

Multiclass Diagnostic Improvement of Interstitial Cystitis (IC) and Overactive Bladder (OAB) via Machine Learning Pipeline on Surface-Enhanced Raman Spectroscopy (SERS)

Minju Cho,[†] Suyeon Kang,[†] Joon Seup Hwang, Miyeon Jue, Jung Hyun Shin,^{*} and Jun Ki Kim^{*}



Cite This: <https://doi.org/10.1021/acssensors.6c02725>



Read Online

ACCESS |



Metrics & More



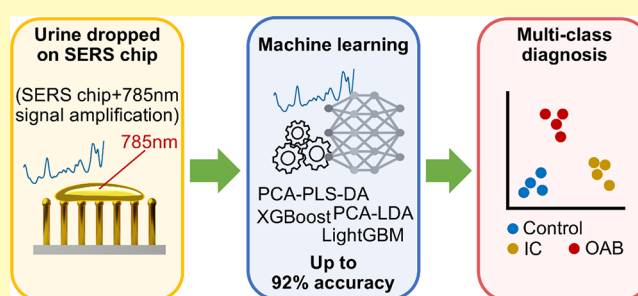
Article Recommendations



Supporting Information

ABSTRACT: Interstitial cystitis (IC) and overactive bladder (OAB) are chronic pelvic conditions with shared symptoms of urinary urgency, increased frequency, and nocturia with pain in the bladder. IC is characterized by bladder pain and inflammation without a clear etiology, whereas OAB involves detrusor overactivity in the absence of infection. Despite this symptomatic overlap, IC and OAB differ in underlying pathophysiology and require distinct treatment and medical care. A specific diagnostic strategy is currently lacking and highlights the need of a robust diagnostic method for effective treatment. Raman spectroscopy provides sensitive and nondestructive molecular fingerprints; when combined with artificial intelligence (AI)-driven analysis of spectral data, it offers a promising approach to overcoming the limitations of exclusion-based diagnoses. In this study, urinary samples were collected from healthy individuals ($n = 117$), IC patients ($n = 19$), and OAB patients ($n = 45$) for the acquisition of Raman spectra using Au-ZnO nanorod surface-enhanced Raman spectroscopy (SERS) chips. To obtain biological interpretable Raman spectra per group and high accuracy of classification performance, the linear models PCA-PLS-DA and PCA-LDA and tree-based nonlinear models XGBoost and LightGBM were applied and reached up to 92% accuracy with interpretable and pathologic relative Raman features for discriminating among healthy control, IC, and OAB. This approach suggests that urine-based SERS Raman spectra combined with machine learning could serve as a promising diagnostic platform to support disease-specific clinical decision.

KEYWORDS: interstitial cystitis, overactive bladder syndrome, surface-enhanced Raman spectroscopy, biomarker candidate, PCA-PLS-DA, PCA-LDA, LightGBM, eXtreme Gradient Boosting, machine learning-based diagnosis



Interstitial cystitis (IC), a chronic pelvic condition characterized by lower urinary tract symptoms and chronic inflammation lasting for >6 weeks, causes discomfort, pressure, or pain in the bladder. IC requires exclusion of possible etiologies such as urinary tract infection, tuberculosis, neurological conditions, malignancy, and diseases of other organs.^{1,2} IC diagnosis is typically reached through medical history, physical examination, image workups, urinalysis, urine culture, cystoscopy, and hydro-distension testing and tends to be one of exclusion. IC, therefore, is prone to misdiagnosis, especially in males, where it may be mistaken for chronic prostatitis/chronic pelvic pain syndrome or overactive bladder (OAB).^{3,4} Overactive bladder syndrome is another chronic disease characterized by the hallmark symptom of urgency, often accompanied by increased frequency and nocturia, without urinary tract infection or other evident pathology.⁵ Urgency, the primary clinical symptom used to diagnose OAB, is closely associated with frequent daytime urination, nocturia, and urinary incontinence.⁶ IC and OAB share similar symptoms, such as urgency and frequency, making it difficult to distinguish them based solely on clinical presentation. Accordingly, a robust and noninvasive diagnostic

strategy is proposed to reduce the misclassification and inform the treatment decisions.

Raman spectroscopy, which exhibits high specificity and sensitivity and offers nondestructive analysis, has been proposed as an IC diagnostic strategy, reinforcing the results of potassium sensitivity testing (PST). Detecting phenylalanine and tryptophan in the Raman spectra of freeze-dried water samples voided following PST can establish a basis for IC diagnosis.⁷ However, significant Raman signals for OAB remain lacking. Inflammatory proteins and molecules, such as VCAM-1, ICAM-1, monocyte chemotactic protein, and macrophage inflammatory protein, are elevated in urine samples from OAB

Received: June 19, 2026

Revised: September 11, 2026

Accepted: September 11, 2026

patients.⁸ This opens an avenue to detect some OAB significant Raman signals with inflammatory-related molecule signals. Additionally, recent advancements in automated data analysis using artificial intelligence (AI) have attracted attention for Raman signal analysis due to the improved sensitivity in signal acquisition based on surface-enhanced Raman spectroscopy (SERS) and computer-based analysis.⁹

Previous studies using Raman spectroscopy have employed multivariate analysis and machine learning algorithm such as principal component-partial least square-discriminant analysis (PCA-PLS-DA), principal component-linear discriminant analysis (PCA-LDA), and tree-based boosting models to discriminate cell type, drug dosages, pathological conditions, and disease states. While PCA-PLS-DA is widely used as a standard approach for Raman spectral analysis, its primary strength is the dimensional reduction and interpretability rather than optimal classification performance. However, spectroscopic differences among disease groups may involve nonlinear patterns arising from peak overlap, nonlinear responses to concentration changes, and peak-to-peak interactions, which may not be captured by the linear classification method.

Thus, the eXtreme Gradient Boosting (XGBoost) model and the Light Gradient Boosting Machine (LightGBM) are often employed to enhance the diagnostic performance. While XGBoost and LightGBM can improve diagnostic accuracy by multiple decision trees in complex Raman spectra patterns, high classification performance alone does not guarantee biologically meaningful feature selection. Moreover, decision tree-based feature importance may be sensitive to the splitting mechanism and can be biased with the limitation due to the tendency of favor features on greater splitting flexibility.¹⁰ Therefore, the integrated feature interpretation approach from linear models with high-performance nonlinear models can compensate for

the limitation of each algorithm method and for the improvement of the feature selection reliability.

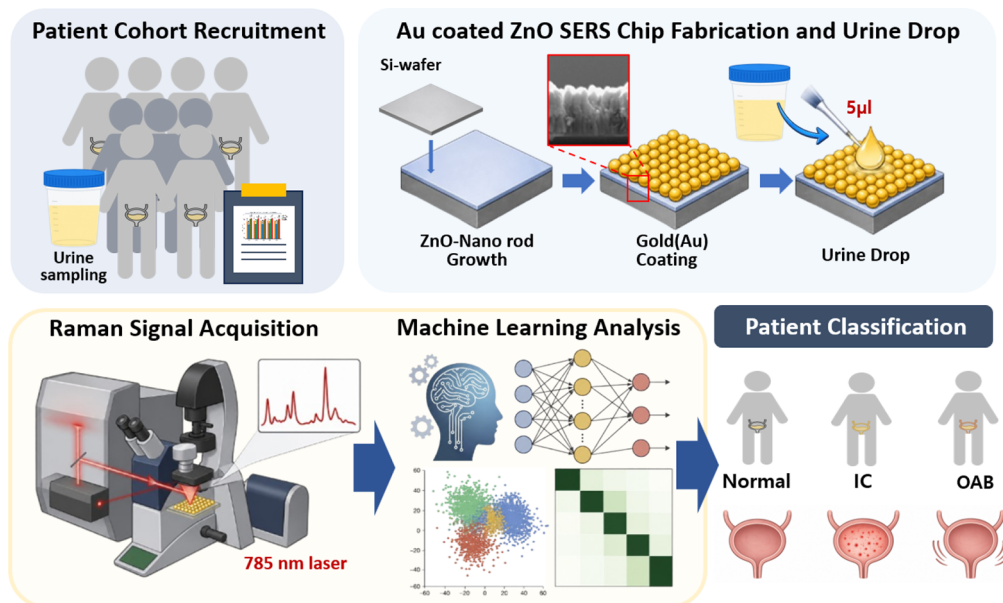
In this study, Raman signals were obtained from urine samples using a SERS chip with a gold (Au) coating zinc oxide (ZnO) nanorod for diagnostic approaches after machine learning application, as illustrated in Scheme 1. We aimed to identify the biomarkers and establish the diagnostic strategies capable of discriminating among IC and OAB. We collected urine samples from healthy individuals ($n = 117$), IC patients ($n = 19$), and OAB patients ($n = 45$). As AI-based diagnostic approaches have attracted research attention, we employed PCA-PLS-DA and PCA-LDA for the interpretable feature exploration step for identifying candidate biomarkers and XGBoost and LightGBM models to effectively model the complex nonlinear relationships and peak-to-peak interactions in Raman spectra. The findings of this study will contribute to establishing a foundation for disease diagnosis, aiming to transcend reliance on diagnosis by exclusion and lay the groundwork for a robust diagnostic approach tool.

