Assembling a consensus on actinic cheilitis: A Delphi study

Juan Seoane | Saman Warnakulasuriya | José Vicente Bagán | José Manuel Aguirre-Urizar | Pía López-Jornet | Gonzalo Hernández-Vallejo | Miguel Ángel González-Moles | Manuel Pereiro-Ferreiros | Javier Seoane-Romero | Pablo Varela-Centelles

Abstract

Aims: To discuss the terminology to define and classify actinic cheilitis (AC) and to build a consensus on the diagnostic and therapeutic approaches to AC.

Methods: Two-round Delphi study using a questionnaire including 34 closed sentences (9 on terminology and taxonomy, 5 on potential for malignant transformation, 12 on diagnostic aspects, 8 on treatment) and 8 open questions. Experts' agreement was rated using a Likert scale (1–7).

Results: A consensus was reached on 24 out of 34 statements (73.5%) and on 5 out of 8 close-ended questions. The response rate was identical in both rounds (attrition of 0%). AC is the term with the highest agreement (median of 7 (strongly agree; IQR: 6–7)) and the lowest dispersion (VC = 21.33). ‘Potentially malignant disorder’ was the preferred classification group for AC (median of 7) and 85.6% of participants showing some level of agreement (CV < 50). Experts (66.75%) consider AC a clinical term (median: 7; IQR: 4–7) and believe definitive diagnosis can be made clinically (median: 6; IQR: 5–7), particularly by inspection and palpation (median: 5; IQR: 4–6). Histopathological confirmation is mandatory for the management of AC (median: 5; IQR: 2.5–7), even for homogeneous lesions (median: 5; IQR: 3.5–6). Consensus was reached on all treatment statements (VC < 50).

Conclusions: AC is a potentially malignant disorder with a significant lack of agreement on diagnostic criteria, procedures, biopsy indications and the importance of techniques to assist in biopsy. A consensus was reached on nomenclature and management of this disorder.

KEYWORDS
actinic cheilitis, consensus, Delphi study, potentially malignant oral disorder
1 | INTRODUCTION

The term actinic cheilitis (AC) in scientific literature describes changes induced by chronic exposure to sunlight (ultraviolet A wavelength 320–400 nm and largely ultraviolet B wavelength 290–320 nm) mostly affecting the lower lip.\(^1\) However, AC is considered a misnomer by some authors who use alternative terms, like solar cheilitis,\(^2\) solar cheilosis,\(^3\) solar keratosis\(^4\) or actinic keratosis of the lip.\(^5\) This terminological controversy is based on the understanding of the pathobiology of the condition as primarily inflammatory\(^1\) or non-inflammatory\(^2\) and on the appellation (actinic: ‘a ray’ vs. solar) of the aetiological radiation.\(^2\)\(^5\)

There is also a lack of consensus on the nature of AC as either a ‘precancerous or potentially malignant’\(^6\)\(^7\) or a primarily neoplastic (intra-epithelial neoplasia) condition,\(^8\)\(^9\)\(^10\) which may explain the divergences in case definition, diagnostic criteria and management. Additionally, existing literature shows a disagreement on whether AC should be used as a clinical term where biopsy is not always recommended\(^11\) or used a preliminary clinical diagnosis for which histopathological examination is mandatory to reach a definitive diagnosis.\(^3\) Even the requirement of epithelial dysplasia of the lip epithelium to diagnose AC is a matter of debate.\(^12\)\(^13\)

These uncertainties may well explain the wide differences in reported AC prevalence when estimated in terms of histopathological (2.08%; 95%CI: 0.94–3.67)\(^11\) or clinical (31.3%; 95%CI: 30.5–35.7)\(^14\) criteria\(^14\) and highlight the need for a consensus on an unequivocal definition of AC.

