

**Dermatology Reports**

https://journals.pagepress.net/dr

**eISSN 2036-7406**

**Publisher's Disclaimer.** E-publishing ahead of print is increasingly important for the rapid dissemination of science. **Dermatology Reports** is, therefore, E-publishing PDF files of an early version of manuscripts that undergone a regular peer review and have been accepted for publication, but have not been through the copyediting, typesetting, pagination and proofreading processes, which may lead to differences between this version and the final one.

The final version of the manuscript will then appear on a regular issue of the journal.

E-publishing of this PDF file has been approved by the authors.

*Please cite this article as:*

*Thi Hong Nguyen N, Huu Le D, Nguyet Vu M, et al*. Validation of a sensitive skin screening questionnaire in a Southeast Asian population. *Dermatol Rep 2026 [Epub Ahead of Print] doi: 10.4081/dr.2026.10944*

© the Author(s), 2026
*Licensee* PAGEPress, Italy

Received: 13 May 2026; Accepted: 2 November 2026.

Note: The publisher is not responsible for the content or functionality of any supporting information supplied by the authors. Any queries should be directed to the corresponding author for the article.
All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article or claim that may be made by its manufacturer is not guaranteed or endorsed by the publisher.

**Validation of a sensitive skin screening questionnaire in a Southeast Asian population**

Nhung Thi Hong Nguyen,[1] Doanh Huu Le,[1,2] Minh Nguyet Vu,[1,2] Hien Thi Thu Do,[2,3] Ha Thai Vu,[1,2] Vinh Thi Ha Nguyen[1,2]

[1]Department of Dermatology and Venereology, Hanoi Medical University; [2]Department of Outpatient, National Hospital of Dermatology and Venereology, Hanoi, Vietnam; [3]VNU University of Medicine and Pharmacy, Vietnam National University, Hanoi, Vietnam

*Correspondence: Vinh Thi Ha Nguyen, Department of Dermatology and Venereology, Hanoi Medical University, Hanoi, 10000, Vietnam. E-mail: ntt.academic@gmail.com; Tel.: +84987934012*

**Key words:** sensitive skin syndrome; lactic acid stinging test; screening questionnaire; diagnostic accuracy; skin sensitivity; Vietnam.

**Contributions:** Nhung Thi Hong Nguyen: formal analysis, data interpretation, writing – review & editing; Doanh Huu Le: methodology, investigation, writing – review & editing; Minh Nguyet Vu: conceptualization, supervision, writing – review & editing; Hien Thi Thu Do: conceptualization, methodology, data curation, writing – original draft, writing – review & editing; Ha Thai Vu: investigation, data collection, writing – review & editing; Vinh Thi Ha Nguyen: data curation, writing – original draft, writing – review & editing. All authors read and approved the final manuscript.

**Conflict of interest:** the authors have no conflict of interest to declare.

**Ethics approval and consent to participate:** the study protocol was approved by the Ethics Committee of the National Hospital of Dermatology and Venereology, Vietnam (Approval No. 56/HĐĐĐ-BVDLTW, 20 November 2024). All participants provided written informed consent prior to enrollment. Reporting adhered to the STARD 2015 guidelines for diagnostic accuracy studies.

**Availability of data and materials:** the datasets used and/or analyzed during the current study are available upon reasonable request from the corresponding author.

**Declaration of generative artificial intelligence and artificial intelligence-assisted technologies in the writing process:** the authors declare that no generative artificial intelligence was used in preparing this manuscript.

**Abstract**

Sensitive skin syndrome (SSS) is common, but the lactic acid stinging test (LAST), a provocation-based reference procedure, is not always feasible in routine practice. This cross-sectional diagnostic accuracy study evaluated a brief trigger-based questionnaire against LAST in Vietnamese adult outpatients attending for health check-ups. A total of 198 participants completed a questionnaire including one self-identification item and 10 trigger items before undergoing LAST. The questionnaire showed acceptable discrimination for LAST positivity (area under the curve [AUC]=0.78; 95% confidence interval [CI]: 0.72-0.84). A *post hoc* cut-off of ≥3 triggers provided 68.9% sensitivity, 78.9% specificity, and 72.7% overall accuracy. LAST positivity was more frequent among women than men, and Fitzpatrick phototype distributions differed significantly between LAST groups. Self-reported sensitive skin, sun exposure, and cosmetic use were associated with LAST positivity in the multivariable analysis, although the estimates require caution because of imprecision and possible overfitting. This concise questionnaire may help identify adults likely to be LAST-positive in comparable outpatient settings. LAST positivity is not a definitive diagnosis of SSS, and external validation is required before broader use.

