Title: Expert opinion on characteristics of vulval LS: initial identification of important clinical features through an international electronic-Delphi consensus study

Word count: Abstract: 349; Body of manuscript: 2,907

Table count: 3

Figure count: -

Authors:
R.C. Simpson[1] (Corresponding Author)
M. Birchall[1]
J. Daniels[2]
S. Gran[1]
G. Kirtschig[3]
K.S. Thomas[1]

[1] Centre of Evidence Based Dermatology, University of Nottingham, Nottingham, UK
[2] Clinical Trials Unit, University of Nottingham, Nottingham, UK
[3]Medical Health Centre Frauenfeld, Frauenfeld, Switzerland

Corresponding author: Rosalind C. Simpson, Rosalind.simpson@nottingham.ac.uk

Funding: This project is funded by the NIHR Advanced Fellowship (NIHR301434). The views expressed are those of the authors and not necessarily those of the NIHR or the Department of Health and Social Care.

Conflicts of interest: There are no conflicts of interest to declare.

Data availability: The data underlying this article will be shared on reasonable request to the corresponding author.

Ethics statement: This project was approved by the University of Nottingham's Faculty of Medicine and the Health Sciences Research Ethics Committee (FMHS 39-0722).

Acknowledgements: We acknowledge the participants who took time to complete the e-Delphi exercise, Miss Rheanne Leatherland at the Centre of Evidence Based Dermatology, University of Nottingham for support in proof-reading the manuscript, and the NIHR for their funding of this work.

What's already known about this topic?

- Vulval lichen sclerosus is a common inflammatory condition that is often misdiagnosed due to lack of awareness of presenting clinical features.
- Delay in diagnosis and treatment can lead to ongoing inflammation and irreversible damage, such as scarring and loss of normal vulval anatomy.
- There are no published diagnostic criteria for vulval lichen sclerosus.

What does this study add?

- Using electronic-Delphi methodology, we have identified a set of 17 diagnostic features (five 'critical/very important and 12 'important but not critical') which are relevant in the **clinical diagnosis** of vulval lichen sclerosus.
- The five most important **clinical diagnostic features** agreed by this group of experts and patients were: whiteness, itch, changes in anatomy, burying of the clitoral area, and improvement in response to topical steroids.

What are the clinical implications of this work?

- Checking validity and subsequent implementation of these diagnostic features will support the diagnosis of vulval lichen sclerosus in different **clinical settings**, particularly non-expert environments.
- Timely diagnosis should lead to earlier, effective treatment, which may minimise the often-irreversible effects typically observed in lichen sclerosus disease progression.

Abstract

Background: Lichen sclerosus (LS) is a chronic inflammatory condition mainly affecting genital skin. It causes distressing symptoms that impact daily quality of life as well as progressive anatomical changes and potential risk of cancer. Vulval LS is often misdiagnosed and treatment delayed. If non-experts are supported to identify clinical features of LS, they may diagnose the condition sooner and commence timely treatment or refer to specialist services for assessment and confirmatory biopsy.

Objectives: To reach international expert consensus on clinical diagnostic features for vulval LS.

Methods: Between March and May 2023, a four-stage electronic-Delphi consensus exercise was conducted. In the first three rounds, participants were asked to rate the importance of a list of clinical features. Responses from round one were summarised and presented in rounds two and three, along with additional features suggested by participants. Any items which met the definition of consensus as 'very important/critical', or 'not important' were not voted upon again in subsequent rounds. In round four, items which met agreement for 'important but not critical' were ranked in order of their importance. Consensus was defined a priori, and all rounds were conducted anonymously.

Results: A total of 47 participants from 14 countries completed round one with 42 (89%) retained by round 3 when consensus was determined. Round four was completed by 36 (77%) participants. Participants completing all four rounds predominantly included healthcare professionals (78%, n=28) and patient support group representatives (19%, n=7). In round one, 21 diagnostic features were voted upon. Participants suggested an additional 10 features which were subsequently added to the round two survey. After three rounds, consensus was achieved for five diagnostic features: whiteness, itch, changes in anatomy, burying of the clitoral area, and improvement in response to topical steroids. There were also 12 features voted as 'important but not critical' and participants subsequently ranked them in the 4th round.

Conclusions: Experts have agreed on five critical diagnostic features for vulval LS in adults and an additional 12 features that may also be important. Future research should assess these clinical features for diagnostic validity through a multicentre diagnostic test accuracy study.

