | + | + | + | x |
| | | | | | Mast cells | ++++ | ++++ | + | x |
| **Bianco et al., 2015** | Brazil (South) | WHO | 30/30 | Quantitative – positive cells were evaluated in 5 equi-distant fields (400x). | MMP1 | + | + | + | + |
| | | | | | MMP2 | ++++ | ++++ | ++++ | + |
| | | | | | MMP9 | ++++ | ++++ | ++++ | + |
| | | | | | Ki-67 | +[a] | +[a] | +[a] | +[a] |
| | | | | | a-SMA | +[a] | +[a] | +[a] | +[a] |
| **Lopes et al., 2016** | Brazil (North-east) | Binary | 65/x | Semi-quantitative score - (1) 1–30% of positive cellls (2) 31–60% (3) more than 60% | | High risk | | Low risk | |
| | | | | | Galectin 1 | + | | + | |
| | | | | | Galectin 3 | + | | +++ | |
| | | | | | Galectin 7 | + | | + | |
| | | | | | Galectin 9 | +++ | | + | |
| **Gomes et al., 2016** | Brazil (South-east) | Binary | 42/21 | Quantitative analysis by spots - seven fields quantified comprising an area of 1mm² | | High risk | | | LSCC |
| | | | | | CD1a | ++++ | | | + |
| | | | | | CD83 | + | | | + |

**Table 2 cont.:** Details of selected studies.

| Study | Country | | Sample | Analysis | Biomarker | | | | |
|---|---|---|---|---|---|---|---|---|---|
| **Oliveira et al., 2014** | Brazil (North-east) | WHO | 40/40 | Quantitative analysis by percentage of positive cells in ten consecutive fields of 400x | p63 | + | + | + | ++++ |
| | | | | | MDM2 | ++++ | + | + | + |
| | | | | | hMHL1 | + | + | + | + |
| **Alves et al., 2014** | Brazil (South-east) | WHO | 26/25 | Quantitative analysis -1,000 cells were counted at 400× and presented as percentage of positive cells | MDM2 | + | + | + | + |
| | | | | | SUMO1 | + | + | + | + |
| **Salvadori et al., 2014** | Brazil (South) | WHO | 29/53 | Ki-67 quantitative analyses, until 1,000 cells/ TGF-b1 and Elastin semi-quantitative analysis, presented as scores | Ki-67 | + | ++ | +++ | ++++ |
| | | | | | Elastin | + | + | + | + |
| | | | | | TGF-b1 | ++++ | +++ | ++ | + |
| **Sarmento et al., 2013** | Brazil (North-east) | WHO | 40/40 | Quantitative analysis until 1,000 cells (400x) – presented as percentages and scores (1 <50% reduced expression; 25%–75% normal Expression; 3 > 75% Over-expression) | MLH1 | +++ | ++ | ++ | + |
| | | | | | MsH2 | +++ | ++ | ++ | + |
| **Araújo et al., 2012** | Brazil (North-east) | WHO | 16/16 | Semi-quantitative analysis - 10 fields at 100x magnification, degree of elastosis was evaluated according to scores | Elastin | +% | +% | +% | x |
| **Von bub-noff et al., 2012** | Germany | WHO | 25/x | Quantitative analysis, - stratified as density of inflammatory infiltrate positive cells/mm$^2$ | IDO | + | ++ | +++ | x |
| | | | | | CD11 | NOT RELATED TO DYSPLASIA GROUPS | | | |
| | | | | | S100 | NOT RELATED TO DYSPLASIA GROUPS | | | |
| | | | | | CD68 | NOT RELATED TO DYSPLASIA GROUPS | | | |
| | | | | | CD1a | NOT RELATED TO DYSPLASIA GROUPS | | | |
| **Souza et al, 2010** | Southeast | WHO | 29/29 | Quantitative analysis with an ocular lattice – presented by mean | CD31 | + | + | + | + |
| **Chou et al., 2010** | U.S.A. | WHO | 20/5 arising from AC | Qualitative analysis in epithelium layers | FGFR3 | + | + | + | + |
| **Costa et al., 2009** | Brazil (Midwest) | WHO | 15/37 | Quantitative analysis in 10 representative and consecutive microscopic high-power fields (x400) | Triptase | + | + | + | +++ |
| | | | | | CD117 | + | + | + | + |
| | | | | | CD31 | + | + | + | + |
| | | | | | Ki-67 | + | + | + | + |
| **Fontes et al., 2009** | Brazil (South-east) | WHO | 36/ 18 | Semi-quantitative analysis – 0 negative; 1 up to 5% of positive cells; 2 5% - 50% of positive cells; 3 more than 50% of positive cells | Maspin | ++++ | +++ | ++ | + |
| ***Xavier et al., 2009*** | Brazil (South-east) | Binary | 21/51 | Semi-quantitative analysis - negative (0, not detectable); 1 (detectable but less than 50% of tumoral or atypical cells stained); 2 (labeling of more than 50% and less than 75% of tumoral or atypical cells); and 3 (widely and highly expressed in more than 75% of the tumoral or atypical cells) | B-catenin | Low risk | | High risk | LSCC |
| | | | | | | + | | + | + |
| | | | | | MMP3 | + | | + | + |
| | | | | | Wnt5a | +++ | | ++ | + |