**Introduction**

Sensitive skin syndrome (SSS) is defined by the International Forum for the Study of Itch (IFSI) as a condition characterized by unpleasant sensory experiences, including stinging, burning, pain, pruritus, and tingling, elicited by stimuli that would not normally provoke such responses. These sensations are not attributable to clinically detectable skin disease; the skin may appear normal or present with transient erythema, and any body site can be affected, with the face most commonly involved.[1] SSS is highly prevalent and can substantially impair quality of life. Population-based surveys in Europe report prevalence estimates approaching 50%, with consistently higher rates among women and younger individuals.[2,3] In contrast, evidence from South-East Asia, including Vietnam, remains scarce, despite the likelihood of region-specific influences related to environmental exposures, cultural practices, and genetic background. Current pathophysiological hypotheses emphasize heightened cutaneous neurosensory responsiveness in conjunction with impairment of the epidermal barrier.[1,4,5]

Several approaches have been proposed to investigate or operationalize SSS. Among provocation-based methods, the lactic acid stinging test (LAST) has long been regarded as a reference procedure.[6] However, its application requires specific reagents and trained personnel and may induce discomfort, limiting feasibility in routine clinical practice, and large-scale studies.[6,7] Consequently,

self-assessment questionnaires have emerged as pragmatic alternatives, particularly in outpatient and epidemiological settings. Most available instruments, however, focus on quantifying symptom intensity rather than establishing a diagnostic classification.[8,9] Importantly, no questionnaire has yet been validated for diagnostic use in a Vietnamese population,[10] representing a critical barrier to standardized case identification for both clinical care and research in this context.

In this study, we assessed the performance of a brief self-administered questionnaire for identifying LAST-positive Vietnamese adult outpatients, using LAST as an operational reference procedure rather than a definitive diagnostic standard for SSS. The specific aims were to evaluate discriminatory accuracy and identify a practical cut-off value, to examine the contribution of individual triggering factors, and to describe key demographic and symptom characteristics. We hypothesized that the questionnaire would demonstrate acceptable discrimination relative to LAST and yield a candidate threshold for further validation in clinical and community settings.

## Materials and Methods

### *Study design and setting*

We conducted a single-center, cross-sectional diagnostic accuracy study at the National Hospital of Dermatology and Venereology, Hanoi, Vietnam, between November 2024 and July 2025. The terms "sensitive skin" and "sensitive skin syndrome" refer to the subjective sensory syndrome defined by the IFSI. LAST served as the operational reference procedure, classifying participants as LAST-positive or LAST-negative. This classification describes a response to a standardized provocation test and does not establish or exclude a definitive clinical diagnosis of SSS.

The expected proportion of LAST-positive participants (66%) was adapted from Pan *et al.*, based on comparable Asian cohorts; a sample size of n=198 was calculated to provide a 95% confidence interval (CI) with ±6.6% precision using the formula for estimation of proportions.[7] The study was reported in accordance with the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) guidelines for cross-sectional studies.[11]

### *Participants*

Adults (≥18 years) attending the dermatology outpatient clinics for health check-ups were recruited consecutively if they met the eligibility criteria. Recruitment was clinic-based and did not use population-based probability sampling. Inclusion criteria were age ≥18 years and provision of written informed consent. Exclusion criteria comprised age <18 years; clinical signs or a history of dermatological disorders that could plausibly explain the symptoms under investigation (*e.g.*, acne,

rosacea, seborrheic dermatitis, atopic dermatitis, psoriasis, or active skin infections), thereby confounding symptom attribution; recent intense sun or ultraviolet exposure, or use of topical or systemic treatments likely to alter skin sensitivity within the preceding 3 months; pregnancy or breastfeeding; and inability to understand or complete the questionnaire.

***Study procedures***

Self-assessment questionnaire

*Development, pilot testing, and reliability*

Ten triggering factors were prespecified *a priori* based on published evidence and clinical relevance to sensitive skin.[2,12] A pilot sample of 30 adults completed the questionnaire twice, with a 3-week interval, to assess test-retest reliability using Cohen's κ. Item wording was refined where necessary, and comprehension difficulties were documented. The pilot cohort was independent of the main study population and was not included in the diagnostic accuracy analyses. In the pilot study, item-specific κ values ranged from 0.72 to 0.89, with minor wording modifications addressing two identified comprehension issues.

