RamanFest Paris | FRANCE
November 04-05 2026

Ali Saeidan[1]
Nai-Hao Yin[2], Helen Birch[2], Tony Parker[2], Allen Goodship[2], Pavel Matousek[3], Jemma G Kerns[1]
1-Lancaster University, School of health and science, Lancaster, United Kingdom
2-University College London, London, United Kingdom
3-University of Exeter, Exeter, United Kingdom

a.saeidan@lancaster.ac.uk

Influence of Spatial Offset on Raman Spectroscopic Discrimination of Bone Diseases Using PCA and PLS-DA

Spatially offset Raman spectroscopy (SORS) enables non-invasive acquisition of Raman signals from subsurface tissue and therefore has potential for detecting biochemical changes associated with bone disease. This study investigated whether SORS spectra could differentiate osteogenesis imperfecta (OI), osteoporosis (OP), and osteoarthritis (OA) from normal controls and examined the influence of spatial offset on disease discrimination. SORS spectra acquired at 0, 2, 5, and 10 mm offsets were pre-processed to reduce baseline and intensity variation before multivariate analysis. Principal component analysis (PCA) was used to investigate spectral variation associated with spatial offset and disease, while partial least squares discriminant analysis (PLS-DA) was applied separately to each disease-versus-control comparison. Disease-discriminating Raman variables were further evaluated using false discovery rate (FDR)-corrected statistical testing, Cohen's *d* effect size, and variable importance in projection (VIP).

PCA demonstrated a clear offset-dependent spectral transition, with lower offsets (0–2 mm) tending to separate from higher offsets (5–10 mm), while disease groups showed substantial overlap in unsupervised PCA space. PLS-DA provided stronger disease discrimination, particularly at the 10 mm offset, where independent-test accuracies of **93.3% for OI, 92.9% for OP, and 83.3% for OA** were obtained. At 10 mm, balanced accuracies were **90%, 95%, and 77.5%**, respectively. Raman variables simultaneously satisfying **FDR q < 0.05, |Cohen's d| ≥ 0.8, and VIP > 1** were identified, with 27 features for OI, 75 for OP, and 80 for OA at 10 mm. These findings indicate that disease-related spectral information is subtle relative to overall spectral variability but can be extracted using supervised multivariate analysis. The consistently strong performance at 10 mm suggests that larger spatial offsets may provide more useful depth-weighted spectral information for distinguishing diseased from normal samples.

References

[1] Buckley, K., Kerns, J. G., Gikas, P. D., Birch, H. L., Vinton, J., Keen, R., Parker, A. W., Matousek, P. & Goodship, A. E. (2014). *Measurement of abnormal bone composition in vivo using noninvasive Raman spectroscopy*. **IBMS BoneKEy, 11**, Article 602. https://doi.org/10.1038/bonekey.2014.97.

[2] Hosseini, M., Afrin, S., Yosick, A., Awad, H. & Berger, A. J. *Machine learning reconstruction of digit bone Raman spectra enables noninvasive transcutaneous detection of systemic osteoporosis*.

[3] Sowoidnich, K., Churchwell, J. H., Buckley, K., Goodship, A. E., Parker, A. W. & Matousek, P. (2017). *Spatially offset Raman spectroscopy for photon migration studies in bones with different mineralization levels*. **Analyst, 142**, 3219–3226. https://doi.org/10.1039/C7AN00408G.

[4] Mosca, S., Dey, P., Salimi, M., Gardner, B., Palombo, F., Stone, N. & Matousek, P. (2021). *Spatially Offset Raman Spectroscopy—How Deep?* **Analytical Chemistry, 93**, 6755–6762.

[5] Matousek, P. & Stone, N. (2016). *Development of deep subsurface Raman spectroscopy for medical diagnosis and disease monitoring*. **Chemical Society Reviews, 45**, 1794–1802.