**Table 2 cont.:** Details of selected studies.

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| **Xavier et al., 2009 (2)** | Brazil (Southeast) | WHO | 40/65 | Quantitative analysis – mean percentage of positive cells from 15 random areas at 400x | P63 | + | + | + | + |
| **Freitas et al., 2008** | Brazil (Northeast) | WHO | 58/x | Semi-quantitative analysis - percentage of positive cells determined from the percentage of total positive and negative cells derived from 10 random areas at ×400 magnification, then classified as scores | p53 | +* | +* | +* | X |
| | | | | | MDM2 | +* | +* | +* | X |
| **Cury et al., 2008** | Brazil (Southeast) | WHO | 25/x | Qualitative analysis in nucleus and cytoplasm and in epithelium layers | STAT3 | -- | -- | -- | -- |
| | | | | | STAT3P | -- | -- | -- | -- |
| **Pimentel et al., 2006** | Brazil (Southeast) | WHO | 70/31 associated to a proximal AC | Quantitative and qualitative analysis - percentage of p53-positive cells and intensity of stain in 10 consecutive 400x fields | p53 | + | + | + | + |
| **Pontes et al., 2005** | Brazil (Southeast) | WHO | 12/x | Quantitative analysis - percentage of positive cells in a 400x field | hMSH2 | + | + | + | + |

+ Indicates similar protein expression between groups
++ and +++ indicate higher protein expression in the group (statistically significant)
X No cases
* Statistically significant – correlated among groups (Spearman's correlation)
-- No quantitative analysis
[a] Proliferation index (Ki-67) and myofibroblasts (α-SMA) increased with the worsening of dysplasia group and in LSCC, but it was not statistically significant
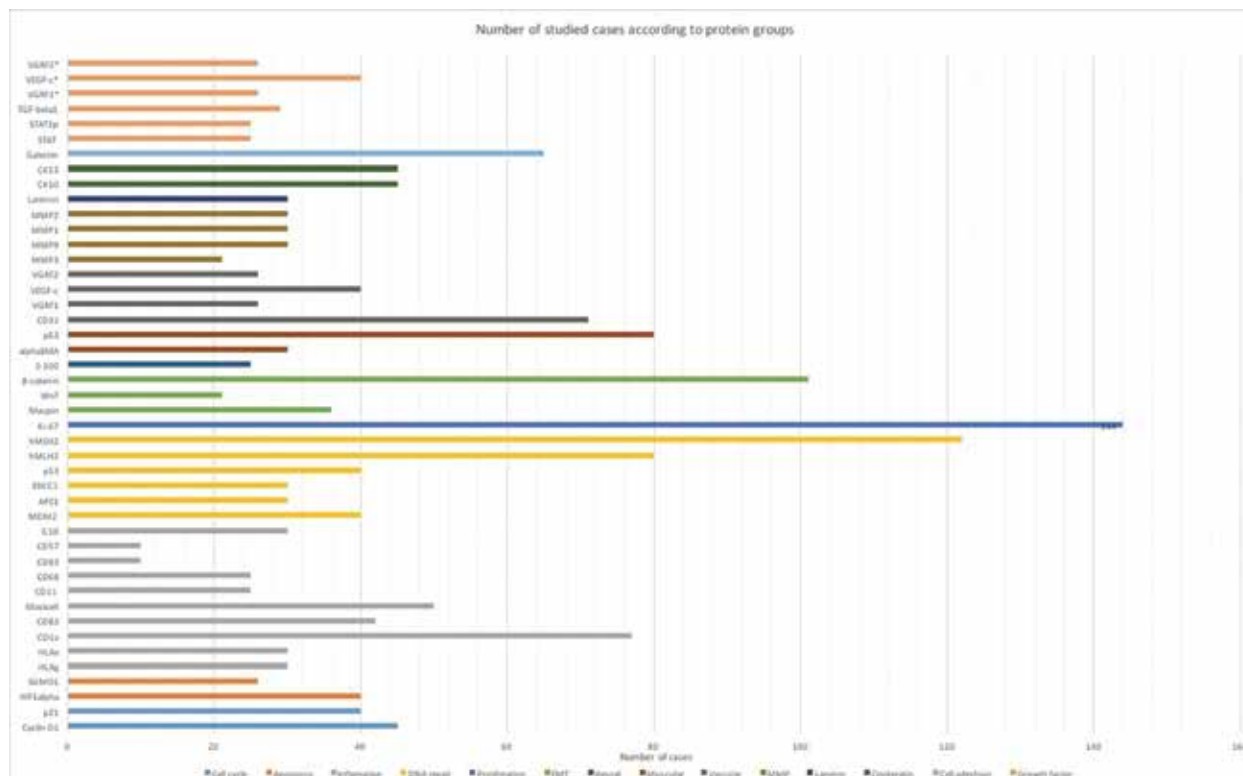