Questionnaire structure

The questionnaire was developed through literature review, expert consensus, and preliminary testing. The expert panel comprised five specialists in dermatology, and consensus was reached through discussion until agreement was achieved. It comprised two sections. Part 1 presented participants with the IFSI definition of sensitive skin and asked whether they considered themselves to have sensitive skin, followed by an overall sensitivity rating on a visual analogue scale (VAS; 0-10). The VAS was used for descriptive purposes only and was not incorporated into the diagnostic score. Part 2 constituted the index test. Participants reported whether each of 10 predefined factors triggered unpleasant cutaneous sensations (stinging, burning, pain, pruritus, or tingling): 1) sun exposure; 2) hot or dry weather or environment; 3) cold or wet weather or environment; 4) wind; 5) contact with water; 6) physical exercise; 7) hygiene soaps or cleansers; 8) cosmetics; 9) smog or pollutants; and 10) psychological stress. Each affirmative response scored one point, yielding a total index-test score ranging from 0 to 10.

Lactic acid stinging test

A volume of 50 μL of 10% lactic acid solution (prepared from an 88% lactic acid stock; Guangdong Guanghua Sci-Tech Co., Ltd., China) was applied to the right nasolabial fold, with saline

applied to the contralateral side as a control. Participants rated stinging intensity at 2.5 and 5 minutes using a 4-point scale (0 = none; 1 = mild; 2 = moderate; 3 = strong). A cumulative score ≥3 defined LAST positivity, in accordance with established criteria.[6]

Blinding

Participants completed the questionnaire immediately before undergoing LAST during the same visit, without intervening treatments. Assessors performing the reference standard had routine access to clinical information and were not fully blinded to questionnaire responses, introducing potential review bias. To mitigate this risk, data collection was temporally structured, and questionnaire scoring was conducted independently of LAST outcomes.

Indeterminate and missing data

Among the 198 participants included in the analysis, no indeterminate LAST results were observed, and questionnaire data were complete. Two of the 200 individuals initially considered for inclusion were excluded because they provided insufficient information. Thus, completeness of questionnaire data refers to the final analytic sample; no missing questionnaire responses were imputed.

***Statistical analysis***

Continuous variables were summarized as mean ± standard deviation or median [interquartile range], as appropriate, and categorical variables as counts and percentages. Between-group comparisons employed Student's *t*-test or the Mann-Whitney U test for continuous variables and $\chi^2$ or Fisher's exact tests for categorical variables. Proportions, including sensitivity and specificity, were reported with 95% CIs calculated using the Wilson method. Diagnostic accuracy for LAST positivity was assessed using receiver operating characteristic (ROC) analyses across all integer cut-offs of the questionnaire trigger score (0-10). Sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and overall accuracy were calculated for each cut-off. The optimal questionnaire threshold was selected *post hoc* by maximizing Youden's J statistic (sensitivity + specificity −1), with the chosen cut-off reported in the *Results*. LAST positivity was predefined as a cumulative score ≥3.

A prespecified multivariable logistic regression model included age, sex, Fitzpatrick phototype, self-reported sensitive skin, and the 10 trigger variables, giving 14 predictor parameters, excluding the intercept. These associations were interpreted cautiously in view of the sample size relative to model

complexity. Results are presented as odds ratios (ORs) with 95% CIs and two-sided p-values. Model diagnostics addressed multicollinearity (variance inflation factors) and calibration (intercept, slope, Brier score, and Hosmer-Lemeshow test).[13-15] Decision-curve analysis evaluated clinical utility across threshold probabilities ranging from 0.10 to 0.70, comparing the questionnaire-based decision rule with treat-all and treat-none strategies. Net benefit and 95% CIs were estimated using bootstrap resampling (B=1000).[15-17] Uncertainty was summarized as follows: 95% CIs for the area under the ROC curve (AUC) were derived by bootstrap (B=1000, percentile method); 95% CIs for sensitivity, specificity, PPV, and NPV were calculated using the Wilson method; and 95% CIs for ORs were obtained using Wald intervals. All statistical tests were two-sided with a significance level of $\alpha$=0.05. Analyses were performed using Python, with NumPy and Pandas for data handling, scikit-learn for ROC and diagnostic metrics, statsmodels for logistic regression, and matplotlib for graphical outputs. Decision-curve analyses were implemented using custom scripts.