RESULTS

Participant Enrolment and Diagnostic Criteria

The interstitial cystitis (IC) cohort included adults who reported consistent symptoms for IC at least 6 months and a pain intensity ≥ 4 on the visual analogue scale (VAS) with cystoscopy record within 12 months to confirm the presence or absence of Hunner lesion. The overactive bladder (OAB) cohort included individuals with persisting OAB symptoms more than 6 months who have urinary frequency of ≥ 8 voids per 24 h with at least one urination urgency, and Overactive Bladder Symptoms Score (OABSS) criteria of a total score ≥ 4

Scheme 1. Overall Scheme for the Diagnosis of Interstitial Cystitis and Overactive Bladder Using SERS and AI Algorithms



Urine samples from a total of normal individuals ($n = 117$), OAB patients ($n = 45$), and IC patients ($n = 19$) were collected and analyzed using a 785 nm laser to measure the Raman spectra of urine on a ZnO nanorod SERS chip with a 200 nm Au coating. Subsequently, multiple machine learning processes such as PCA-PLS-DA, PCA-LDA, XGBoost, and LightGBM were applied to distinguish between OAB and IC patients from normal and visualized the model performance with confusion matrix and ROC curves.

and item-3 (urgency) score ≥ 2 .¹¹ For the healthy control group defined as “Normal group”, men and women aged 20 years or older who were nonsmokers and who do not have prior bladder disease, urinary infection, neurologic or systemic disorder, recent urogenital procedures, or other condition affecting lower urinary tract function were recruited and enrolled after the providing written informed consent. Detailed demographic and clinical characteristics of each cohort are provided in Table S1.

Raman Signal Amplification and Urine Sample Analysis Process Using the ZnO-Au Nanorod SERS Chip

To enhance the Raman signal, we fabricated ZnO nanorods up to 500 nm in length with the deposition of gold coating with a

thickness of 200 nm on top of the nanorods. To verify the growth of nanorods with the Au coating on the SERS chips, we obtained a cross-sectional image and top-view image of the SERS chip using SEM (Figure 1A,B). The fabrication scheme (Figure S1A) and the cross-sectional and top-view images of the ZnO nanorod (Figure S1B) and the Au-coated ZnO nanorod (Figure S1C) are provided in Figure S1. We measured the Raman signal of rhodamine B to confirm the signal enhancement shown in Figure 1C. The average of the calculated enhancement factor (EF) of the rhodamine B peak is 4.03×10^6 , and the EF value of each peak is indicated. We also performed the reproducibility of the SERS chip for quality control by measuring five points of three different SERS chips

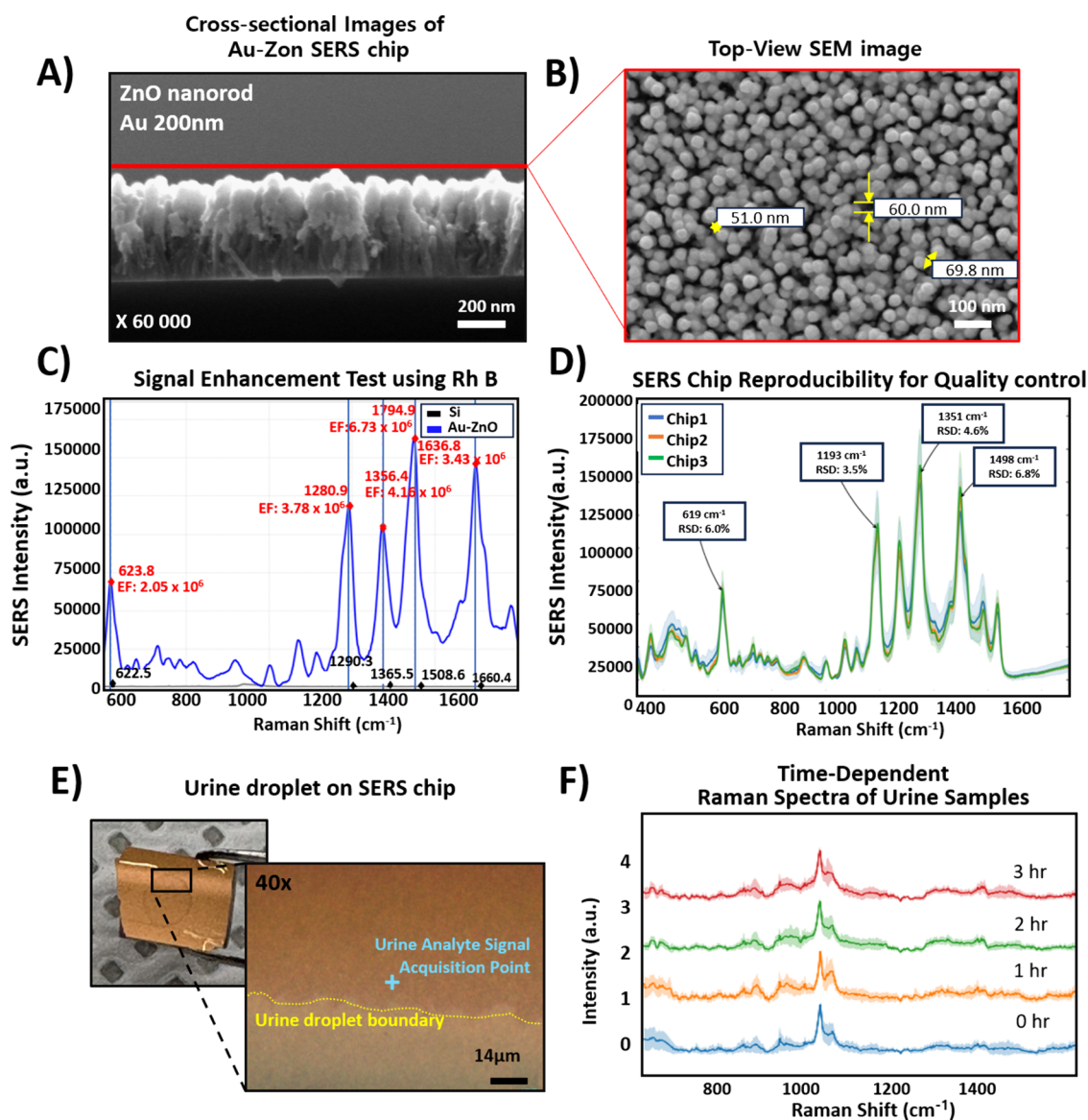


Figure 1. Raman signal acquisition test using the Au-ZnO nanorod SERS chip. (A) Cross-sectional image of 200 nm gold coating on the ZnO SERS chip using SEM ($\times 60,000$) for morphological confirmation. (B) Top view of the Au-ZnO SERS chip with representative inter-nano gap widths of 51.0, 60.0, and 69.8 nm ($\times 5000$; scale bar, 100 nm). (C) Comparison of Rhodamine B Raman spectra obtained on the SERS chip and bare Si wafer visualizing the signal enhancement. (D) Chip-to-chip reproducibility test using three Au-ZnO SERS chips. The relative standard deviations (RSD) are at 619, 1193, 1351, and 1498 cm^{-1} with 6.0%, 3.5%, 4.6%, and 6.8%, respectively. (E) Urine droplet boundary with the urine analyte signal acquisition point on the 40 $\times$ magnified image of the SERS chip. (F) Repeated SERS spectra (0–3 h, room temperature) of a single diabetic urine sample excluded from the diagnostic cohort. The spectra remain consistent over 0–3 h, confirming short-term signal stability. This panel demonstrates temporal stability only and is not directly comparable with the cohort-averaged spectra.