**Fig. 2:** Number of studied cases according to protein groups.

- Risk of bias

After analyzing all included studies with SUMARI critical appraisal tool, we observed that most articles (n=22) had unclear criteria for inclusion of the cases. We considered a clear inclusion criterion when the study reported when the cases were diagnosed, where they were collected and what were the sample inclusion/exclusion criteria (even if it was a convenience sample). Also, only three studies reported to have consecutive inclusion of cases. However, all studies reported a reliable measurement of AC degrees of dysplasia according to the grading systems proposed by WHO (4) and Kujan *et al* (6).

Thirteen articles reported the patients' demographics, while clinical information was available in only four of the reviewed studies. Information regarding the patient's outcome/follow-up was available in only one study (9) and partially available in another one (10).

Almost all studies reported to use clear statistical methods to compare the variables, however they were very diverse, since the researchers applied different methodologies for cell counting, with no clear cutoff of what was positive/negative or low/high. Also, maybe due to limited number of cases, authors tended to group AC cases for statistical analysis, and this grouping was very heterogeneous among studies.

- Meta-analysis

Three studies were selected for meta-analysis (11,12,23). We compared the mean expression of Ki-67 among the groups of AC, LSCC and control and observed a high heterogeneity among the studies (tau$^2$=241.02; I$^2$=95.91%). We observed that Ki-67 mean expression was similar in control groups and was higher in LSCC than in AC. However, it varied remarkably among AC subgroups. This information is summarized in a forest plot in Fig. 3.

## Discussion

One of the main purposes of investigating immunohistochemical biomarkers in lip lesions is observing if there are differences between protein expressions in different grades of AC or between AC and LSCC or control/normal lip. Usually, this is done not for diagnostic purposes, but with the expectation to develop future prognostic markers, which could possibly set apart cases that will undergo malignant transformation.

In the past two decades a significant number of studies have been conducted in lip carcinogenesis and many biomarkers have been identified in lip lesions, yet researchers are still not able to determine its usefulness in the clinical setting or in histopathological routine (37). The histological grading system proposed by WHO (4) for epithelial dysplasia in AC is extensively used by oral pathologists, anyhow, clinical experience has shown that even cases of mild oral dysplasia can develop into LSCC (38). In 2006, Kujan *et al.* (6) proposed a new binary grading system for AC, however a decade later few studies applied this system for histological assessment (5,39).