**Results**

Two hundred individuals were initially considered for inclusion. Two were excluded because they provided insufficient information, leaving 198 adults in the analysis (age range, 18-65 years; mean [standard deviation], 34.2 [11.3]). Of these, 122 (61.6%) were LAST-positive and 76 (38.4%) were LAST-negative. The 61.6% is the observed proportion of LAST-positive participants in this clinic sample and is not an estimate of SSS prevalence in the Vietnamese general population. LAST positivity was more frequent among women than men (67.2% *vs.* 51.4%; p=0.029), whereas mean age did not differ between LAST-positive and LAST-negative groups (p=0.807). Fitzpatrick phototype distributions differed significantly (p=0.010), with a higher proportion of phototype III relative to IV among LAST-positive individuals. Baseline characteristics are summarized in *Supplementary Table 1*.

Overall, 120 of 198 participants (60.6%) self-identified as having sensitive skin. When compared with LAST as the reference standard, self-report demonstrated a sensitivity of 86.9% (95% CI: 79.8-91.8), specificity of 81.6% (95% CI: 71.4-88.7), PPV of 88.3% (95% CI: 81.4-92.9), NPV of 79.5% (95% CI: 69.2-87.0), and an overall accuracy of 84.8% (95% CI: 79.2-89.2). Agreement with LAST yielded Cohen's $\kappa$=0.68 (Table 1).

ROC analysis of the questionnaire trigger score showed moderate discriminatory ability, with an AUC of 0.78 (95% CI: 0.72-0.84) (Figure 1A). At the *post hoc* selected cut-off of ≥3 triggers, sensitivity was 68.9% (95% CI: 60.2-76.4), specificity 78.9% (95% CI: 68.5-86.6), PPV 84.0% (95% CI: 75.6-89.9), NPV 61.2% (95% CI: 51.3-70.3), and accuracy 72.7% (95% CI: 66.1-78.5). As

anticipated, increasing the cut-off improved specificity at the expense of sensitivity (for example, at ≥5 triggers, sensitivity 26.2% and specificity 93.4%). Figure 1B illustrates threshold-specific performance. No adverse events related to LAST were observed.

Item-level analyses revealed heterogeneity in diagnostic performance across individual triggers (*Supplementary Table 2*). Sun exposure and hot or dry environments showed relatively balanced sensitivity and specificity, whereas psychological stress exhibited low sensitivity (18.9%) but high specificity (88.2%). Trigger prevalence and full item-level sensitivity and specificity estimates with 95% CIs are reported to facilitate interpretation.

In the prespecified multivariable logistic regression model, self-reported sensitive skin was associated with LAST positivity (OR: 48.27; 95% CI: 15.20-153.25; $p<0.001$). The wide CI indicates substantial uncertainty about the magnitude of this association, and the estimate should be interpreted cautiously. Adjusted associations were also observed for sun exposure (OR: 3.18; 95% CI: 1.06-9.58; $p=0.04$) and cosmetic use (OR: 3.73; 95% CI: 1.01-13.79; $p<0.05$). Fitzpatrick phototype III did not meet the conventional significance threshold ($p=0.06$), and age and sex were not significantly associated with LAST positivity. Full model estimates are presented in Table 2.

The calibration intercept was 0.03 and the slope was 1.06. The decile-based calibration plot showed deviations from the 45° reference line, including overestimation in some high-risk groups (Figure 2). The Hosmer-Lemeshow test was significant ($p<0.001$), indicating evidence of lack of fit by this test. These findings warrant caution: calibration summaries within this cohort do not establish model stability or performance in new participants.

Across threshold probabilities from 0.10 to 0.70, the ≥3-trigger questionnaire rule had a positive estimated net benefit relative to classifying all participants as LAST-negative. It exceeded the net benefit of classifying all participants as LAST-positive only above approximately 0.39; at lower thresholds, the classify-all-as-positive strategy had greater estimated net benefit (Figure 3). Bootstrap-derived 95% confidence bands illustrate uncertainty around the questionnaire rule and do not establish superiority at every threshold.

**Discussion**

In this cross-sectional diagnostic accuracy study, a brief trigger-based questionnaire identified LAST-positive adult outpatients with moderate discrimination (AUC 0.78). A *post hoc* threshold of ≥3 reported triggers achieved sensitivity of 68.9% and specificity of 78.9%. Decision-curve analysis suggested potential utility over a restricted range of decision thresholds, while the calibration results and sample size warrant caution. These findings support further evaluation of the questionnaire in

comparable outpatient settings; they do not establish a definitive diagnosis of SSS or justify population-wide use.