with the average relative standard deviation (RSD) of 5.2%, as shown in Figure 1D. We dropped 5 μL of urine samples on the SERS chip for Raman signal amplification by comparing the signal same condition on the Si wafer in Figure S2A for the bright-field image of urine drop and Figure S2B for signal enhancement with 1.18×10^4 . Finally, we measured urine analytes as indicated with a cross mark beyond the droplet boundary, as shown in Figure 1E. Raman spectra with high intensity were more frequently obtained when the analyte was measured from 5 to 10 μm beyond the boundary of the sample droplet.¹² To further characterize this measurement region, we performed 2D Raman mapping across the dried urine region (Figure S2C). High-intensity acquisition points were broadly distributed throughout the mapped area but occurred more frequently near the droplet boundary (Figure S2D). This boundary-associated enrichment may be related to the coffee-ring effect, which can concentrate analytes during droplet drying and thereby enhance local SERS signals. This effect has also been exploited for SERS detection of phenylalanine in urine.¹³ Therefore, the region approximately 5–10 μm beyond the dried droplet boundary consistently used as a standardized measurement zone. Storage-time testing at room temperature (0–3 h; Figure 1F) showed no significant change in the Raman signal over time—only a slightly larger standard deviation at 0 h near 600 and 1400 cm^{-1} . Spectra were acquired within 1–3 h of urine collection to ensure spectral stability and reproducibility.

Raman Spectra Profile and Intensity Patterns of Urine Samples from Healthy Control, IC, and OAB Groups

By comparing the full Raman spectrum between 600 and 1460 cm^{-1} , which is within the biological fingerprint range, the spectra showed a similar spectral profile, but certain Raman shift regions showed intensities that are different between the control group and the patient groups, as shown in Figure 2A. Especially, the Raman shift near 1000 cm^{-1} can be associated with phenylalanine and also 900–1000 and 1200–1400 cm^{-1} regions, which may correspond to protein-, nucleic acid-, amide-, and CH/COO^- -related molecular vibrations. These different patterns within the Raman spectra reflect the biological composition within the urine sample depending on the pathological conditions. From a previous study, IC significant Raman spectra for diagnosis are related with phenylalanine signal.¹⁴ For quantitative evaluation, major Raman intensities from major Raman peaks 1135 cm^{-1} (C–N/C–C stretching, protein-related vibration)^{15,16} and 1220 cm^{-1} (Amide III, protein backbone)¹⁷ were compared between groups as violin plots shown in Figure 2B,C, respectively, and both are significantly different (P values $*** < 0.001$ and $* < 0.05$). The corresponding effect sizes were medium for the disease-versus-control comparisons (Cohen's $d = 0.49$ – 0.70), and these differences can serve as IC- or OAB-related molecular signatures. The IC–OAB difference at these two bands was small ($d = 0.23$), as these protein-backbone vibrations may be shared by both conditions. To visualize Raman spectral

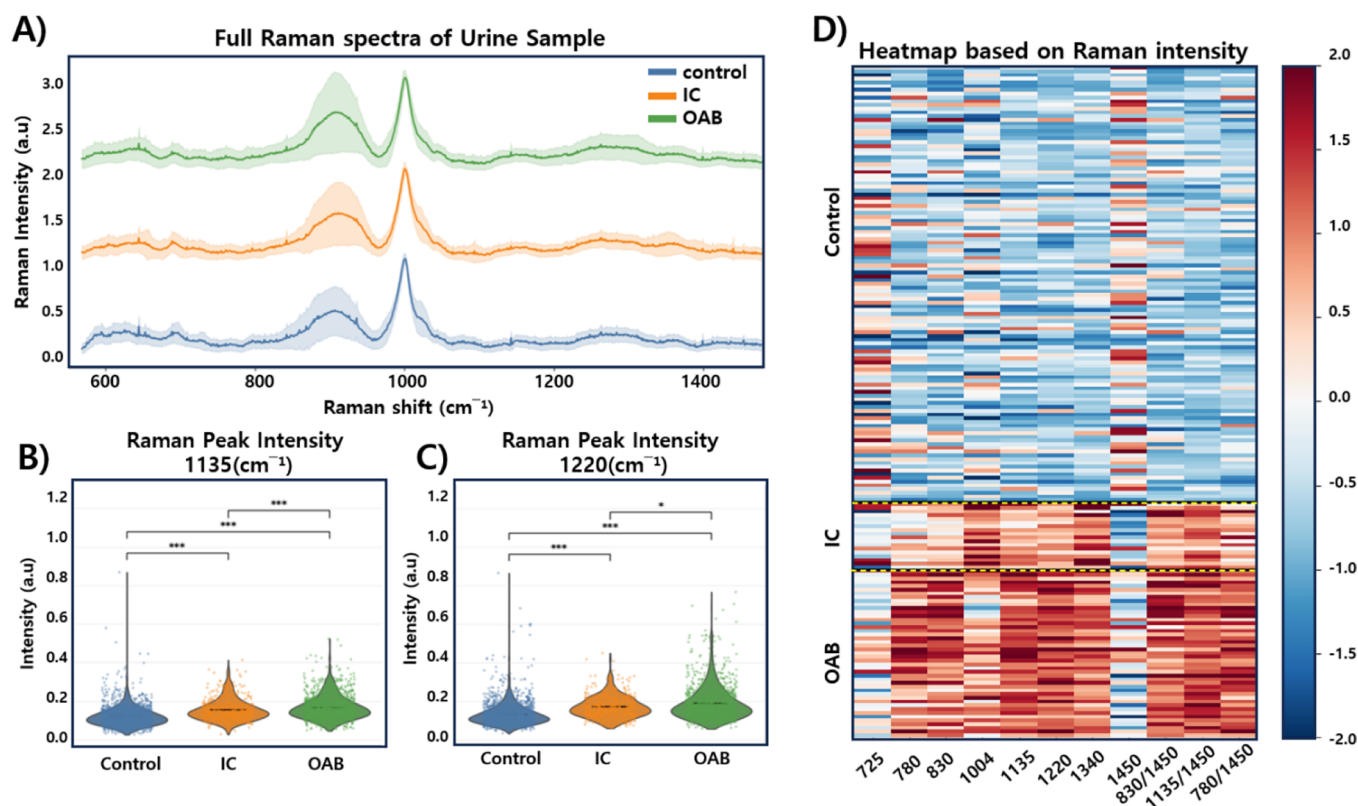


Figure 2. Raman spectral profiling and intensity pattern of urine samples from healthy control group, IC, and OAB patients. (A) Full Raman spectra with the solid average line with the shaded standard deviation area. Raman peak intensity violin plots (B) 1135 (cm^{-1}) and (C) 1220 (cm^{-1}) showing the intensity difference. (D) Heatmap based on Raman intensity visualizing disease-related intensity patterns that are different from the healthy control group.

intensity patterns, a Raman intensity-based heatmap was constructed, as shown in Figure 2D. The heatmap clearly showed the Raman intensity pattern differences between the control group and patient groups IC and OAB. These Raman shifts showed elevation or reduction patterns between control and patients. Samples from IC and OAB share similar intensity patterns compared to control. These visualizations of the heatmap reflect that urine SERS-Raman spectra contain distinguishable spectral differences among the control and patient groups. The relative assignment of major Raman shifts for heatmap analysis is listed in Table S2.