A range of medical and surgical procedures (5% topical 5-FU, 5% topical imiquimod, 3% diclofenac gel, chemical peels, dermabrasion, photodynamic therapy, cryotherapy, electrosurgery, Er:YAG laser and vermilionectomy (cold blade and CO\(_2\) laser), together with Mohs micrographic surgery) have been used for treating AC.\(^3\)\(^15\)\(^18\) However, and despite the superiority of surgery compared with medical therapies (92.8% vs. 65.9% showing remission with lower recurrence rates (8.4% vs. 19.2%)\(^16\)\(^16\)), there is no consensus on the optimal strategy for treatment of AC\(^16\)\(^18\) because both approaches have benefits and drawbacks.\(^19\)

Thus, the current comprehension regarding AC clearly requires an effort to minimize disparities and reach agreements on terminological and taxonomical aspects, as well as in its diagnosis and treatment. In this vein, an expert consensus (Delphi method) approach can support clinical decision-making, particularly when evidence about a given topic is contradictory.\(^20\) The Delphi method is aimed at structuring a consensus and it is based upon the anonymity of the experts among themselves and on the iteration with controlled feedback (experts are consulted more than once).\(^21\)\(^26\) This technique is a procedure for group communication meant to reach a convergence of opinion on a specific ‘real-world’ issue.\(^26\)

The objectives of this investigation were i) to discuss and agree upon the most appropriate terminology to define and classify AC and ii) to build a consensus about the diagnostic and therapeutic approaches to AC.

2 | MATERIALS AND METHODS

A two-round Delphi was designed according to the guidance on conducting and reporting Delphi studies.\(^21\) A steering committee was established, and 21 experts (9 dermatologists, 9 specialists in Oral Medicine, and 3 oral pathologists) were recruited from a consortium of 11 academics and 10 clinical specialists across Europe and Latin America. Experts were selected through a bibliographical search (contributing at least one publication on this topic) and from their clinical experience (>15 years).

Our previous systematic review\(^15\) permitted the identification of terminological and taxonomical inconsistencies, along with discordances in diagnostic and therapeutic criteria for AC. As a result, a questionnaire was built including 34 closed sentences (9 on terminology and taxonomy, 5 on potential for malignant transformation, 12 dealing with diagnostic aspects, as well as 8 statements on AC treatment) together with 8 open questions. The instrument was electronically circulated to the panel.

The first round (R1: March 2020) used a Likert scale (1 to 7; 1: strongly disagree – 7: strongly agree) to assess the level of concordance with the closed sentences presented to the experts.\(^26\) Participants were also asked to cast their degree of confidence on their responses using another Likert scale (1: very uncertain – 7: very certain). Once the obtained information was analysed and synthesized, and considering the feedback from the panel, the open questions in the instrument were changed into close-ended ones and the second round (R2: June 2020) was initiated. In addition to the second questionnaire, the experts received information about the overall response of the group in R1. In case a participant disagreed in one or more of his/her responses with the panel (agree/strongly agree or disagree/strongly disagree), and to favours consensus, the person was asked to reconsider the score allocated to his/her R1 responses. This permitted the expert to bring his/her responses closer to the group’s opinion or to remain with his/her previous response and to justify his/her score.

For purposes of this study, a consensus for close-ended questions was defined as ≥75% agreement by the experts. With reference to the sentences, a <50 variation coefficient (minor variability among experts) was considered an indicator for agreement. The criterion for dropping items at R1 was pre-established considering a minimum acceptable level of confidence in expressing an opinion (≥ 5).

This study was approved by the University of Santiago de Compostela Research Ethics Committee (#USC-17/2020). A declaration of interests was completed by all participants.

2.1 | Statistical analysis

A quantitative analysis of the results of both R1 and R2 was undertaken using the SPSS v.24 statistical package (SPSS Inc). The descriptive study used the median as a central trend indicator and the interquartile range as a measure of spread. The variability of the
experts’ scores (dispersion) for a given item was calculated using the coefficient of variation (standard deviation/mean). The level of agreement (percentage) was also described for close-ended questions. The opinions of the experts were also considered for non-consensual questions.

3 | RESULTS

Twenty five experts were invited, and 21 agreed to participate in the study, resulting a recruitment rate of 84%.

The Delphi technique permitted a consensus on 24 out 34 statements (73.5%) and on 5 out of 8 (62.5%) close-ended questions. The response rate was identical in both rounds, with an attrition of 0%, and just one statement was dropped because of a low level of confidence among the participants in their responses. Results are displayed in Tables 1 to 4.