Most prior research on sensitive skin has relied on self-identification or symptom severity instruments, such as the Sensitive Scale-10 (SS-10), without formal classification against a reference stimulus.[1-5,8,9] While SS-10 captures short-term symptom burden, it does not support dichotomous case definition. By directly benchmarking questionnaire performance against the LAST in an adult Asian outpatient population, the present study addresses a methodological gap.[6,7,10] Anchoring the instrument to a reference provocation test enabled explicit threshold selection, calibration assessment, and decision-analytic evaluation. These analyses concern LAST-defined outcomes and do not establish clinical SSS status. The observed agreement between self-reported sensitive skin and LAST ($\kappa$=0.68) supports convergent validity, while the structured trigger score adds standardization beyond narrative clinical assessment.

The 11-item instrument requires no specialized equipment and may be feasible in busy outpatient clinics. However, at the ≥3-trigger threshold, 38 of 122 LAST-positive participants were not identified, so a negative questionnaire result should not be used to exclude SSS. An initial self-identification question followed by structured trigger assessment may help organize symptom discussions and identify candidates for further assessment, but this proposed workflow requires prospective evaluation. Predictive values depend on the proportion of LAST-positive participants and the case mix. The decision curve does not establish transportability to populations with a different baseline risk.

Item-level analyses describe potential symptom patterns. Triggers related to sun exposure and hot or dry environments showed relatively balanced sensitivity and specificity, consistent with prevailing hypotheses of neurosensory hyperresponsiveness and epidermal barrier impairment.[1,4,5] Cosmetic use was associated with LAST positivity after adjustment, but its wide CI and the cross-sectional design limit conclusions about effect size and causality. In contrast, psychological stress demonstrated high specificity but low sensitivity, limiting its utility for triage, likely reflecting interindividual variability and contextual influences. Mechanistically, the prominence of sunlight and heat or dryness aligns with evidence implicating heat-sensitive transient receptor potential channels, small-fiber excitability, and barrier susceptibility, with potential modulation by skin phototype.[1,2,4,5]

The large adjusted association between self-reported sensitive skin and LAST positivity (OR: 48.27; 95% CI: 15.20-153.25) should not be interpreted as a precise or causal effect. The model included 14 predictor parameters, excluding the intercept, in 198 participants with 122 LAST-positive and 76 LAST-negative outcomes. This corresponds to approximately 8.7 LAST-positive and 5.4 LAST-negative participants per predictor parameter. These ratios are descriptive indicators of limited

information relative to model complexity, rather than definitive thresholds for model adequacy. Sparse predictor-outcome combinations may produce unstable or inflated estimates, and overfitting cannot be excluded.[13,14] The sample-size calculation was based on estimation of a proportion, not on the requirements of multivariable modeling. The adjusted associations should therefore be regarded as exploratory and require confirmation in larger independent samples.

Several considerations temper interpretation. Self-reported triggers may be misclassified, and the direction of any resulting bias is uncertain. The smaller number of LAST-negative participants limits precision for specificity. The ≥3-trigger cut-off was selected *post hoc* using Youden's index and evaluated in the same cohort; its performance may therefore be optimistic. Decision-curve analysis does not remove this selection bias. Likewise, calibration summaries and bootstrap CIs do not by themselves demonstrate stable predictor effects or rule out overfitting. Independent validation, including calibration assessment, is needed.[13,14,18]

LAST positivity was observed in 122 of 198 participants (61.6%), and 120 participants (60.6%) self-identified as having sensitive skin. These proportions describe adults attending a dermatology outpatient service for health check-ups. Although eligible attendees were recruited consecutively, the sample was not drawn from the general population. Attendance at a specialist dermatology service may select individuals with greater skin-related concerns, cosmetic intolerance, or awareness of cutaneous sensations, even when the stated reason for attendance is a health check-up. Such enrichment is plausible but was not quantified. Consequently, the observed 61.6% LAST positivity cannot be interpreted as the prevalence of SSS in the Vietnamese general population. It is not directly comparable with the approximately 50% estimates from population-based surveys, which differ in sampling frame, population characteristics, and often rely on self-reported sensitivity rather than a provocation-test definition.[2,3,7] These differences may explain part of the higher proportions in this clinic sample, but do not demonstrate a higher population prevalence in Vietnam.