Raman Spectral Comparison between the IC and OAB Groups

To seek differential potential between IC and OAB, the mean Raman spectra were directly compared between those two groups. As shown in Figure 3A, IC and OAB showed broadly similar spectral profiles reflecting their partially overlapping urinary molecular composition. However, distinct intensity differences were observed at specific Raman shift regions. To visualize the specific Raman intensity differences, a difference spectrum was calculated from the mean Raman spectra of IC and OAB, as shown in Figure 3B. This result possesses that the possibility of Raman shift regions that showed relatively higher intensity in either IC or OAB can be disease discrimination features. Based on the previously reported Raman assignment, the increased region of Raman spectra in IC may be tentatively associated with nucleic acid-related ring vibration, phenylalanine, and CH₃ deformation molecular vibrations. On

the other hand, cytosine/uracil, protein-related C–C stretching, amide III/protein-related signals, and CH vibration/lipid contribution-related Raman signal were increased in the OAB group. Although these molecular assignments should be interpreted as tentative, they support the possibility that IC and OAB have distinct urinary molecular fingerprints associated with inflammatory responses, urothelial alterations, protein composition, and lipid-related molecular changes. Raman features may further serve as candidate spectral markers for subsequent machine learning-based classification.

Integrated Raman Feature and Group-Dependent Intensity Pattern Using Multiple Machine Learning Models

Finally, top Raman features were extracted from PCA, XGBoost, and LightGBM and ranked according to the contributing order. In Figure 4A, adjacent wavenumbers within a single Raman band are strongly correlated; this per-wavenumber ranking lists several neighboring points of the same peak (e.g., 1005–1011 cm^{−1}, phenylalanine) as separate top features and therefore serves only as an intermediate feature-localization step rather than as the basis of biological interpretation. The biological interpretation was instead performed at the Raman-band level, using the consensus features shared across the PCA, XGBoost, and LightGBM models. Therefore, adjacent points of the same peak are collectively treated as a single spectral signature rather than as separate biomarkers. In Figure 4B,C, important Raman features based on nonlinear relation and a tree-based decision were listed. Although the feature rankings differed among the classification models, several Raman shifts

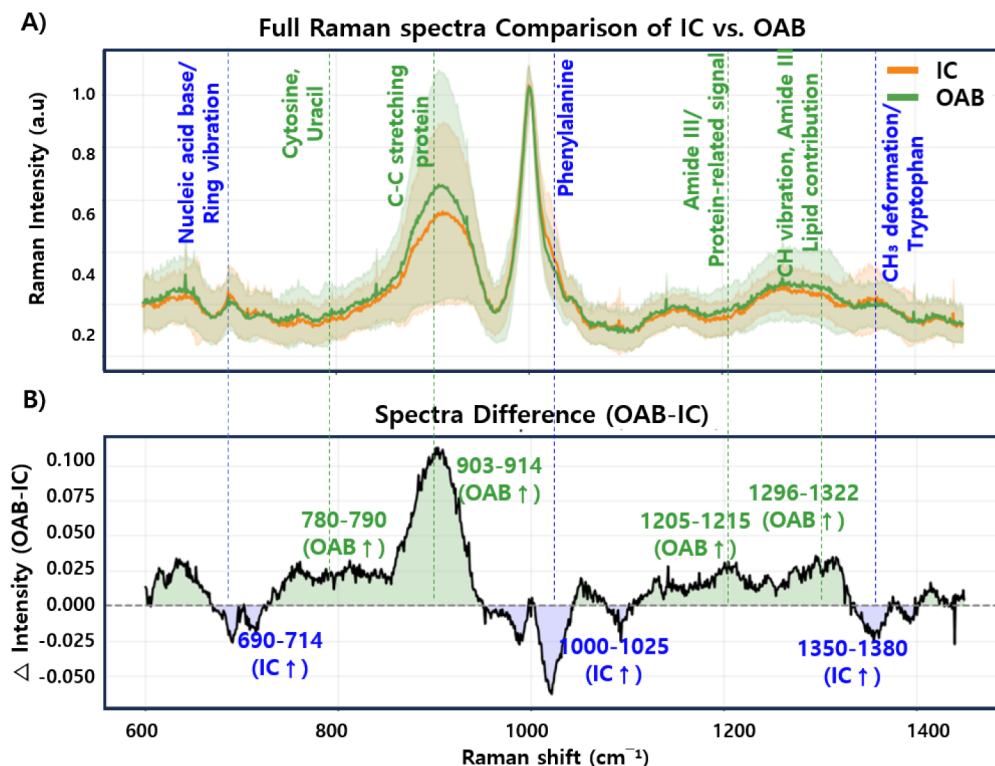


Figure 3. Differential spectral features between patient groups IC and OAB. (A) Full spectra of IC and OAB with indication of tentative biochemical assignments of selected Raman shift regions. The solid line indicates the average of full Raman spectra, and the shaded area stands for standard deviation. (B) Subtracted intensity from OAB to IC showing Raman shift regions relatively high or low compared to each group. Green-labeled regions indicate Raman features that increased in OAB, while blue-labeled regions indicate increase in IC. This suggests that the IC and OAB spectra possess distinct urinary Raman spectra.