In this review, only few papers investigated the same biomarkers, which made it impractical to make comparisons between studies. The proliferation marker Ki-67 was the most studied biomarker, anyhow only three papers studying this protein met the inclusion criteria for meta-analysis. We observed that Ki-67 is differently expressed among AC and control and between AC and LSCC, however its expression was highly variable among AC groups.

Despite we have identified many studies in this review, almost all of them comprise case series, with cross-sectional analyses, lacking quality follow-up data for AC cases and are therefore unsuitable to in use in prognosis analysis. We recognize that this is one of the main issues of studying oral potentially malignant disorders, since
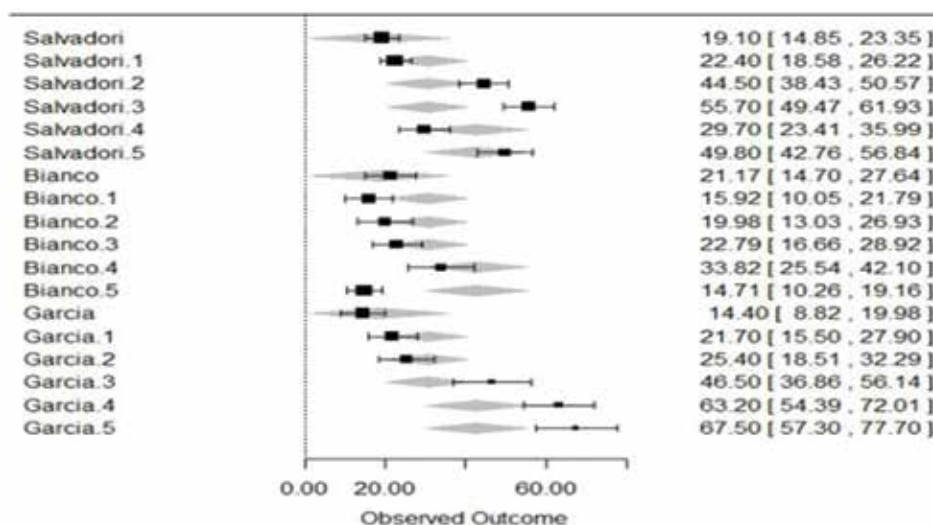


**Fig. 3:** Forest plot showing the differences in Ki-67 expression among control (no number), AC with mild dysplasia (.1), AC with moderate dysplasia (.2), AC with severe dysplasia (.3), low grade LSCC (.4) and moderate/high grade LSCC (.5).

they may or not undergo malignant transformation in an unknown period of time, and it is challenging to follow-up patients continuously.

Also, most reviewed studies failed to specify the inclusion criteria for studied cases and samples were chosen by convenience, which is a potential source of bias and it can reduce the level of evidence of the studies. Besides, many articles did not report important clinic and demographic data of the patients, nor did they report the presenting site's clinic/demographic information. Therefore, it is not possible to make comparisons between patient's characteristics and expression of immunohistochemical biomarkers. Furthermore, the literature is inconsistent regarding the evaluation of biomarkers positivity or low/high expression, since researchers uses different quantitative or semi-quantitative methodologies for cell evaluation, which makes it difficult to compare studies with the same biomarkers.

Notwithstanding, another possible source of bias is the statistical analysis performed for each study. Even though most authors statistically analysed their results, we observed a tendency in grouping AC subgroups (e.g. all cases of AC independently of dysplasia grading) and comparing it only to LSCC or control, to achieve statistically significant results. This could be due to small sample sizes, as it could possibly be related to selective reporting bias, which may happen as a result of the belief that scientific journals will not to accept papers reporting only "negative" results (not statistically significant). Additionally, nearly all evaluated studies are not replicable or reproducible, since important data are often not reported. According to Peng [2015] (40), there are two major components to a reproducible study: that the raw data from the experiment are available; and that the statistical code and documentation to reproduce the analysis are also available.

At last, we acknowledge another risk of bias within this review, since almost all included studies are from Brazilian research groups. This can be explained by the fact that in tropical countries, rural workers are chronically exposed to high levels of solar radiation throughout the year, which explains an AC prevalence of up to 28.4% in Brazilian populations (41). Considering that, a great number of studies in this field are conducted in this country. Also, we have thoroughly analysed all published studies in AC and although there are papers from USA, Chile, Australia, Greece, Spain and Germany, for example, only two of them met the inclusion criteria in this review. The other researches that investigated biomarkers in AC did not report comparisons between epithelial dysplasia groups and therefore were excluded.