Plain Language Summary

Establishing diagnostic criteria for vulval lichen sclerosus

Vulval lichen sclerosus (LS) typically causes itching, splitting, thinning and whitening of the vulva (the skin around the outside of the vagina). Other symptoms may occur including soreness, tightening and thickening of the skin. At least 1 in 100 of women are affected. It impacts on daily living, making activities like going to the toilet and sexual intercourse difficult. LS also increases the risk of vulval cancer.
LS can be difficult to diagnose, causing delays in treatment. This is problematic as symptoms can worsen, leading to irreversible vulval scarring and increasing the risk of vulval cancer.

We aimed to address this issue through asking international experts and patients to agree on a set of clinical features (symptoms and visual signs) to be included in a 'checklist' for diagnosing LS. We used a recognised research method called the electronic-Delphi technique where participants were asked a series of questions across four rounds.

A total of 47 participants from 14 countries contributed to the first survey, with 36 of these completing all four rounds. They were asked to score which LS features they felt were the most important for diagnosis.

This study agreed five features to be 'critical' in diagnosing LS (scarring of the area surrounding the clitoris, changes in genital structure, itch, improvement in response to topical steroids, and whiteness). A further 12 features were considered important, but not critical. These diagnostic features will go on to be tested in a clinical study for their accuracy. After this, a freely available 'checklist' will be created for helping health professionals and patients to diagnose LS more easily.

Introduction

Lichen sclerosus (LS), as defined by the International Classification of Diseases-11 (ICD-11) is an inflammatory disorder of unknown aetiology affecting the skin of the vulva and perianal area. Typically, it affects women in the fifth and sixth decades of life though it can occur at any age and is not uncommon in prepubertal girls.[1] Men may also be affected, although less commonly than women. LS is estimated to affect 0.1–0.3% of new patients in a general hospital patient population and 1.7% of patients referred to general gynaecological practice. However, the exact prevalence and incidence is not known. It could affect up to 3% of postmenopausal women.[2-5] The scale of the problem is likely to be underestimated due to reluctance in presentation and the fact that a proportion of women with LS are asymptomatic.[2,6,7]

Symptomatic LS can cause significant distress, including intense itch, pain, and skin splitting. This can negatively impact daily functioning, psychosocial well-being, and sexual activity.[8,9] LS symptoms may be commonly misdiagnosed as candidiasis or genitourinary symptoms of menopause (GSM). Often women are not examined on first presentation of these symptoms. When they are examined, visible findings may be mistaken for other vulval conditions such as candidiasis or eczema. Delay in diagnosis and treatment can lead to ongoing inflammation and irreversible damage, such as scarring and loss of normal vulval anatomy. There is also a 3-5% risk of developing vulval squamous cell carcinoma in some circumstances.[10-12] Furthermore, untreated LS may contribute to self-harm or suicidal thoughts, highlighting the importance of early detection and management.[8,13] For those who are diagnosed early, effective treatment may reduce the risk of scarring and malignant transformation[14] as well as improve overall quality of life.

Despite its significant impact, genital disease remains an under-researched area and clinical practice is often limited by lack of high-quality evidence.[15]

The British Association of Dermatologists state that LS is a clinical diagnosis and confirmatory biopsy is not always necessary when the typical clinical features are present.[16] Clinical history and examination are therefore the first step in diagnosing LS. Healthcare professionals to whom these patients may present need to be aware of LS and have the skills to identify clinical features. However, there are no published diagnostic criteria to support healthcare professionals in making their diagnosis and this has been identified as a key research priority in the field of LS. [17] It is important to note that histological examination is recommended if there are atypical features or diagnostic uncertainty, and it is essential if there is any suspicion of neoplastic change. Furthermore, biopsy is a test that is usually carried out in a specialist setting under the guidance of an expert rather than in the community. [16]

Via an international multi-perspective electronic-Delphi consensus exercise,[18,19] this study's objective was to undertake initial identification of important LS clinical features in women, as agreed by experts including patients.

The results will inform future research to assess the validity of these clinical diagnostic features through a multicentre diagnostic test accuracy study.