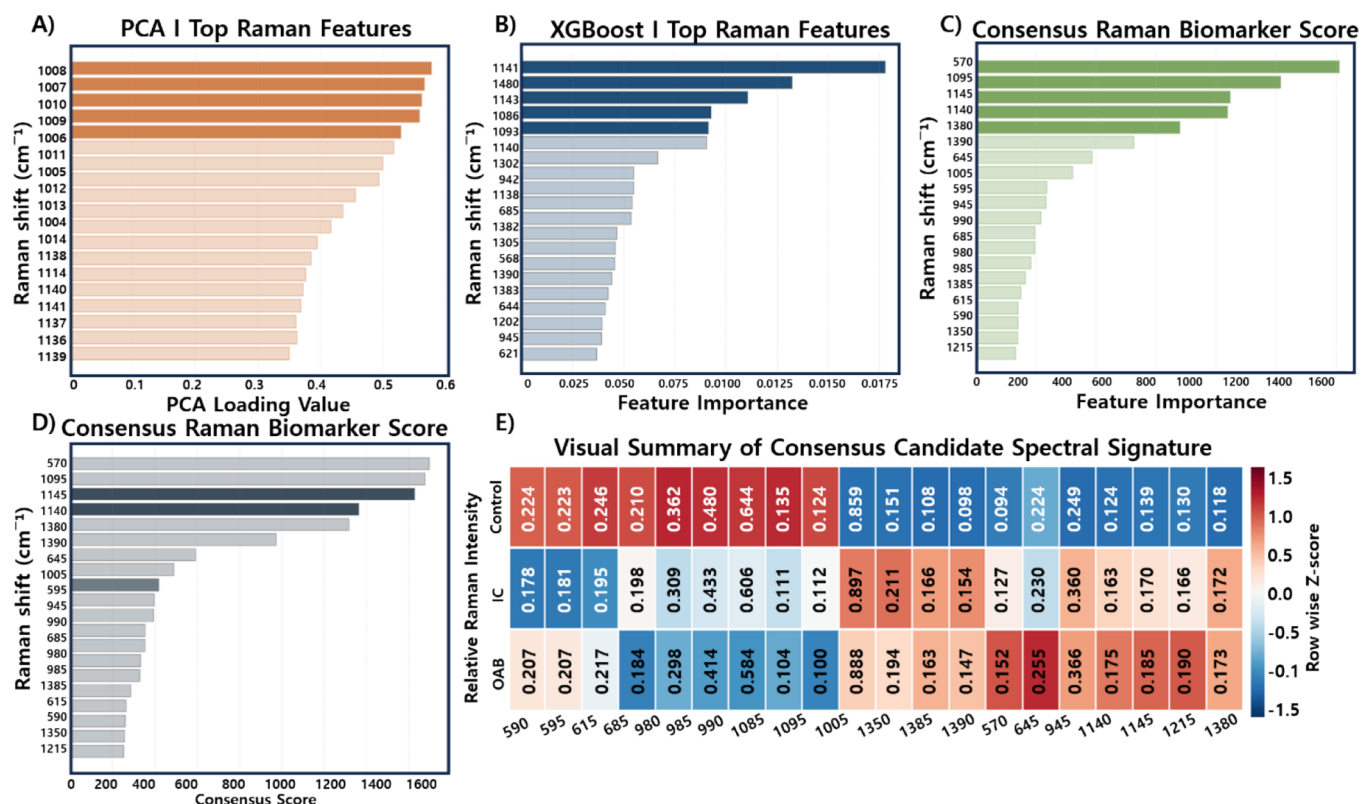


Figure 4. Integrated model-derived Raman biomarker candidates and the summary of relative intensity pattern. Top feature derived from multivariate statistical analysis. (A) PCA, machine learning-derived top feature of (B) XGBoost. (C) LightGBM, integrated Raman biomarker from the (D) multiple analytical model. (E) Visual summary of integrated Raman biomarker candidate intensity in each group. These patterns support the evidence of machine learning-based discrimination and suggest that this elevation or reduction of selected Raman shifts may serve as a spectral marker for distinguishing bladder disease subtypes.

were commonly identified as important discriminative features. This multimodel feature interpretation method is designed based on the previous Raman spectroscopy studies, which highlight the importance of multiple feature selection approaches for explainable machine learning to identify interpretable spectral markers.¹⁸ Thus, the integrated common Raman shifts extracted from all models are listed in Figure 4D. To further characterize the feature intensity elevation or reduction, Raman intensities were compared between groups using those features listed in Figure 4D. The visual summary of these feature elevation or reduction is shown in Figure 4E, and the relative Raman assignments with *P* value are summarized in Table S3. This Raman intensity comparison showed that many Raman features differed significantly between the control and disease groups, while only selected features showed significant differences between IC and OAB. These results indicate that model-derived common Raman spectral features capture group-dependent spectral differences among control, IC, and OAB urine samples.

Multiple Machine Learning Model Application for Development of Diagnosis

To compare the classification ability of linear and nonlinear-based classification, PCA-PLS-DA, PCA-LDA, XGBoost, and LightGBM were applied. For PCA, 100 principal components were retained, which accounted for 88.4% of the total spectral variance, as shown in Figure S3A. Cross-validated classification accuracy increased with the number of retained components up

to approximately 100 and then plateaued, with no appreciable further improvement (Figure S3B). The scatter distribution of the three groups was visualized using the PCA-PLS-DA-based score plot in Figure 5A in 2D and Figure 5B in 3D space. Even though some IC and OAB showed partial overlap, the overall tendency of cluster separation is in different directions. To reach more clear separation between the groups, LightGBM was applied to evaluate classification performance. In the confusion matrix of LightGBM, controls and OAB were correctly classified as predicted label to true label, as shown in Figure 5C. For IC, 76 samples were correctly classified as IC whereas some IC samples were misclassified as OAB or controls. However, the LightGBM model demonstrated strong overall classification performance. ROC curve analysis showed AUC values of 0.99, 0.98, and 0.98 for healthy control, IC, and OAB, respectively, with a micro-average AUC of 0.992 shown in Figure 5D. Confusion matrices for the rest of the machine learning model are provided in Figure S4. Compared with linear models such as PCA-PLS-DA and PCA-LDA, tree-based nonlinear models, including XGBoost and LightGBM, showed superior classification performance. Five-cross-validation was performed for all classification models, and the results are presented in a bar graph in Figure S5 showing robustness of each model. As a negative control, repeating the identical pipeline with randomly permuted class labels reduced accuracy to chance level (0.47 ± 0.01 ; permutation $p = 0.03$), confirming that the classification performance was not an artifact of preprocessing or feature extraction. To

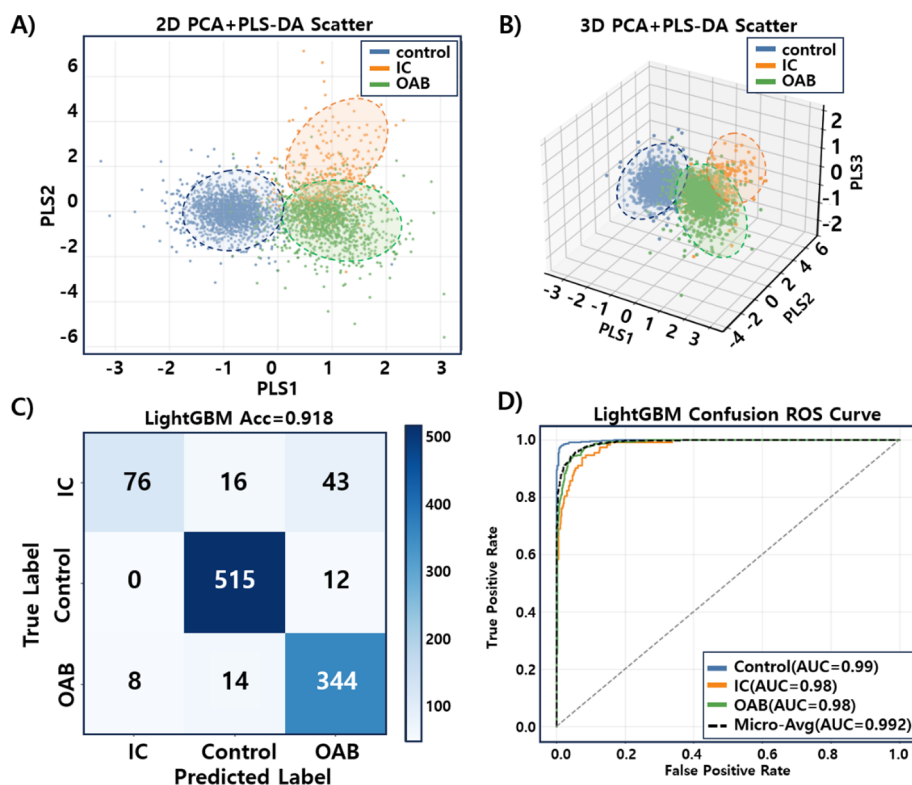


Figure 5. Machine learning-based classification summary using PCA-PLS-DA and LightGBM. The distribution of Raman spectra from healthy control, IC, and OAB are displayed in (A) 2D PCA-PLS-DA and (B) 3D-PCA-PLS-DA score plots. All group shows visual tendency of clear separation from each other. (C) Confusion matrix of LightGBM and (D) ROC curve showing that these groups are effectively discriminated.

further assess generalization to unseen individuals, all four models were additionally evaluated using participant-level cross-validation (Stratified GroupK Fold). Overall discrimination remained high across all models (participant-level macro AUC 0.90–0.94; Figure S6).

DISCUSSION

Since both IC and OAB are clinically difficult to distinguish due to the overlapping lower urinary tract symptoms with urinary urgency and frequency,¹⁹ a novel noninvasive diagnostic method can complement the symptom and clinical observation-based diagnosis. Urine volume, drying time, acquisition conditions, and spectral preprocessing were standardized, and first-morning midstream urine samples were collected between 6:00 and 8:00 a.m. after an overnight fast of at least 12 h. However, urine specific gravity, osmolality, and creatinine were not measured. Therefore, normalization to the phenylalanine band at approximately 1000–1004 cm^{-1} was performed to reduce intensity variation.