As regards the meta-analysis, we also acknowledge its limitation. Since only three articles could be included for statistical analysis, we observed high heterogeneity among results, specially between AC groups, and this may not represent the reality of Ki-67 staining in AC or LSCC. One of the studies that investigated this protein (10) had to be excluded from meta-analysis since we believe the authors analysed the same sample used in their previous study (11), which was included.

- Implications for research and practice

In this review, we have identified 76 studies that investigated biomarkers in the field of lip carcinogenesis. Despite this significant number, well documented cohort studies are still limited. We still ought to understand the behaviour of AC and its progression to cancer, in order to apply it clinically. We emphasize the difficulty in accessing complete follow-up data and highlight the need for further clinical research in potentially malignant disorders. As suggested in a systematic review by Smith *et al.* (7), multi-centre cohort studies, with patients stratified by treatment type, could be the solution to further address the issues of investigating those lesions.

We recommend that in studies of biomarkers of lip carcinogenesis histological grading is performed for AC and LSCC, preferentially using more than one grading system. Also, comparisons with normal epithelium are indispensable. Thoroughly describing the methodology used for quantifying the antibodies is crucial for reproducibility of the study and, ideally, a unified methodology should be adopted. Maybe with the aid of an image software, to reduce examiners' observation variability, this could be achieved. Likewise, results should be described more carefully, with tables showing the results for each AC and LSCC subgroup, as well as control groups. Clinic and demographic information are also important to be described.

## Conclusions

We observed that the different studied proteins are similarly expressed in AC epithelial dysplasia grades, therefore are not useful in differentiating them. However, the potential use of some biomarkers to differentiate AC and LSCC has been demonstrated. We believe that soon some of them could become useful in identifying cancer risk in patients with actinic cheilitis. If we can develop reliable and reproducible follow-up studies, we will be able to change clinical practice in terms of individualizing patients' treatment and prognosis prediction. Clearly, further research is needed to exploit the many possibilities in lip carcinogenesis.

## References

1. de Visscher JG, van der Wall I. Etiology of cancer of the lip. A review. Int J Oral Maxillofac Surg. 1998;27:199-203.

2. Jadotte YT, Schwartz RA. Solar cheilosis: An ominous precursor Part I. Diagnostic insights. J Am Acad Dermatol. 2012;66:173-84.

3. Jadotte YT, Schwartz RA. Solar cheilosis: An ominous precursor Part II. Therapeutic perspectives. J Am Acad Dermatol. 2012;66:187-98.

4. Barnes L, Eveson JW, Reichart P, Sidransky D. World Health Organization Classification of Tumours. Pathology and Genetics of Head and Neck Tumours. Lyon: IARC Press;2005.

5. Câmara PR, Dutra SN, Takahama Junior A, Fontes KBFC, Azevedo RS. A comparative study using WHO and binary oral epithelial dysplasia grading systems in actinic cheilitis. Oral Dis. 2016;22:523-9.

6. Kujan O, Oliver RJ, Khattab A, Roberts SA, Thakker N, Sloan P. Evaluation of a new binary system of grading oral epithelial dysplasia for prediction of malignant transformation. Oral Oncol. 2006;42:987-93.

7. Smith J, Rattay T, Mcconkey C, Helliwell T, Mehanna H. Biomarkers in dysplasia of the oral cavity : A systematic review. Oral Oncol. 2009;45:647-53.

8. Dekkers O, Egger M, Altman D, Vandenbroucke J. Distinguishing case series from cohort studies. Ann Intern Med. 2012;156:37-40.

9. Ariotti C, Wagner VP, Salvadori G, Carrard VC, Martins MAT, Filho JJ da C, et al. VEGFR1 and VEGFR2 in lip carcinogenesis and its association with microvessel density. Tumor Biol. 2015;36:7285-92.

10. Garcia NG, Oliveira DT, Pereira Lauris JR, Custodio Domingues MA, Minicucci EM, Soares CT. Loss of cytokeratin 10 indicates malignant transformation in actinic cheilitis. Clin Oral Investig. 2016;20:745-52.