3.1 | Nomenclature and taxonomy

Actinic cheilitis (AC) is the term with the highest level of consensus among the experts, the lowest dispersion of scores (VC = 21.33) and a median of 7 (strongly agree) (IQR: 6–7). No differences (p > 0.05) were observed between dermatologists and oral medicine specialists in terms of degree of agreement. Conversely, ‘solar cheilosis’ gathered the strongest disagreement, with 85.7% of experts showing some level of discrepancy and a median of 1 (strongly disagree) (IQR: 1–3). Actinic cheilosis and solar cheilitis have also been discarded as inappropriate terms (CV >50).

Regarding taxonomy, ‘potentially malignant disorder’ was the most favoured by the group for classifying this entity, with a median of 7 and 85.6% of participants showing some level of agreement (CV < 50). Other options, such as ‘precancerous lip lesion’, ‘precursor lesion’ or ‘preinvasive malignant lesion’ have demonstrated lack of agreement (CV > 50) (Table 1). However, 23.8% of experts considered AC a primarily malignant lesion, and no significant differences were identified between specialists regarding the degree of agreement about whether AC should be considered a preinvasive malignant lesion (p > 0.05). Furthermore, the panel exhibited their agreement with the statement that histopathological findings are not sufficient to predict a malignant transformation (median 6; IQR: 4–6) and also with the absence of methodologically sound prospective studies reporting on malignant transformations of AC (median: 7; IQR: 5–7) (Table 1).

| TABLE 1 | Level of agreement (1–3: Disagreement; 5–7: Agreement) about nomenclature, taxonomy and potential for malignant transformation of actinic cheilitis in both rounds |
|---|---|---|---|---|---|---|---|---|
| | Round I | | | Round II | | | |
| | Score 1–3 (%) | Score 5–7 (%) | Median (IQR) | VC | Score 1–3 (%) | Score 5–7 (%) | Median (IQR) | VC |
| Nomenclature | | | | | | | | |
| Actinic cheilitis | (–) | 90.4 | 7 (6–7) | 18.78 | 95.2 | 7 (6–7) | 21.33 |
| Actinic cheilosis | 57.1 | (–) | 3 (1–5) | 68.96 | 66.6 | (–) | 2 (1–4) | 72.85 |
| Solar cheilitis | 52.4 | (–) | 3 (1–4) | 56.42 | 57.1 | (–) | 3 (1.5–4) | 50.80 |
| Solar cheilosis | 66.7 | (–) | 2 (1–4) | 73.15 | 85.7 | (–) | 1 (1–3) | 63.00 |
| Taxonomy | | | | | | | | |
| Preinvasive malignant lesion | 81.0 | (–) | 2 (1–3) | 61.94 | 71.4 | (–) | 2 (1–4.5) | 67.3 |
| Precancerous lip lesion | 57.1 | (–) | 4 (1.5–6.5) | 61.53 | 52.4 | (–) | 4 (1.5–6) | 58.00 |
| Potentially malignant disorder | (–) | 71.5 | 6 (3–7) | 40.07 | (–) | 85.6 | 7 (5.5–7) | 34.31 |
| Potentially malignant lesion | (–) | 70.0 | 6 (3–7) | 45.45 | (–) | 66.7 | 6 (3–7) | 45.51 |
| Precursor lesion | 52.4 | (–) | 2 (1–4.5) | 67.14 | 90.5 | (–) | 3 (1–3) | 56.22 |
| AC potential for malignant transformation | | | | | | | | |
| AC, good clinicopathological correspondence | 60.0 | (–) | 4 (3–6) | 36.97 | 57.2 | (–) | 4 (4–6) | 28.88 |
| AC, has a moderate rate (3%) of malignant transformation | (–) | 47.7 | 4 (3–7) | 45.07 | (–) | 47.6 | 4 (3.5–6.5) | 41.84 |
| AC, there are no prospective studies reporting malignant transformation | (–) | 75.0 | 6 (5–7) | 29.38 | (–) | 80.9 | 7 (5–7) | 27.25 |
| Pathological findings are not sufficient to predict malignant transformation | (–) | 61.9 | 5 (4–6) | 34.08 | (–) | 66.7 | 6 (4–6) | 32.72 |
| Lip cancer can develop on non-dysplastic lip epithelia | (–) | 80.9 | 6 (5–6.5) | 26.38 | (–) | 90.5 | 6 (5–7) | 21.87 |

Abbreviations: IQR, interquartile rage; VC, variation coefficient.
From a clinical standpoint, the presence of erosions or ulcerations, white patches and heterogeneous lesions was considered the best clinical predictors for malignant transformation in AC (76.9%). In addition, the presence of crusts, hyperkeratotic lesions, paleness and blurring of vermilion border can also indicate clinical suspicion for malignant transformation (80.95%) (Table 1).