First, comparison of the full Raman spectra revealed that all groups showed an intensity difference especially between control and patient groups IC and OAB. The Raman intensity pattern difference in controls and patient groups was shown mostly in protein- and amino acid-related Raman regions, including the phenylalanine/protein-related band at 1004 cm^{-1} , C–N/C–C stretching at 1135 cm^{-1} , and amide III/protein backbone signal at 1220 cm^{-1} . This reflects that SERS-Raman spectra detect the molecular Raman signature associated with bladder disease status. For IC and OAB direct comparison, IC showed relatively

high intensities at 690–714, 1000–1025, and 1350–1380 cm^{-1} . For OAB, the intensity of Raman spectra at 780–790, 903–914, 1205–1215, and 1296–1322 cm^{-1} were higher than for IC. These differences indicate that IC and OAB may have distinct urinary molecular signatures despite the overlapping clinical symptoms and possibility of signature separation for diagnostic purposes.

Raman spectra of these groups were then extracted from PCA, XGBoost, and LightGBM to select Raman shifts that were repeatedly identified as important features from multiple models for disease-associated spectral features. Therefore, Raman shifts that are commonly identified as important features across multiple models should have greater reliability as candidate disease-associated spectral markers rather than a single algorithm.^{20,21} The integrated features identified for control were 590, 595, 615, 685, 980–990, and 1085–1095 cm^{-1} , and these Raman shifts are associated with protein/lipid skeletal vibrations, C–C skeletal modes, nucleic acid-related signals, and PO_2^- /nucleic acid vibrations.

These Raman shifts can be the baseline molecular background of normal urine, in contrast to inflammatory or protein-related alterations that increased in the disease groups. Regarding the disease associated band, the band near 1005 cm^{-1} is assigned to phenylalanine/protein-related vibration, while the 1350–1390 cm^{-1} region may be associated with CH deformation and CH/COO[−] vibration. From previous Raman studies, a phenylalanine-associated band was reported in control and mild IC samples, while a tryptophan-associated band was reported in severe IC samples.¹⁴ This suggests that these Raman features may vary with the IC phenotype or

symptom severity. In addition, phenylalanine has been reported as an IC-discriminating metabolite in urinary metabolomic analysis, although its urinary level was lower, rather than higher, in IC patients than in nondisease controls.²² Therefore, the increased SERS intensity at the phenylalanine-associated band should not be interpreted as a direct increase in urinary phenylalanine concentration. Alterations in urinary metabolites have also been reported in IC/BPS patients.²³ Thus, the increased Raman bands at 1005 and 1350 cm^{-1} may reflect disease-associated alterations in aromatic amino acid- and protein-related molecular composition, rather than a direct increase in urinary phenylalanine or tryptophan concentration. Similarly, the 1385–1390 cm^{-1} region, associated with broader CH_3/COO^- -related vibrations, may reflect compositional changes accompanying urothelial dysfunction and inflammation rather than a specific inflammatory protein.

On the other hand, high Raman spectral intensities identified from OAB patients were at 570, 645, 945, 1140–1145, 1215, and 1380 cm^{-1} . The Raman features are mainly associated with nucleic acid, hemoglobin, protein backbone, C–N/C–C stretching, amide III/protein-related signals, and CH/COO^- vibration. OAB-associated Raman structural patterns may reflect the altered protein structural and inflammatory-related molecular increase. In OAB patients, the elevation of inflammatory or protein-related urinary biomarkers such as MCP-1 and VCAM-1 has been reported in previous studies.^{8,24} The protein, amide, and lipid-related Raman features' intensities observed in OAB may reflect urinary molecular composition changes associated with these pathophysiological alterations. However, these Raman bands may contain overlapping contributions from proteins, amino acids, nucleic acids, lipids, and metabolites and therefore cannot be directly assigned to specific biomarkers based on label-free SERS alone. The observed features are thus interpreted as candidate OAB-associated spectral regions that may reflect broader molecular changes related to OAB pathophysiology. Previous studies have reported alterations in urinary proteins and inflammatory mediators in OAB, including NGF, BDNF,²⁵ MCP-1, and distinct cytokine profiles,²⁶ supporting the presence of broader protein- and inflammation-associated molecular changes in OAB. Overall, the discriminative Raman bands identified for IC and OAB should be regarded as candidate disease-associated spectral signatures rather than definitive molecular biomarkers. Definitive identification and quantification of the underlying molecular contributors will require complementary targeted analyses, such as LC-MS/MS for metabolites and amino acids and immunoassays for inflammatory or urothelial proteins, in future studies.

Machine learning-based classification analyses were performed to evaluate the diagnostic potential of urine Raman spectra. In the model performance comparison, PCA-PLS-DA and PCA-LDA showed accuracies of 0.855 and 0.879, respectively. By using tree-based decision models such as XGBoost and LightGBM, higher accuracies of 0.909 and 0.918 were achieved. Although biological SERS spectra show large sample-to-sample variance, multivariate machine-learning models still achieve high diagnostic accuracy by exploiting subtle but consistent covariance patterns across many wavenumbers rather than variance visible in individual spectra, as demonstrated for exosome-SERS cancer diagnosis (AUC 0.970) and serum-SERS diagnosis of immune and kidney diseases.^{27,28} IC was frequently predicted as OAB while OAB was reliably identified. This may reflect the clinical overlap of

certain IC phenotypes with OAB rather than class imbalance alone, as this pattern persisted under class-weighted analysis. The misclassified IC cases show that the errors were not random but reflected clinically overlapping, borderline cases. Patients misclassified as OAB were predominantly non-Hunner-lesion IC or IC with comorbid OAB whose clinical presentation overlaps with OAB, whereas those misclassified as healthy controls were mostly elderly patients with diagnostically confounding features (e.g., age-related changes in pain perception or anti-inflammatory medication). Because VAS pain scores did not differ between correctly and incorrectly classified patients, these errors reflect the clinical and molecular overlap of these IC subtypes with OAB, with IC frequently classified as OAB while OAB itself was reliably identified. Although these observations are limited by the relatively small number of IC patients, which constrains statistical power and generalizability, they provide a clinically coherent account of the misclassifications and warrant confirmation in larger cohorts. Building on these results, external validation in larger, independent cohorts with participant-level analysis will be essential to establishing clinical generalizability and to translating this urine SERS–machine learning approach into a clinically applicable diagnostic tool.

CONCLUSIONS

SERS-Raman spectroscopy with machine learning-based analysis of urine can distinguish bladder diseases such as IC and OAB from healthy controls. By applying multiple machine learning models such as PCA-PLS-DA, PCA-LDA, XGBoost, and LightGBM, integrated Raman features appeared in group-dependent Raman intensities. The identified candidate Raman spectra were distinct between IC and OAB despite their similar clinical symptoms. A tree-based boosting model such as LightGBM supports the diagnostic approach for urine Raman spectra. These findings suggest that urine SERS-Raman spectroscopy with the combination of machine learning can provide a noninvasive diagnostic aid for bladder diseases.

EXPERIMENTAL SECTION/METHODS

Study Design and Urine Sample Acquisition

The study protocol was approved by the IRB of Asan Medical Center (IRB No. 2022-1416). Urine samples were acquired from study participants, categorized into the following groups: healthy control (Normal), interstitial cystitis (IC), and overactive bladder (OAB). The healthy control group included those who visited the general health examination center without any history of urological diagnosis and who agreed on use of their samples for any research purposes. Diagnosis of IC and OAB was based on definition by the International Continence Society. In accordance with the IRB-approved participant recruitment and sample-collection protocol, first-morning midstream urine samples were collected between 6:00 and 8:00 a.m. after an overnight fast of at least 12 h. The sample was analyzed within 3 h following urine collection.

Preparation of ZnO Nanorod on the SERS Chip

ZnO nanorods (~500 nm) were grown on a Si wafer by hydrothermal synthesis and coated with a 200 nm Au layer by sputtering. Nanorod growth and Au coating were confirmed by SEM (FE-SEM, JEOL). Detailed fabrication conditions are provided in the [Supporting Information](#).