Methods

Study Type

This study employed a four-stage international electronic Delphi (e-Delphi) exercise that was conducted between March and May 2023. The Delphi method is a widely used technique in clinical and health services research, as it allows for the establishment of consensus among a panel of expert participants through the scoring and revision of a series of structured statements.[18,19] The electronic format of this study allowed for ease of participation across multiple countries and was moderated by a central coordinator (R.C.S.). This technique has been used for developing diagnostic criteria in the past.[20]

Participants

In the recruitment phase, an invitation email was sent to a diverse group of expert stakeholders, including clinicians, researchers, industry representatives, and patient representatives. The invitees were identified as potential participants due to their specialist interest in this field or their lived experience with LS. Participants included members from the International Society for the Study of Vulvovaginal Disease (ISSVD), the British Society for the Study of Vulval Disease (BSSVD), and the leads of patient support groups. The ISSVD and BSSVD were selected as suitable organisations from which to invite participants as their members are actively engaged in the management of patients with vulval disease and are therefore considered expert in the field. Patient support group leads were invited as they have a broad knowledge and wide understanding of the lived experience of LS through liaising with their members.

No inconvenience allowance was offered, and ethical approval was granted by the University of Nottingham's Faculty of Medicine & Health Sciences Research Ethics Committee (FMHS 39-0722). The study protocol was made available in the public domain prior to starting the study.[21]

Study Procedures

To inform the consensus exercise, candidate diagnostic features to include were harvested from Cochrane systematic reviews,[22] qualitative work with health professionals and patients,[23,24] and think-aloud discussions with patient representatives. The lead researcher extracted a long list of potential diagnostic criteria from these sources.

The e-Delphi exercise was conducted in English. Participants' details were kept anonymous except for the coordinator who required these for administrative purposes. Participants completed the questionnaires using the online Welphi tool[25] and were given approximately 2-weeks to submit their responses for each round, with reminders sent to non-responders.[26] Informed consent was obtained at the beginning of the first round. Those who did not participate in a round of the survey were not invited to participate in the subsequent rounds.

In the first round, participants rated the importance of a long list of 21 diagnostic features on a nine-point Likert scale. They were asked *'In your opinion, how important are the following features to make a diagnosis of vulval lichen sclerosus?'*. A score of one indicated the feature was *"Not important*", a score of 4-6 indicated they were "*Important but not critical*," and a score of 7-9 indicated they were "*Critical*" features for making a clinical diagnosis of LS. Participants were also

given the option to select "*Unable to Score*" if they felt unable to comment based on their experience. Participants were not limited in how many they could rate as 'critical' nor were they asked to rank the diagnostic criteria at this stage.

During voting, participants were able to make comments about the individual features and they were able to suggest additional diagnostic features they believed were missing from the initial list. The survey instrument was amended following these comments from round one and additional items added for round two.

During the second round, participants were provided with feedback on the group's average response for each diagnostic feature and were given the opportunity to revise their rating for each item. Items meeting consensus as very important/critical, or not important, were removed prior to the subsequent round. During round three, participants were asked to vote again on the remaining features and were provided with the group's average response from round two.

A fourth round asked participants to rank order the features that did not meet consensus for "*critical"* but where 70% or more had voted *"important but not critical"* or *"critical" (*scores of 4-9).

Definition of Consensus

The definition of consensus was determined *a priori* and included two criteria: 1) at least 70% of participants score the outcome as 7, 8 or 9; and 2) 15% or less of participants score it as 1, 2, or 3. If 70% or more of participants vote an item as 'not important' (a score of 1,2 or 3) an item could also be dropped.

Results

A total of 53 invitees expressed interest to participate and were subsequently invited via email to complete the first round of the study. Of these, 47 completed round one. Subsequent retention rates were 44 (97%) in round two, 42 (89%) in round three, and 36 (77%) in round four. Participants represented three distinct stakeholder groups and 14 different nationalities. The United Kingdom and Unites States comprised half 30% (n=14) and 21% (n=21) of respondents respectively. Although the group had varying levels of experience with LS patients, over half reported having more than 10 years of experience. Of the healthcare professionals who completed the first three rounds, 36% were dermatologists and 32% were gynaecologists, suggesting that a large proportion of participants had significant knowledge and experience in this field. Participant demographics are outlined in Table 1.

In the first round, 21 potential diagnostic features were presented, and an additional 10 features were added in subsequent rounds based on participant suggestions. Moreover, three participants provided seven comments about the diagnostic features, and these were considered before amending the features present in round 2. Table 2 shows the complete list of the features and descriptive statistics.