11. Garcia NG, Gonzalez-Moles MA, Ruiz-Avila I, Bravo M, Ramos-Garcia P, Minicucci EM, et al. Asymmetrical proliferative pattern loss linked to cyclin D1 overexpression during malignant transformation of the lip epithelium. J Eur Acad Dermatol Venereol. 2016;30:1315-20.

12. Bianco BC, Scotti FM, Vieira DSC, Biz MT, Castro RG, Modolo F. Immunohistochemical expression of matrix metalloproteinase-1, matrix metalloproteinase-2 and matrix metalloproteinase-9, myofibroblasts and Ki-67 in actinic cheilitis and lip squamous cell carcinoma. Int J Exp Pathol. 2015;96:311-8.

13. Chrun E da S, Modolo F, Vieira DSC, Borges ALSJ, Castro RG, Daniel FI. Immunoexpression of HDAC1, HDAC2 and HAT1 in actinic cheilitis and lip squamous cell carcinoma. Oral Dis. 2016;38:42-9.

14. Barbosa NG, Souza LB, Nonaka CFW, Silveira EJD. Evaluation of hypoxia, angiogenesis, and lymphangiogenesis in actinic cheilitis. Int J Dermatol. 2016;55:E573-8.

15. Gonçalves AS, Oliveira JP, Oliveira CFP, Silva TA, Mendonça EF, Wastowski IJ, et al. Relevance of HLA-G, HLA-E and IL-10 expression in lip carcinogenesis. Hum Immunol. 2016;77:785-90.

16. Lopes MLD de S, de Oliveira DHIP, Sarmento DJ de S, Queiroz LMG, Miguel MC da C, da Silveira ÉJD. Correlation between cell cycle proteins and hMSH2 in actinic cheilitis and lip cancer. Arch Dermatol Res. 2016;308:165-71.

17. Ariotti C, Wagner VP, Salvadori G, Carrard VC, Trevizani Martins MA, da Cunha Filho JJ, et al. VEGFR1 and VEGFR2 in lip carcinogenesis and its association with microvessel density. Tumor Biol. 2015;36:7285-92.

18. Araujo CP, Silva Gurgel CA, Goncalves Ramos EA, Freitas VS, Barbosa Junior A de A, Pedreira Ramalho LM, et al. Accumulation of CD1a-positive Langerhans cells and mast cells in actinic cheilitis. J Mol Histol. 2010;41:357-65.

19. Lopes MLD de S, Nonaka CFW, Queiroz LMG, de Souza LB, Miguel MC da C, da Silveira EJD. Pattern of galectins expression in actinic cheilitis with different risks of malignant transformation. J Oral Pathol Med. 2016;45:621-6.

20. Gomes JO, Carvalho M de V, Fonseca FP, Gondak RO, Lopes MA, Vargas PA. CD1a+and CD83+Langerhans cells are reduced in lower lip squamous cell carcinoma. J Oral Pathol Med. 2016;45:433-9.

21. De Oliveira DHIP, de Sousa Lopes MLD, de Santana Sarmento DJ, Queiroz LMG, da Costa Miguel MC, da Silveira ÉJD. Relationship between the epithelial expression of hMLH1, MDM2, and p63 and lower lip carcinogenesis. J Oral Pathol Med. 2014;43:357-63.

22. Alves MGO, Delgado A da M, Balducci I, Carvalho YR, Rodrigues Cavalcante AS, Almeida JD. Study of MDM2 and SUMO-1 expression in actinic cheilitis and lip cancer. Arch Dermatol Res. 2014;306:837-41.

23. Salvadori G, dos Santos JN, Trevizani Martins MA, Vasconcelos AC, Meurer L, Rados PV, et al. Ki-67, TGF-beta 1, and elastin content are significantly altered in lip carcinogenesis. Tumor Biol. 2014;35:7635-44.

24. Sarmento DJ de S, de Almeida WL, Miguel MC da C, Queiroz LMG, Godoy GP, Da Cruz MCFN, et al. Immunohistochemical analysis of mismatch proteins in carcinogenesis of the lower lip. Histopathology. 2013;63:371-7.