Raman Signal Acquisition

Using a sensing chip with Au-coated ZnO nanostructures on the surface, the Raman signal of urinary molecular signatures in urine was amplified. A total of 5 μL of urine drop at room temperature was placed on SERS with the Au-ZnO nanorod for 30 min before Raman acquisition, and measurements were taken using an upright Raman spectroscopy system (FEX-INV, WEVE, Seoul, South Korea) equipped with a 785 nm excitation diode laser operating at an output power of 10 mW,²⁹ with 1200 grooves/mm gratings; the center value was set at 1090 cm^{-1} , an ND filter at 1%, and the pinhole at 100 μm using a 40 $\times$ lens (LUCPLFLN, OLYMPUS, Japan). The exposure time was set to 10 s, with an accumulation number of 5. After acquiring the Raman spectrum of the urine sample (10–30 spectra per participant), the Raman signals underwent polynomial fitting and noise removal using the Raon-Vu program for further data processing. Prior to multivariate analysis, each spectrum was baseline-corrected using a fifth-order polynomial fit and normalized to the mean intensity of the phenylalanine band ($\sim 1000\text{--}1004\text{ cm}^{-1}$).

Machine Learning Classification and Feature Interpretation

Four classifiers were compared—the linear PCA-PLS-DA and PCA-LDA and the nonlinear tree-based XGBoost and LightGBM—implemented in Python (scikit-learn, xgboost, and lightgbm). Spectra were standardized and, for PCA-based models, reduced to 100 principal components. These steps were fitted only on the training partition of each split to prevent information leakage. Data were split into training and test sets (7:3, stratified) and were additionally evaluated by stratified fivefold and participant-wise (StratifiedGroupKFold) cross-validation. Model hyperparameters were fixed a priori. Performance was assessed by balanced accuracy, confusion matrices, and ROC curves. For biological interpretation, feature importances were extracted from each model, and the top 30 Raman shifts commonly identified across models were compared by groupwise intensity to identify consensus disease-associated features. Full model configurations and hyperparameters are provided in the [Supporting Information](#).

Statistical Analysis

For statistical analysis, pairwise Mann–Whitney U tests were performed to compare selected Raman features and ratio features among groups.³⁰ The p -values were adjusted for multiple comparisons using the Benjamini–Hochberg false discovery rate method.³¹ Statistical significance was determined based on adjusted p -values, and the significance levels were indicated as n.s. not significant, $* < 0.05$, $** < 0.01$, and $*** < 0.001$. Group comparisons were performed on 117 control, 19 IC, and 45 OAB participants (10–30 spectra per participant), and the significance threshold was set at $\alpha = 0.05$ (Benjamini–Hochberg adjusted). Effect sizes were quantified as Cohen's d with 95% bootstrap confidence intervals (concordant with the nonparametric rank-biserial correlation). Reproducibility of the control baseline was assessed across the 117 healthy controls: the control-dominant bands showed moderate interindividual variability ($\text{CV} = 23\%$) with high within-participant reproducibility (average ICC = 0.90), and their intensities were not associated with age or sex (Spearman $|r| \leq 0.11$; sex: rank-biserial $r = -0.23$, Benjamini–Hochberg FDR $p > 0.05$).

■ ASSOCIATED CONTENT

Data Availability Statement

Data will be made available on request.

SI Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acssensors.6c02725>.

Experimental methods, participant characteristics, supplementary figures, supplementary tables, and additional machine-learning results (DOCX)

■ AUTHOR INFORMATION

Corresponding Authors

Jung Hyun Shin – Department of Urology, Ewha Womans University Mokdong Hospital, Seoul 07985, Korea; Email: jshinuro@ewha.ac.kr

Jun Ki Kim – Department of Convergence Medicine, BK 21 Project, University of Ulsan, College of Medicine, Seoul 05505, Republic of Korea; Apollon, Inc., Seoul 07797, Republic of Korea; Biomedical Engineering Research Center, Asan Medical Center, Seoul 05505, Republic of Korea; Institute of Digital Healthcare, University of Ulsan College of Medicine, Seoul 05505, Republic of Korea;  orcid.org/0000-0002-0099-9681; Email: kim@amc.seoul.kr

Authors

Minju Cho – Department of Convergence Medicine, BK 21 Project, University of Ulsan, College of Medicine, Seoul 05505, Republic of Korea

Suyeon Kang – Department of Convergence Medicine, BK 21 Project, University of Ulsan, College of Medicine, Seoul 05505, Republic of Korea

Joon Seup Hwang – Department of Convergence Medicine, BK 21 Project, University of Ulsan, College of Medicine, Seoul 05505, Republic of Korea

Miyeon Jue – Apollon, Inc., Seoul 07797, Republic of Korea;  orcid.org/0009-0006-0260-1258

Complete contact information is available at:

<https://pubs.acs.org/doi/10.1021/acssensors.6c02725>

Author Contributions

[†]M.C. and S.K. contributed equally to this work. CRediT Authorship Contribution Statement Minju Cho: writing-original draft, visualization, validation, software, methodology, formal analysis, and data curation. **Suyeon Kang**, **Joon Seup Hwang**, and **Miyeon Jue**: writing-original draft, validation, resources, software, methodology, formal analysis, and data curation. **Jung Hyun Shin**: conceptualization and writing-original draft. **Jun Ki Kim**: conceptualization, writing-original draft, validation, supervision, and funding acquisition. Writing-reviewing and editing: all authors reviewed and approved the final version of the manuscript.

Funding

This work was supported by the National Research Foundation of Korea (NRF) funded by the Korean government (MSIT) (RS-2024-00450201 and RS-2025-25407039) and the Korea Health Industry Development Institute (KHIDI), funded by the Ministry of Health & Welfare, Republic of Korea (RS-2026-25516119 and RS-2026-25547118).

Notes

No generative AI or AI-assisted technologies were used in the writing, editing, or content generation of this manuscript. All text was written and reviewed solely by the authors.

J.K.K. and M.J. serve as the scientific advisor and Chief Technology Officer, respectively, at Apollon Inc. (Korea), a

startup that may develop diagnostic devices based on this work. The remaining authors declare no other competing interests.

ACKNOWLEDGMENTS

We gratefully acknowledge Dr. Kyung Yoon Soo of the Health Screening and Promotion Center at Asan Medical Center for his support in sample collection.