Strengths include complete questionnaire and LAST data in the final analytic sample, evaluation of discrimination and decision-curve performance, and preliminary test-retest reliability evidence from an independent pilot.

Important limitations remain. LAST assesses a provoked sensory response and is an imperfect reference procedure for the multidimensional clinical syndrome; a positive result does not confirm SSS, and a negative result does not exclude it. Partial blinding may introduce review bias. The single-center setting, clinic-based recruitment, and exclusion of two individuals with insufficient information may affect generalizability. Seasonal and climatic factors may influence reported triggers, and the cross-sectional design precludes causal inference.

Future multicenter studies should validate the questionnaire and the selected cut-off in independently recruited outpatient and community samples, assess subgroup performance by phototype, age, and exposure profile, and compare the questionnaire with existing instruments. Larger samples and penalized regression could help assess the stability of adjusted associations. Any clinical workflow based on the questionnaire should also be evaluated for its effects on referrals, patient outcomes, and costs before widespread implementation.

## Conclusions

This brief trigger-based questionnaire showed moderate discrimination for LAST positivity among Vietnamese adults attending a dermatology outpatient service for health check-ups. The *post hoc* ≥3-trigger cut-off is a candidate threshold for further validation in comparable settings. Neither LAST positivity nor questionnaire positivity constitutes a definitive diagnosis of SSS. External validation is required before broader clinical or population use.

## References


1. Misery L, Weisshaar E, Brenaut E, et al.; Special Interest Group on sensitive skin of the International Forum for the Study of Itch (ISFI). Pathophysiology and management of sensitive skin: position paper from the special interest group on sensitive skin of the International Forum for the Study of Itch (IFSI). J Eur Acad Dermatol Venereol 2020;34:222-9.
2. Misery L, Jourdan E, Huet F, et al. Sensitive skin in France: a study on prevalence, relationship with age and skin type and impact on quality of life. J Eur Acad Dermatol Venereol 2018;32:791-5.
3. Farage MA. The Prevalence of Sensitive Skin. Front Med 2019;6:98.
4. Richters R, Falcone D, Uzunbajakava N, et al. What is sensitive skin? A systematic literature review of objective measurements. Skin Pharmacol Physiol 2015;28:75-83.
5. Buhé V, Vié K, Guéré C, et al. Pathophysiological Study of Sensitive Skin. Acta Derm Venereol 2016;96:314-8.
6. Corazza M, Guarneri F, Montesi L, et al. Proposal of a self-assessment questionnaire for the diagnosis of sensitive skin. J Cosmet Dermatol 2022;21:2488-96.
7. Pan Y, Ma X, Song Y, et al. Questionnaire and Lactic Acid Sting Test Play Different Role on the Assessment of Sensitive Skin: A Cross-sectional Study. Clin Cosmet Investig Dermatol 2021;14:1215-25.
8. Gougerot A, Vigan M, Bourrain JL, et al. SIGL: Is this a clinical evaluation tool for reactive skins? Nouv Dermatol 2007;26:461-4.