Following round two, three features (i.e., changes in the anatomy of the genital area, itch, and whiteness) met the predetermined criteria for consensus and so were removed from further voting. Bladder pain was also removed from further voting as it met the pre-defined threshold for exclusion.

During round three, two additional features (i.e., burying of the clitoral area and response to topical corticosteroids) received enough votes to meet the criteria for consensus, bringing the total number of critical features for which a consensus had been established to five. These features are listed in Table 2.

After round three, 12 additional features were identified as "important" or "critical" by over 70% of participants (scoring them between 4 and 9). Due to the large number of features meeting consensus during round 3, participants were asked to complete a final round where they ranked these items in order of their importance for making a clinical diagnosis of LS. This was to elucidate whether some had greater weighting within this category of importance, compared to others. The three most highly ranked features included fissuring, crinkly skin, and bruising/bleeding under the skin. Table 3 presents the results of this final round.

Discussion

This four-round e-Delphi consensus exercise has enabled expert consensus on important clinical features to diagnose vulval LS. Whilst the results of this work provide expert opinion, an algorithm or score has not yet been developed. Therefore, the validity and utility of the features in the clinical setting have not been proven and they need to be tested for diagnostic accuracy and consistency. Differential diagnosis of vulval itch with visible skin changes include dermatitis (atopic, seborrhoeic, allergic contact or irritant), lichen planus, vulvovaginal candidiasis and vitiligo. It is likely that a combination of the identified diagnostic features will be needed, rather than an individual feature being diagnostic on its own, to differentiate from these other conditions. This may be particularly important in the case of LS overlapping with other disorders, such as lichen planus.

Strengths of this study are that we engaged experienced individuals as the majority of the participants had over six years' experience in managing or supporting others with LS. We are therefore confident that the group have the necessary skills to contribute to the derived diagnostic dataset. There was international reach with representation of 14 countries over five continents.

The study did not have representation from genitourinary medicine or histopathology. These are considerations when interpreting the results as these specialities may have provided different diagnostic viewpoints. However, as we sought to identify important *clinical* diagnostic features for LS, histopathological findings were *not* being judged at this time. LS is almost always diagnosable on biopsy, however, incorrect biopsy site selection can lead to false negative results and national guidelines indicate that LS can be diagnosed clinically without need for biopsy, except for certain situations.[16] There was some attrition of participants during the process, particularly between rounds 3 and 4. However, decision for consensus was made after round 3 where 89% remained in participation.

Clinical scoring systems for LS have previously been proposed in the literature. In a clinical study with 24 LS patients and 49 patients with other vulval diseases, a six-feature physician (erosions, hyperkeratosis, fissures, agglutination, stenosis, and atrophy) and four-feature (pruritus, burning, soreness, dyspareunia) patient administered score was developed with the aim of assessing symptoms and evaluating treatment response over time. Indeed, the system has been applied in several clinical studies since its publication. Items for inclusion in the scale were selected by two physicians from the same centre and applied to a cohort of 73 patients coming through their

centre.[27] The tool has not been widely adopted for use and it does not appear in guidelines as a method for diagnosing LS. [16,28] A reason may be that it uses technical language and it is not suitable for non-experts or those not familiar with vulval disease.

Scoring systems for identifying and monitoring vulval architectural changes have also been published,[29,30] but these are not intended for diagnostic use.

Shenis and Selk carried out an e-Delphi consensus exercise for item generation to create an LS severity scale (rather than for a diagnostic tool). Patients were not included in the generation of their e-Delphi agreed list.[23] In a subsequent study with six experts, they were unable to agree on any signs, architectural changes, or an overall global impression to assess LS disease severity based on analysis of vulval photographs.[31] They concluded that standardized descriptions regarding what constitutes mild, moderate, and severe signs and anatomical changes are required before further scale development can occur.

A recent online survey study collected self-reported information on disease presentation in pre-menopausal women with biopsy proven LS.[32] Of 503 respondents, average delay in diagnosis was 4 years. Most common symptoms and those which had greatest impact on the individual were dyspareunia (68%) and tearing with intercourse or vaginal penetration (63%). Other common symptoms or skin changes noticed by patients were hypopigmentation, vulval fissures and labial resorption. The findings from this survey study are representative of pain, fissuring, whiteness and anatomical change, all of which are in our list of critical or important diagnostic features for LS.