### 3.2 Diagnosis of actinic cheilitis

Most experts (66.75%) consider AC a clinical term, not a histopathological description (median: 7; IQR: 4–7), and believe a definitive diagnosis can be made solely by clinical criteria (median: 5; IQR: 4–6), particularly by inspection and palpation (median: 5; IQR: 4–6). However, they find histopathological confirmation mandatory for the management of AC (median: 5; IQR: 2.5–7), even for homogeneous lesions (median: 5; IQR: 3.5–6).

An incisional biopsy of AC should be performed when a clinical diagnosis is established and when clinical changes or signs of suspicion for malignant transformation are detected during follow-up (95.23%). In this vein, an adequate histopathological diagnosis may be critical for selecting therapeutic approaches (median: 6; IQR: 3.5–7). However, the maximum diagnostic certainty for AC is achieved by surgical excision (vermilionectomy) after incisional biopsy (ie by histopathological assessment of the whole specimen) (median: 7; IQR: 6.5–7).

Moreover, the experts consider solar elastosis as a pathognomonic microscopic finding in AC (median: 7; IQR: 5–7), but no consensus was reached about the requirement of epithelial dysplasia for a histopathological diagnosis of AC (median: 3; IQR: 1–4; VC: 72.75; Table 2), with dermatologists consistently assigning significantly higher scores in favour of this diagnostic criterion as a requirement ($p = 0.019$).

### 3.3 Therapeutic approaches

Consensus was reached on all eight statements related to the treatment of AC (VC < 50). According to the panel, treatment should be started once the diagnosis is established (including preventive measures) (80.95%), and thereafter, follow-up should include clinical visits every 3 or 6 months, depending on the clinical presentation and on the severity of the epithelial dysplasia (85.71%), Tables 3, 4.
In mild AC cases, once clinical diagnosis is established, treatment could start without awaiting histopathological confirmation (median: 5; IQR: 2.5–7). Dermatologists elicited a higher level of agreement with this approach than did experts in Oral Medicine and Pathology (p = 0.014). Particularly, non-dysplastic AC lesions could be treated by non-surgical approaches (drugs and/or photodynamic therapy) (median: 6; IQR: 5–6), which offer good cosmetic outcomes after treatment and minor adverse events (median: 6; IQR: 5–6).
IQR: 4–6), but higher recurrence could be expected (median: 6; IQR: 4.5–6.5).

Surgical approaches are better than non-surgical ones, and laser vaporization techniques could be used for multicentric or diffuse AC lesions with mild epithelial dysplasia (median: 6; IQR: 4–7). Besides, AC lesions with moderate/severe epithelial dysplasia (median: 6; IQR: 5–7), and particularly those with diffuse presentations and with severe epithelial dysplasia are best treated by vermilionectomy techniques to prevent the development of a carcinoma (median: 6; IQR: 6–7), Table 3.

4 | DISCUSSION

4.1 | Strengths and limitations of the study

The Delphi methodology was used to address consensus with the objective of contributing to a common effort targeted towards agreement. A standardized Delphi procedure was used on a non-homogeneous consensus group to minimize the impact of individual opinions. Considering the composition of the panel influences the rating of Delphi studies, we have used strict criteria for the recruitment of the 21 experts (6 of them are university full professors). In addition, this methodology permits the identification of any different opinions between specialties.