25. Araújo CP, De Aquino Xavier FC, Silva Gurgel CA, Gonçalves Ramos EA, Souza Freitas V, Schlaepfer-Sales CB, et al. Elastin accumulation in actinic cheilitis with different degrees of epithelial dysplasia. Int J Morphol. 2012;30:627-33.

26. von Bubnoff D, Zahn S, Wenzel J, Wilms H, Bieber T, Lüftl M. Indoleamine 2,3-dioxygenase expression in early keratocyte neoplasia of the lower lip correlates to the degree of cell atypia. Pathol Int. 2012;62:105-11.

27. Souza LR, Fonseca-Silva T, Pereira CS, Santos EP, Lima LC, Carvalho H a., et al. Immunohistochemical analysis of p53, APE1, hMSH2 and ERCC1 proteins in actinic cheilitis and lip squamous cell carcinoma. Histopathology. 2011;58:352-60.

28. Freitas VS, de Andrade Santos PP, Freitas R de A, Pinto LP, de Souza LB. Mast cells and matrix metalloproteinase 9 expression in actinic cheilitis and lip squamous cell carcinoma. Oral Surg Oral Med Oral Pathol Oral Radiol Endodont. 2011;112:342-8.

29. Chou A, Dekker N, Jordan RCK. Identification of novel fibroblast growth factor receptor 3 gene mutations in actinic cheilitis and squamous cell carcinoma of the lip. Oral Surg Oral Med Oral Pathol Oral Radiol Endodont. 2009;107:535-41.

30. Costa NL, Oton-Leite AF, Cheim-Júnior AP, Alencar Rde C, Bittar GO, Silva TA, et al. Density and migration of mast cells in lip squamous cell carcinoma and actinic cheilitis. Histol Histopathol. 2009;24:457-65.

31. Fontes A, de Sousa SM, dos Santos E, Martins MT. The severity of epithelial dysplasia is associated with loss of maspin expression in actinic cheilitis. J Cutan Pathol. 2009;36:1151-6.

32. Aquino Xavier FC, Rodini CO, Pedreira Ramalho LM, Mantesso A, Nunes FD. WNT-5A, but not matrix metalloproteinase 3 or beta-catenin protein, expression is related to early stages of lip carcinogenesis. J Oral Pathol Med. 2009;38:708-15.

33. Aquino Xavier FC, Takiya CM, Almeida Reis SR, Pedreira Ramalho LM. p63 Immunoexpression in lip carcinogenesis. J Mol Histol. 2009;40:131-7.

34. de Freitas Mda C, Ramalho LM, Xavier FC, Moreira AL, Reis SR. p53 and MDM2 protein expression in actinic cheilitis. J Appl Oral Sci. 2008;16:414-9.

35. Cury PR, Furuse C, de Araujo NS, de Araujo VC. Signal transducer and activator of transcription-3 expression and activation is dysregulated in actinic cheilitis. J Cutan Pathol. 2007;34:606-11.

36. Pontes HAR, Aguiar MCF de, Mesquita RA, Pontes FSC, Silveira Junior JB da. Imunoexpressão da proteína de reparo hMSH2 em queilite actínica e mucosa labial normal. Rev Bras Cancerol. 2005;51:23-30.

37. Brennan M, Migliorati CA, Lockhart PB, Wray D, Rcps FDS, Rcsed FDS, et al. Management of oral epithelial dysplasia : a review. Oral Surg Oral Med Oral Pathol Oral Radiol Endod. 2007;103:e1-e12.

38. Lumerman H, Freedman P, Kerpel S. Oral epithelial dysplasia and the development of invasive squamous cell carcinoma. Oral Surg Oral Med Oral Pathol Oral Radiol Endod. 1995;79:321-9.

39. Pilati SFM, Bianco BC, Vieira DSC, Modolo F. Histopathologic features in actinic cheilitis by the comparison of grading dysplasia systems. Oral Dis. 2017;23:219-24.

40. Peng, R. The reproducibility crisis in science: A statistical counterattack. Significance. 2015;12:30-2.

41. Ferreira AM, de Souza Lucena EE, de Oliveira TC, da Silveira EJD, de Oliveira PT, de Lima KC. Prevalence and factors associated with oral potentially malignant disorders in Brazil's rural workers. Oral Dis. 2016;22:536-42.

*Conflict of interest*
The authors declare that there are no conflicts of interest.