REFERENCES

- (1) Hanno, P.; Dmochowski, R. Status of international consensus on interstitial cystitis/bladder pain syndrome/painful bladder syndrome: 2008 snapshot. *Neurourol. Urodyn.* **2009**, *28* (4), 274–286.
- (2) Clemens, J. Q.; Erickson, D. R.; Lai, H. H. Diagnosis and Treatment of Interstitial Cystitis/Bladder Pain Syndrome Reply. *J. Urol.* **2022**, *208* (6), 1178–1179.
- (3) Forrest, J. B.; Schmidt, S. Interstitial cystitis, chronic nonbacterial prostatitis and chronic pelvic pain syndrome in men: a common and frequently identical clinical entity. *J. Urol.* **2004**, *172* (6 Pt 2), 2561–2562.
- (4) Arora, H. C.; Shoskes, D. A. The enigma of men with interstitial cystitis/bladder pain syndrome. *Transl. Androl. Urol.* **2015**, *4* (6), 668–676.
- (5) Haylen, B. T.; de Ridder, D.; Freeman, R. M.; Swift, S. E.; Berghmans, B.; Lee, J.; Monga, A.; Petri, E.; Rizk, D. E.; Sand, P. K.; et al. An International Urogynecological Association (IUGA)/International Continence Society (ICS) joint report on the terminology for female pelvic floor dysfunction. *Int. Urogynecol. J.* **2010**, *21* (1), 5–26.
- (6) Billecocq, S.; Bo, K.; Dumoulin, C.; Aigon, A.; Amarenco, G.; Bakker, E.; Cornillet-Bernard, M.; Cretinon, S.; Deffieux, X.; Lartigues, G.; et al. An International Urogynecological Association (IUGA)/International Continence Society (ICS) joint report on the terminology for the conservative and non-pharmacological management of female pelvic floor dysfunction. *Prog. Urol.* **2019**, *29* (4), 183–208.
- (7) Gregoire, M.; Liandier, F.; Naud, A.; Lacombe, L.; Fradet, Y. Does the potassium stimulation test predict cystometric, cystoscopic outcome in interstitial cystitis. *J. Urol.* **2002**, *168* (2), 556–557.
- (8) Futyma, K.; Nowakowski, L.; Zietek-Strobl, A.; Kaminska, A.; Taoussi, N.; Rechberger, T. Urine Proteomic Study in OAB Patients-Preliminary Report. *J. Clin. Med.* **2020**, *9* (5), No. 1389.
- (9) Jue, M.; Lee, S.; Paulson, B.; Namgoong, J. M.; Yu, H. Y.; Kim, G.; Jeon, S.; Shin, D. M.; Choo, M. S.; Joo, J.; et al. Optimization of ZnO Nanorod-Based Surface Enhanced Raman Scattering Substrates for Bio-Applications. *Nanomaterials* **2019**, *9* (3), No. 447.
- (10) Zhou, Z.; Hooker, G. Unbiased measurement of feature importance in tree-based methods. *arXiv preprint arXiv* **2021**, pp 1–21.
- (11) Chuang, F. C.; Hsiao, S. M.; Kuo, H. C. The Overactive Bladder Symptom Score, International Prostate Symptom Score-Storage Subscore, and Urgency Severity Score in Patients With Overactive Bladder and Hypersensitive Bladder: Which Scoring System is Best? *Int. Neurourol. J.* **2018**, *22* (2), 99–106.
- (12) Kim, M. G.; Jue, M.; Lee, K. H.; Lee, E. Y.; Roh, Y.; Lee, M.; Lee, H. J.; Lee, S.; Liu, H.; Koo, B.; Jang, Y. O.; Kim, E. Y.; Zhen, Q.; Kim, S.-H.; Kim, J. K.; Shin, Y. Deep learning assisted surface-enhanced Raman spectroscopy (SERS) for rapid and direct nucleic acid amplification and detection: Toward enhanced molecular diagnostics. *ACS Nano* **2023**, *17*, 18332–18345.
- (13) Murugesan, B.; Yang, J. Tunable Coffee Ring Formation on Polycarbonate Nanofiber Film for Sensitive SERS Detection of Phenylalanine in Urine. *ACS Omega* **2019**, *4* (12), 14928–14936.
- (14) Hsieh, T. F.; Yu, K. J.; Lin, S. Y. Possible application of Raman microspectroscopy to verify the interstitial cystitis diagnosis after potassium sensitivity test: phenylalanine or tryptophan as a biomarker. *Dis. Markers* **2007**, *23* (3), 147–152.
- (15) Lukose, J.; Barik, A. K.; Mithun, N.; Sanoop Pavithran, M.; George, S. D.; Murukeshan, V. M.; Chidangil, S. Raman spectroscopy for viral diagnostics. *Biophys. Rev.* **2023**, *15* (2), 199–221.
- (16) Adigal, S. S.; Bhandary, S. V.; Kuzhupilly, N. I. R.; George, S. D.; Kartha, V. B.; Chidangil, S. Analysis of tear fluid protein fractions using surface-enhanced Raman spectroscopy (SERS), followed by a high-performance liquid chromatography-light emitting diode-induced fluorescence (HPLC-LED-IF) method. *RSC Adv.* **2025**, *15* (50), 42728–42739.
- (17) Stone, N.; Kendall, C.; Smith, J.; Crow, P.; Barr, H. Raman spectroscopy for identification of epithelial cancers. *Faraday Discuss.* **2004**, *126*, 141–157.
- (18) Rossberg, N.; Gautam, R.; Komolibus, K.; O'Sullivan, B.; Visentin, A. Explainable AI-Based Feature Selection Approaches for Raman Spectroscopy. *Diagnostics* **2025**, *15* (16), No. 2063.
- (19) Macdiarmid, S. A.; Sand, P. K. Diagnosis of interstitial cystitis/painful bladder syndrome in patients with overactive bladder symptoms. *Rev. Urol.* **2007**, *9* (1), 9–16.
- (20) Abeel, T.; Helleputte, T.; Van de Peer, Y.; Dupont, P.; Saeys, Y. Robust biomarker identification for cancer diagnosis with ensemble feature selection methods. *Bioinformatics* **2010**, *26* (3), 392–398.
- (21) Shahrjooiaghghi, A.; Frigui, H.; Zhang, X.; Wei, X.; Shi, B.; Trabelsi, A. An Ensemble Feature Selection Method for Biomarker Discovery. In *2017 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT)*; IEEE **2017**, 416–421.
- (22) Kim, J.; De Hoedt, A.; Wiggins, E.; Haywood, K.; Jin, P.; Greenwood, B.; Narain, N. R.; Tolstikov, V.; Bussberg, V.; Barbour, K. E.; et al. Diagnostic Utility of Serum and Urinary Metabolite Analysis in Patients with Interstitial Cystitis/Painful Bladder Syndrome. *Urology* **2021**, *157*, 85–92.
- (23) Fiehn, O.; Kim, J. Metabolomics insights into pathophysiological mechanisms of interstitial cystitis. *Int. Neurourol. J.* **2014**, *18* (3), 106–114.
- (24) Farhan, B.; Chang, H.; Ahmed, A.; Zaldivair, F.; Ghoniem, G. Characterisation of urinary monocyte chemoattractant protein 1: Potential biomarker for patients with overactive bladder. *Arab. J. Urol.* **2019**, *17* (1), 58–60.
- (25) Rachaneni, S.; Arya, P.; Latthe, P. Urinary nerve growth factor: a biomarker of detrusor overactivity? A systematic review. *Int. Urogynecol. J.* **2013**, *24* (10), 1603–1609.
- (26) Jiang, Y. H.; Jhang, J. F.; Hsu, Y. H.; Ho, H. C.; Wu, Y. H.; Kuo, H. C. Urine biomarkers in ESSIC type 2 interstitial cystitis/bladder pain syndrome and overactive bladder with developing a novel diagnostic algorithm. *Sci. Rep.* **2021**, *11* (1), No. 914.
- (27) Shin, H.; Choi, B. H.; Shim, O.; Kim, J.; Park, Y.; Cho, S. K.; Kim, H. K.; Choi, Y. Single test-based diagnosis of multiple cancer types using Exosome-SERS-AI for early stage cancers. *Nat. Commun.* **2023**, *14* (1), No. 1644.
- (28) Yang, J.; Chen, X.; Luo, C.; Li, Z.; Chen, C.; Han, S.; Lv, X.; Wu, L.; Chen, C. Application of serum SERS technology combined with deep learning algorithm in the rapid diagnosis of immune diseases and chronic kidney disease. *Sci. Rep.* **2023**, *13* (1), No. 15719.
- (29) Lee, S.; Namgoong, J. M.; Jue, M.; Joung, Y.; Ryu, C. M.; Shin, D. M.; Choo, M. S.; Kim, J. K. Selective Detection of Nano-Sized Diagnostic Markers Using Au-ZnO Nanorod-Based Surface-Enhanced Raman Spectroscopy (SERS) in Ureteral Obstruction Models. *Int. J. Nanomed.* **2020**, *15*, 8121–8130.
- (30) Mann, H. B.; Whitney, D. R. On a Test of Whether one of Two Random Variables is Stochastically Larger than the Other. *Ann. Math. Stat.* **1947**, *18* (1), 50–60.
- (31) Benjamini, Y.; Hochberg, Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J. R. Stat. Soc. Ser. B (Methodological)* **1995**, *57* (1), 289–300.