In an attempt to evaluate the agreement level of the answers, a seven-point scale was adopted to be more discriminative, together with the selection of quality indicators for the study (number of rounds, reproducible criteria for the selection of experts and criteria for dropping items). Furthermore, the degree of confidence with each of the answers was incorporated as a subrogated indicator for the quality of the data obtained. However, reaching a consensus does not indicate the results of the consensus being factual, and it does not substitute for established clinical guidelines based on strong scientific evidence. Unfortunately, existing evidence on diagnosis and treatment of AC is scarce and discordant and a guideline development by professional societies may be very useful supporting future clinical practice.

4.2 | Taxonomy and potential for malignant transformation

Although ‘solar cheilitis’ has been recommended as an appropriate term for labelling this entity (inflammation is not the starting event, and ‘solar’ is more accurate than ‘actinic’, that includes other types of radiation as aetiological agent), ‘actinic cheilitis’ is the most frequently term used in the literature and the best valued by the panel of experts.

Classically, the terms ‘potentially malignant’, ‘premalignant’ and ‘precancerous’ have been used interchangeably, although ‘potentially malignant disorder’ has been the term adopted by the WHO Collaborating Centre in 2007. This concept combines both precancerous lesions and conditions separately grouped in the earlier WHO definitions. Potentially, malignant is more suited as not all ACs will eventually develop into lip cancer. This nomenclature has reached the highest level of consensus in the current study, although some authors still consider AC as micro-invasive, superficial, incipient malignant epithelial neoplasia.

The chances for a patient with actinic changes and lip carcinoma to develop a second primary lip cancer are 2.5 times higher than for a patient without AC. In addition, 95% of squamous cell carcinomas of the lip occur on a background of AC. Conversely, and considering that more than two decades of chronic sun exposure is needed for AC to turn into lip cancer and with the limited long-term, prospective, observational studies on non-treated AC, it is difficult to know the actual rate of malignant transformation of this entity. It is not surprising that the moderate rate of malignant transformation, 3%, reported in a recent systematic review supported by a single study has not reached a consensus of experts.

AC may present a wide range of clinical features which do not unveil the severity of the lesion. Thus, the presence of ulceration, severe inflammation, blurring of the vermilion, nodularity, atrophy, friability and chronic bleeding is indicative of epithelial dysplasia or presence of malignization. Besides, the panel has agreed the presence of hyperkeratotic lesions, paleness and heterogeneous lesions as additional clinical predictors for malignant transformation. Although histopathological changes are not sufficient for predicting the risk for malignant transformation of AC because malignant transformation can also occur in non-dysplastic epithelia, epithelial dysplasia becomes relevant with AC progression. In addition, different biomarkers (including overexpression of p53) have proved able to identify early events in lip carcinogenesis but the precise contribution of biomarkers to clinical practice is very limited.

4.3 | Diagnosis of actinic cheilitis

Considering AC being a clinical term and obtaining a biopsy would be particularly recommended from clinically suspicious sites, clinicians should avoid treating AC relying on clinical findings alone. Considering the divergence between the clinical appearance of the lesion and its histopathological characteristics, combined with the difficulties for a single biopsy to represent the whole lesion, the use of multiple incisional biopsies or other procedures to enhance diagnostic accuracy would be desirable. Autofluorescence imaging and staining with toluidine blue combined with clinical examination may well be one of the adjunctive chair-side procedures for detecting high-risk zones and selecting a biopsy site. However, the panel of experts considered this option as of minor importance and the item was removed after 2nd round because of the low level of confidence expressed in the responses (median:4; IQR: 2–7).

Besides, pre-treatment diagnosis has usually been based upon the information obtained from incisional biopsies although the nature of AC as a non-homogeneous, multifocal lesion is a well-known
limitation for this approach. This can result in the underdiagnosis of dysplastic lesions and masking of non-contiguous foci of squamous cell carcinoma, even in diffuse and poorly demarcated lesions.9,38 

The chances for this to occur seem to be supported by the frequently reported carcinomas in vermilionectomy surgical specimens (excisional biopsies) from patients with clinical diagnosis of AC40,41 even in those cases with a previous incisional biopsy confirming no invasion.9