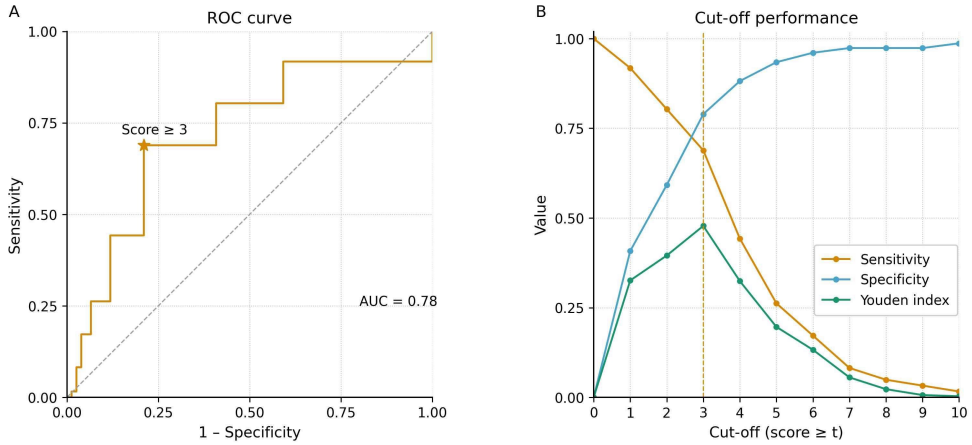
9. Misery L, Jean-Decoster C, Mery S, et al. A new ten-item questionnaire for assessing sensitive skin: the Sensitive Scale-10. Acta Derm Venereol 2014;94:635-9.
10. Guerra-Tapia A, Serra-Baldrich E, Prieto Cabezas L, et al. Diagnosis and Treatment of Sensitive Skin Syndrome: An Algorithm for Clinical Practice. Actas Dermosifiliogr (Engl Ed) 2019;110:800-8. English, Spanish.
11. von Elm E, Altman DG, Egger M, et al.; STROBE Initiative. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. J Clin Epidemiol 2008;61:344-9.
12. Brenaut E, Barnetche T, Le Gall-Ianotto C, et al. Triggering factors in sensitive skin from the worldwide patients' point of view: a systematic literature review and meta-analysis. J Eur Acad Dermatol Venereol 2020;34:230-8.
13. Steyerberg EW. Clinical Prediction Models. Springer; 2009.
14. Steyerberg EW, Vickers AJ, Cook NR, et al. Assessing the performance of prediction models: a framework for traditional and novel measures. Epidemiology (Cambridge, Mass) 2010;21:128-38.
15. Vickers AJ, Holland F. Decision curve analysis to evaluate the clinical benefit of prediction models. Spine J 2021;21:1643-8.
16. Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. Med Decis Making 2006;26:565-74.
17. Van Calster B, Wynants L, Verbeek JFM, et al. Reporting and Interpreting Decision Curve Analysis: A Guide for Investigators. Eur Urol 2018;74:796-804.
18. Stevens RJ, Poppe KK. Validation of clinical prediction models: what does the "calibration slope" really measure? J Clin Epidemiol 2020;118:93-9.

**Table 1.** Agreement between self-reported sensitive skin and LAST.

| | **LAST-positive (n=122)** | **LAST-negative (n=76)** | **Total** |
|---|---|---|---|
| Self-reported sensitive skin | 106 | 14 | 120 |
| No self-reported sensitive skin | 16 | 62 | 78 |
| Total | 122 | 76 | 198 |
| Metric | Estimate (95% CI) | | |
| Sensitivity | 86.9% (79.8-91.8) | | |
| Specificity | 81.6% (71.4-88.7) | | |
| PPV | 88.3% (81.4-92.9) | | |
| NPV | 79.5% (69.2-87.0) | | |
| Accuracy | 84.8% (79.2-89.2) | | |
| Cohen's κ | 0.68 | | |

LAST, lactic acid stinging test; PPV, positive predictive value; NPV, negative predictive value; CI, confidence interval; κ, Cohen's kappa.

**Figure 1.** Questionnaire discrimination and threshold performance. A) ROC curve showing moderate discriminatory performance of the questionnaire for LAST positivity, with an AUC of 0.78 (95% CI: 0.72-0.84). B) Diagnostic performance across integer cut-offs of the questionnaire trigger score (decision rule: score ≥*t*), with the *post hoc* selected cut-off at ≥3 triggers highlighted. At this threshold, the confusion matrix comprised 84 true positives, 16 false positives, 60 true negatives, and 38 false negatives, using LAST status as the reference. The denominators were LAST-positive n=122 and LAST-negative n=76. Wilson 95% CIs for sensitivity and specificity are reported in the text.




AUC, area under the ROC curve; ROC, receiver operating characteristic.

**Table 2.** Multivariable logistic regression for factors associated with a positive LAST result.

| Variable | p-value | OR | 95% CI |
|---|---|---|---|
| Age (per year) | 0.76 | 1.01 | 0.963-1.053 |
| Sex (female *vs.* male) | 0.32 | 1.75 | 0.583-5.25 |
| Skin phototype (III *vs.* IV) | 0.06 | 2.44 | 0.964-6.199 |
| Self-reported sensitive skin (yes *vs.* no) | <0.01* | 48.27 | 15.2-153.251 |
| Sun exposure (yes *vs.* no) | 0.04* | 3.18 | 1.055-9.577 |
| Hot/dry weather/environment (yes *vs.* no) | 0.86 | 1.10 | 0.393-3.067 |
| Cold/wet weather/environment (yes *vs.* no) | 0.49 | 1.46 | 0.504-4.233 |
| Wind (yes *vs.* no) | 0.16 | 0.33 | 0.072-1.553 |
| Contact with water (yes *vs.* no) | 0.80 | 0.76 | 0.091-6.383 |
| Physical exercise (yes *vs.* no) | 0.23 | 2.71 | 0.54-13.619 |
| Soaps/cleansers (yes *vs.* no) | 0.21 | 0.42 | 0.104-1.656 |
| Cosmetics (yes *vs.* no) | <0.05* | 3.73 | 1.009-13.79 |
| Smog/pollutants (yes *vs.* no) | 0.93 | 1.05 | 0.371-2.96 |
| Psychological stress (yes *vs.* no) | 0.09 | 0.33 | 0.091-1.181 |