The SWIFT (soreness, whitening, incontinence, fissures, and thickening of the clitoral hood) tool has been created for LS diagnosis in children.[33] Following a pilot study which identified 5 predictors for LS in premenarchal girls, a larger study involving 105 patients retrospective case records demonstrated a 97% accuracy in predicting LS. The model is yet to be validated in a prospective, larger multicentre cohort. The clinical features represented in the tool include those in our list of critical or important features, except for incontinence. SWIFT refers to paediatric rather than adult vulval LS which is likely to account for this difference. However, incontinence is indeed reported to be associated with genital LS and it may represent a causative feature or triggering factor for the disease.[34]

The results of our e-Delphi exercise in terms of clinical LS findings are comparable with the aforementioned published studies. However, some additional points were raised which deserve highlighting. A positive response to topical corticosteroids is true of any inflammatory skin condition and therefore it is not *specific* to LS. Around 70% of patients with LS will have disease remission after a 3-month course of superpotent topical corticosteroid.[35,36] Complete non-response to topical corticosteroid would be a way to rule out LS, but there are many variables, such as non-compliance, inadequate potency or duration of treatment, that need to be considered. If the condition does not respond positively, an alternative diagnosis, or non-adherence, are strong possibilities. Dermatologists frequently encounter other inflammatory skin conditions in the vulval clinic, but gynaecologists may not. Patients are unlikely to have knowledge of other vulval conditions. This may be the reason that this item was considered a potential diagnostic feature for LS, despite being non-specific.

Participants commented that the differences between early and late findings in LS should be acknowledged as some changes occur over time in stages. For example, in early LS, whitening with hyperkeratosis but without anatomical changes can occur. In later, non-active LS, scarring with whiteness but not hyperkeratosis might be seen. Explanation of the type of whiteness seen in LS was commented upon: the whiteness seen in hyperkeratosis (seen in other vulval conditions) is different from the yellowish-white seen in sclerosis (specific to LS). Furthermore, in younger women, dyspareunia is the most common symptom[32] and in this group skin changes may be subtle, for example fissuring occurring only after sex.

Finally, the differentiation between features that are *specific* for LS versus those that are *common* in LS but seen in other conditions needs to be clarified. Many vulval conditions itch and burn, cause dyspareunia and may respond to topical corticosteroids. At this initial stage, we do not define the number of critical criteria required to rule in LS, nor any algorithm or score. We will investigate how well the diagnostic criteria voted as 'critical' and 'important but not critical' function in practice to diagnose LS through a multicentre diagnostic test accuracy study using expert clinical diagnosis as the reference standard.[37]

Our goal is unique as we aim to create a tool which can be used by non-experts and patients to support the earlier recognition and diagnosis of LS. This is an area that has not been addressed by any other studies whose purpose has been to monitor LS in the secondary care/specialist environment. If a combination of the Delphi-agreed features in a diagnostic tool can reasonably pick-up LS, further validity could be investigated through including histological assessment of people identified as well as longer term outcomes of those patients.

Conclusion

This 4-round e-Delphi exercise has elicited a list of 17 clinical diagnostic features for vulval LS which may improve identification of LS by non-experts. The results will inform future research in the field of LS diagnosis.

References:

1. International Classification of Diseases for Mortality and Morbidity Statistics, 11th Revision, v2024-01 EB60.0 Lichen sclerosus of vulva - ICD-11 MMS (findacode.com). Last accessed 19th April 2024.
2. Goldstein AT, Marinoff SC, Christopher K, Srodon M. Prevalence of vulvar lichen sclerosus in a general gynecology practice. J Reprod Med. 2005;50(7):477-80.
3. Lansdorp CA, van den Hondel KE, Korfage IJ, et al. Quality of life in Dutch women with lichen sclerosus. Br J Dermatol. 2013;168(4):787-93.
4. Kirtschig G, Becker K, Gunthert A, Jasaitiene D, Cooper S, et al. Evidence-based (S3) Guideline on (anogenital) Lichen sclerosus. J Eur Acad Dermatol Venereol. 2015;29(10):e1-43.
5. Leibovitz A, Kaplun VV, Saposhnicov N, Habot B. Vulvovaginal examinations in elderly nursing home women residents. Arch Gerontol Geriatr. 2000;31(1):1-4.
6. Tasker GL, Wojnarowska F. Lichen sclerosus. Clin Exp Dermatol. 2003;28(2):128-33.
7. Virgili A, Borghi A, Toni G, et al. Prospective clinical and epidemiologic study of vulvar lichen sclerosus: analysis of prevalence and severity of clinical features, together with historical and demographic associations. Dermatology. 2014;228(2):145-51.
8. Sargeant HA, O'Callaghan FV. The impact of chronic vulval pain on quality of life and psychosocial well-being. Aust N Z J Obstet Gynaecol. 2007;47(3):235-9.
9. Hickey S, Bell H. Quality of life in the vulvar clinic: a pilot study. J Low Genit Tract Dis. 2010;14(3):225-9.
10. Halonen P, Jakobsson M, Heikinheimo O, et al. Lichen sclerosus and risk of cancer. Int J Cancer. 2017;140(9):1998-2002.
11. Bleeker MC, Visser PJ, Overbeek LI, van Beurden M, Berkhof J. Lichen Sclerosus: Incidence and Risk of Vulvar Squamous Cell Carcinoma. Cancer Epidemiol Biomarkers Prev. 2016;25(8):1224-30.
12. Meani R, Howard A, Veysey E. Incidence of vulval squamous cell carcinoma in women with vulval lichen sclerosus in an Australian tertiary referral centre. Australas J Dermatol. 2019;60(1):76-7.
13. Sargeant HA, O'Callaghan F. Predictors of psychological well-being in a sample of women with vulval pain. J Reprod Med. 2009;54(2):109-16.
14. Lee A, Bradford J, Fischer G. Long-term Management of Adult Vulvar Lichen Sclerosus: A Prospective Cohort Study of 507 Women. JAMA Dermatol. 2015;151(10):1061-7.
15. Claire Drumm, Sumir Chawla, Georgios Kravvas, et al. Representation of genital dermatoses in dermatological literature, *Br J Dermatol*. 2023; 188(4):563-4.
16. Lewis FM, Tatnall FM, Velangi SS, Bunker CB, Kumar A, Brackenbury F, et al. British Association of Dermatologists guidelines for the management of lichen sclerosus, Br J Dermatol. 2018;178(4):839-53.
17. Simpson RC, Cooper SM, Kirtschig G, Larsen S, Lawton S, et al. Future research priorities for lichen sclerosus - results of a James Lind Alliance Priority Setting Partnership. Br J Dermatol. 2019;180(5):1236-7.
18. Murphy MK, Black NA, Lamping DL, McKee CM, Sanderson CF, et al. Consensus development methods, and their use in clinical guideline development. Health Technol Assess. 1998;2(3):i-iv, 1-88.
19. Vernon W. The Delphi technique: a review. International Journal of Therapy and rehabilitation. 2009;16(2):69-76.
20. Simpson RC, Thomas KS, Leighton P, Murphy R. Diagnostic criteria for erosive lichen planus affecting the vulva: an international electronic-Delphi consensus exercise. Br J Dermatol. 2013;169(2):337-43.
21. Simpson RC. Lichen sclerosus diagnosis: which are the most important clinical diagnostic features? University of Nottingham. ls-protocol.pdf (nottingham.ac.uk)
22. Chi CC, Kirtschig G, Baldo M, et al. Topical interventions for genital lichen sclerosus. Cochrane Database Syst Rev. 2011;2011(12):CD008240.