4.4 | Therapeutic approaches

Different reports acknowledge a lack of consensus15-18 and insufficient evidence for recommending a particular treatment option.18 This circumstance reflects the absence of both controlled clinical trials and methodological standardization of AC studies.15-18 This heterogeneity is favoured by the unclear clinico-histopathological definition of AC. Several studies have defined their AC cases before treatment relying only on clinical criteria,42-44 and even on restrictive clinical criteria: 'refractory cases', 'only severe diffuse damages cases' or 'non-ulcerated AC'.45 Conversely, certain reports defined AC according to Kaugar's criteria, which requires the presence of epithelial dysplasia.46-49

Notwithstanding, five different systematic reviews have proved the superiority of surgical treatments over of non-surgical options in terms of histopathological control,15-18,50 particularly partial surgery and laser therapy (CO2, Er-Yag, Thulium) alone or in combination with photodynamic therapy elicited the best clinical response,18 followed by 5% imiquimod—topical chemotherapy drug—(once a day, 3 to 5 days per week for 4 to 6 weeks) and photodynamic therapy.17,18 However, topical and photodynamic therapy is associated with a higher recurrence.50 In any case, other non-surgical measures such as regular follow-up after treatment and intense photoprotection should also be implemented especially for high-risk patients.15,16,41

4.5 | Implications for clinical practice and research

The choice of therapy should be based upon the clinical and histopathological features of the lesion. Thus, mild- or non-dysplastic lesions can benefit from non-surgical approaches15,17 and surgical treatments should be prioritized for those cases with severe epithelial dysplasia, particularly vermilionectomy with serial analysis of the vermilion which could be the best option for diffuse and severe dysplastic lesions.

New strategies should be implemented to reach a consensus for the non-consensual statements identified in the current study, especially about the definition of a case and on the significance of epithelial dysplasia as a diagnostic criterion. More randomized clinical trials incorporating pre-treatment histopathological information and considering among their outcomes variables such as the remission rate, clinical response, changes to dysplasia grade, adverse events, cosmetic outcomes, healing time and recurrence rate, as well as follow-up, or malignization rate after treatment are needed to ease evidence-based decision-making. In addition, multicentric, longitudinal (prospective cohort design) studies are required to identify biomarkers able to predict malignant transformation in AC.35

5 | CONCLUSIONS

AC is a potentially malignant disorder with a significant lack of agreement on diagnostic criteria and procedures, biopsy indications and on the importance of ancillary techniques to assist in biopsy. However, and despite the insufficient evidence, a consensus was reached on the non-consensual statements identified in the current study, especially about the definition and the therapeutic management of this disorder, which can be useful to the clinicians in their decision-making processes.

CONFLICT OF INTEREST

The authors have no conflicts of interest to declare.

AUTHOR CONTRIBUTION

Juan Manuel Seoane: Conceptualization; Data curation; Formal analysis; Investigation; Methodology; Project administration; Resources; Supervision; Writing-original draft; Writing-review & editing. Saman Warnakulasuriya: Conceptualization; Formal analysis; Investigation; Methodology; Validation; Writing-original draft; Writing-review & editing. Jose V Bagan: Conceptualization; Formal analysis; Methodology; Validation; Writing-original draft; Writing-review & editing. JM Aguirre: Conceptualization; Formal analysis; Investigation; Methodology; Validation; Writing-review & editing. Pia Lopez-Jornet: Conceptualization; Formal analysis; Investigation; Methodology; Validation; Writing-review & editing. Gonzalo Hernández: Conceptualization; Formal analysis; Investigation; Methodology; Validation; Writing-review & editing. MA González-Moles: Conceptualization; Formal analysis; Investigation; Methodology; Validation; Writing-review & editing. Manuel Pereiro-Ferreiros: Conceptualization; Investigation; Methodology; Project administration; Resources; Supervision; Writing-review & editing. Javier Seoane Romero: Conceptualization; Data curation; Investigation; Methodology; Project administration; Resources; Supervision; Writing-original draft; Writing-review & editing. Pablo Varela-Centelles: Conceptualization; Data curation; Investigation; Methodology; Supervision; Writing-original draft; Writing-review & editing.

PEER REVIEW

The peer review history for this article is available at https://pubns.com/publon/10.1111/jop.13200.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available in the supplementary material of this article.

ORCID

Saman Warnakulasuriya https://orcid.org/0000-0003-2103-0746
José Vicente Bagan https://orcid.org/0000-0002-0295-2916
REFERENCES