OR, odds ratio; CI, confidence interval; LAST, lactic acid stinging test; *p<0.05. Adjusted ORs were derived from a prespecified multivariable logistic regression model including age (per year), sex, Fitzpatrick phototype, self-reported sensitive skin, and 10 trigger variables (sun exposure; hot/dry environment; cold/wet environment; wind; contact with water; physical exercise; soaps/cleansers; cosmetics; smog/pollutants; psychological stress). Reference categories were male sex, Fitzpatrick phototype IV, absence of each trigger, and no self-reported sensitive skin. P-values were two-sided and obtained using Wald tests. The analysis included n=198 participants. The model included 14 predictor parameters, excluding the intercept, with 122 LAST-positive and 76 LAST-negative participants. The large estimate for self-reported sensitive skin and its wide CI should be interpreted cautiously because of possible sparse-data instability and overfitting; adjusted associations are exploratory.

**Figure 2.** Calibration of the logistic regression model. Points represent observed proportions of LAST positivity with 95% Wilson CIs across deciles of predicted risk, plotted against the corresponding mean predicted probabilities. The solid line connects the decile estimates; the dashed diagonal indicates perfect calibration. The calibration intercept was 0.03 and the calibration slope was 1.06. The Hosmer-Lemeshow test was significant (p<0.001).

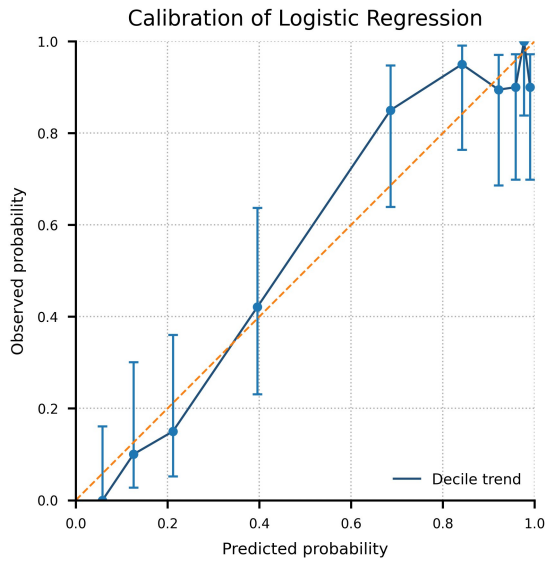
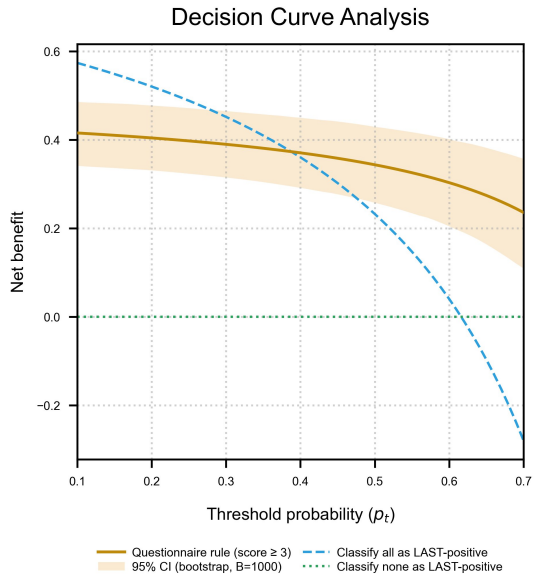


**Figure 3.** Decision-curve analysis. Net benefit of the questionnaire-based decision rule (score ≥3) compared with two default strategies, classifying all participants as LAST-positive and classifying all as LAST-negative, across threshold probabilities ranging from 0.10 to 0.70. Shaded areas represent 95% confidence bands derived from bootstrap resampling.




LAST, lactic acid stinging test; CI, confidence interval.

***Online supplementary material:***

***Supplementary Table 1.*** *Baseline characteristics of the study population.*

***Supplementary Table 2.*** *Item-level diagnostic performance of trigger items.*