23. Sheinis M, Selk A. Development of the Adult Vulvar Lichen Sclerosus Severity Scale-A Delphi Consensus Exercise for Item Generation. J Low Genit Tract Dis. 2018;22(1):66-73.
24. Green N, Sheinis M, Selk A. Vulvar Lichen Sclerosus: Outcomes Important to Patients in Assessing Disease Severity. J Low Genit Tract Dis. 2020;24(3):299-304.
25. Welphi. Welphi [Internet]. 2020. Available from: http://www.welphi.com.
26. Hsu C-C, Sandford BA. The Delphi technique: making sense of consensus. Practical assessment, research, and evaluation. 2007;12(10).
27. Gunthert AR, Duclos K, Jahns BG, Krause E, Amann E, et al. Clinical scoring system for vulvar lichen sclerosus. J Sex Med. 2012;9(9):2342-50.
28. Kirtschig G, Kinberger M, Kreuter A, Simpson R, Günther A, van Hees C, et al. EuroGuiDerm guideline lichen sclerosus version [1] [internet]. June 2023 [cited 2023 September 27]. Available from: https://www.guidelines.edf.one//uploads/attachments/cljzpyj9r1hqqyxjrdbpy2szy-lichen-sclerosus-guideline.pdf
29. Almadori A, Zenner N, Boyle D, Swale V, Reid W, et al. Development and Validation of a Clinical Grading Scale to Assess the Vulvar Region: The Vulvar Architecture Severity Scale. Aesthet Surg J. 2020;40(12):1319-26.
30. Boero V, Liverani CA, Brambilla M, Monti E, Tozzi R, et al. The "CIV Classification," a New Proposal for the Architectural Grading of Vulvar Lichen Sclerosus. J Low Genit Tract Dis. 2021;25(4):291-5.
31. Sheinis M, Green N, Vieira-Baptista P, Carriero C, Fischer G, et al. Adult Vulvar Lichen Sclerosus: Can Experts Agree on the Assessment of Disease Severity? J Low Genit Tract Dis. 2020;24(3):295-8.
32. Krapf JM, Smith AB, Cigna ST, Goldstein AT. Presenting Symptoms and Diagnosis of Vulvar Lichen Sclerosus in Premenopausal Women: A Cross-Sectional Study. J Low Genit Tract Dis. 2022;26(3):271-5.
33. Wang M, Wininger M, Vash-Margita A. The SWIFT Model for Lichen Sclerosus Among Premenarchal Girls. J Low Genit Tract Dis. 2022;26(1):46-52.
34. Kirby L, Gran S, Orekoya F, et al. Is urinary incontinence associated with vulval lichen sclerosus in women? A cross-sectional study. Br J Dermatol. 2021;185(5):1063-1065.
35. Cooper SM, Gao XH, Powell JJ, Wojnarowska F. Does treatment of vulvar lichen sclerosus influence its prognosis? Arch Dermatol. 2004;140(6):702-6.
36. Renaud-Vilmer C, Cavelier-Balloy B, Porcher R, Dubertret L. Vulvar lichen sclerosus: effect of long-term topical application of a potent steroid on the course of the disease. Arch Dermatol. 2004;140(6):709-12.
37. Establishing Effective Diagnostic Criteria for Lichen Sclerosus (SHELLS) study webpage SHELLS - The University of Nottingham. Last accessed 19th April 2024.

Tables

Table 1: *Demographics of participants in the e-Delphi exercise who completed all three rounds.*

	Round 1 N (%)	**Round 2 N (%)**	**Round 3 N (%)**	**Round 4 N (%)**
Total participants	47	44	42	36
Stakeholder Group*				
Healthcare professionals	37 (79%)	34 (77%)	32 (76%)	28 (78%)
Patients/Patient Support Group Representatives	9 (19%)	9 (20%)	9 (21%)	7 (19%)
Researchers	2 (4%)	2 (5%)	2 (5%)	2 (6%)
Healthcare Profession				
Dermatologist	16 (34%)	16 (36%)	15 (36%)	12 (33%)
General/Family Practitioner	1 (2%)	1 (2%)	1 (2%)	1 (3%)
Gynaecologist	15 (32%)	13 (30%)	12 (29%)	11 (31%)
Nurse	2 (4%)	2 (5%)	2 (5%)	2 (6%)
Urologist	1 (2%)	0	0	0
Other Healthcare Professional	9 (19%)	9 (20%)	9 (21%)	3 (8%)
Not Applicable	3 (6%)	3 (7%)	3 (6%)	7 (19%)

Country*				
Australia	1 (2%)	1 (2%)	1 (2%)	1 (3%)
Brazil	1 (2%)	1 (2%)	1 (2%)	1 (3%)
Canada	5 (11%)	5 (11%)	5 (12%)	4 (11%)
Denmark	4 (9%)	3 (7%)	3 (7%)	3 (8%)
France	1 (2%)	1 (2%)	1 (2%)	1 (3%)
Germany	2 (4%)	2 (5%)	2 (5%)	2 (6%)
Israel	1 (2%)	1 (2%)	1 (2%)	1 (3%)
Italy	1 (2%)	0	0	0
Luxembourg	1 (2%)	1 (2%)	1 (2%)	1 (3%)
Netherlands	3 (6%)	3 (7%)	2 (5%)	0
Russia	1 (2%)	1 (2%)	1 (2%)	1 (3%)
Switzerland	3 (6%)	3 (7%)	3 (7%)	2 (6%)
United Kingdom	14 (30%)	14 (32%)	13 (31%)	12 (33%)
United States of America	10 (21%)	9 (20%)	9 (21%)	8 (22%)
Experience with LS Patients				
Not Applicable	4 (9%)	3 (7%)	3 (7%)	3 (8%)

<1 year	0	0	0	0
2-5 years	5 (11%)	5 (11%)	5 (12%)	4 (11%)
6-10 years	10 (21%)	10 (23%)	10 (21%)	8 (22%)
11-20 years	12 (26%)	10 (23%)	9 (21%)	8 (22%)
>20 years	16 (34%)	16 (36%)	15 (36%)	13 (36%)

*Please note that the total number of participants in these groups may be higher than 47 as some participants represented more than one category of stakeholder group.

Table 2: *Proportion of participants rating diagnostic features as 'not important', 'important but not critical', and 'critical' after three rounds of voting in the e-Delphi survey. Key – Green = meets consensus as critical (>70% participants scored 7-9); Amber=important but not critical (>70% score 4-9); Red = meets consensus as not important (>70% participants score 0-3); White = consensus not met.*

Diagnostic Feature	Critical (7-9)	Important But Not Critical (4-6)	Not Important (1-3)	Unable to Score
Changes in the anatomy of the genital area	**84%**	**16%**	**0%**	**0%**
Response to topical corticosteroids**	**79%**	**12%**	**6%**	**2%**
Whiteness	**78%**	**23%**	**0%**	**0%**
Itch	**71%**	**25%**	**6%**	**0%**
Burying of the clitoral area**	**71%**	**19%**	**7%**	**5%**
Absence of vaginal involvement**	65%	27%	7%	2%
Fissuring	58%	33%	6%	3%
How important do you think it is to separate out early and late changes of lichen sclerosus when making a diagnosis?**	55%	39%	2%	5%
Crinkly skin	53%	35%	13%	2%
Bruising/bleeding under the skin	41%	50%	8%	0%
Fissuring at the back entrance to the vagina**	41%	38%	19%	2%
Pain/soreness unrelated to sexual activity	41%	32%	23%	3%
Pain/soreness related to sexual activity	40%	44%	13%	3%
Skin thickening	40%	37%	21%	2%
Loss of skin stretchiness	40%	36%	21%	3%
Erosions	37%	40%	21%	3%
Irritation	35%	39%	21%	5%
Perianal fissure**	26%	48%	26%	0%
Changes affect non-hair bearing skin**	40%	19%	38%	2%
Skin hardening	38%	23%	38%	2%
Blistering	21%	14%	65%	0%
Anorectal itch**	19%	50%	31%	0%
Dryness of skin	19%	19%	60%	2%
Brownish/darker colour change	17%	20%	62%	0%
Sensory disturbance other than pain in the vulval area**	14%	19%	67%	0%
Urinary incontinence**	12%	18%	69%	0%
Constipation	10%	30%	59%	2%
Redness	10%	53%	36%	0%
Disturbed urinary stream	8%	36%	54%	2%
Swelling of the skin	5%	29%	66%	0%
Bladder pain*	5%	24%	71%	2%

*Bladder pain met consensus for exclusion after round 2 (>70% rate it as 'not important') and therefore did not appear in subsequent rounds. **Additional features added before round two based on suggestions made in round one. Please note that due to rounding, some totals add to more than 100%

Table 3: *Average ranking of the 12 diagnostic features (signs and symptoms) which were voted as being 'important but not critical'.* ***The lower the average ranking, the higher the importance.*** *The table is ordered highest ranking to lowest ranking.*

Diagnostic Feature	Average Ranking
Fissuring	3.53
Crinkly skin	3.78
Bruising/bleeding under the skin	4.58
Fissuring at the back entrance to the vagina	6.19
Absence of vaginal involvement	6.22
Erosions	6.31
Loss of skin stretchiness	6.31
Irritation	6.67
Skin thickening	7.00
Pain/soreness related to sexual activity	8.28
Perianal fissure	9.50
Pain/soreness unrelated to sexual activity	9.64
How important do you think it is to separate out early and late changes of lichen sclerosus when making a diagnosis?	Not included in ranking