Figures

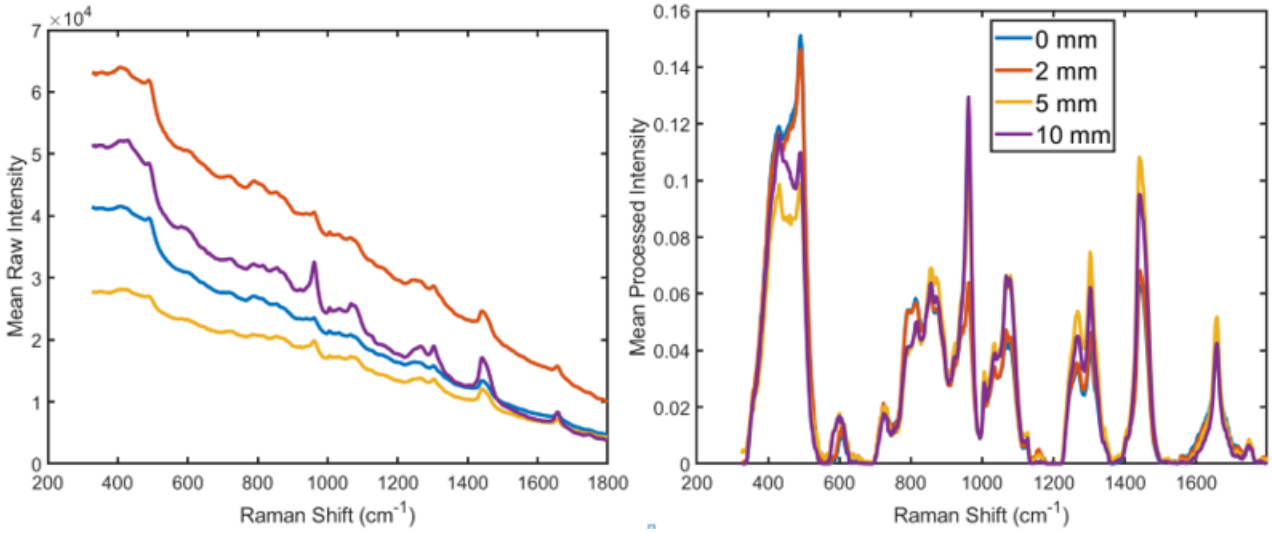

Figure 1. Raw and processed SORS spectra, showing individual spectra (top) and mean spectra at 0, 2, 5, and 10 mm offsets (bottom). Preprocessing reduces baseline and intensity variations and enhances Raman features.

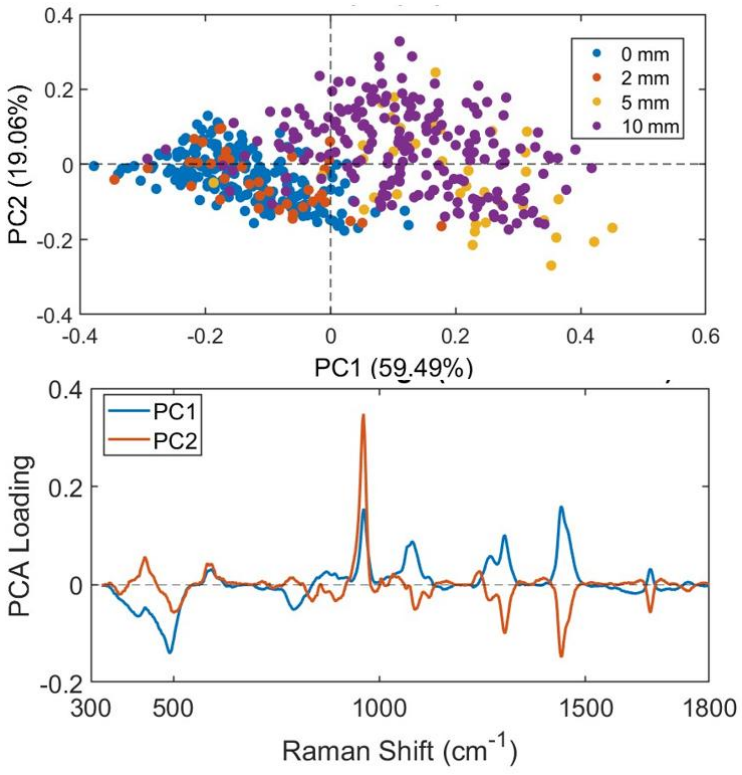

Figure 2. PCA of SORS spectra by spatial offset: (a) PC1 vs PC2 score plot for 0, 2, 5, and 10 mm offsets; (b) corresponding PC1 and PC2 loading plots.

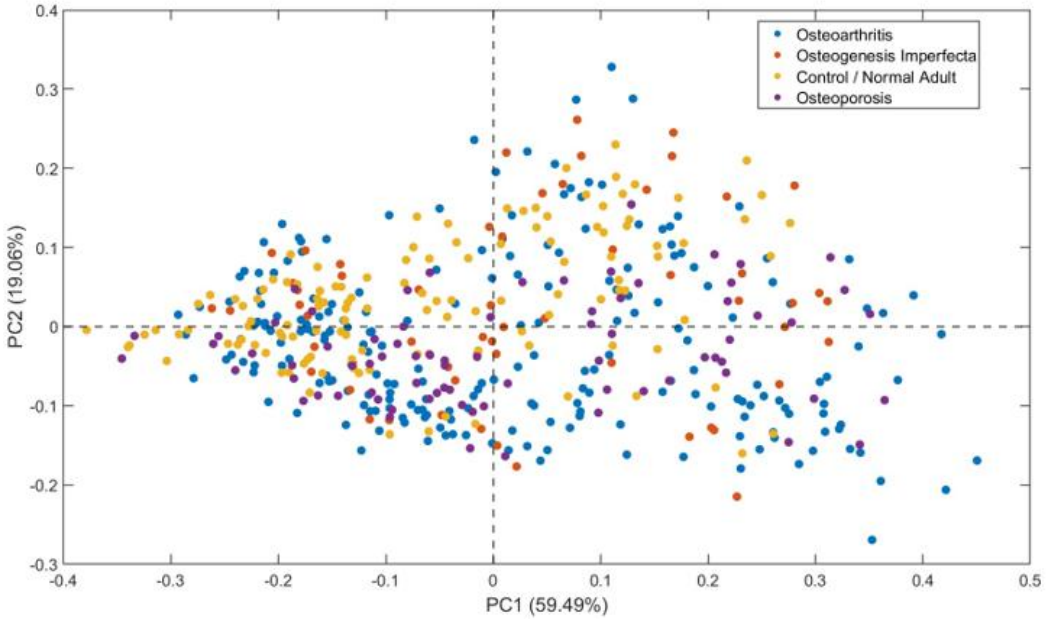

Figure 3. PCA score plot of SORS spectra across all offsets. PC1 and PC2 explain 59.49% and 19.06% of the total spectral variance, respectively. Samples are coloured according to disease group: osteoarthritis, osteogenesis imperfecta, control/normal adult, and osteoporosis.

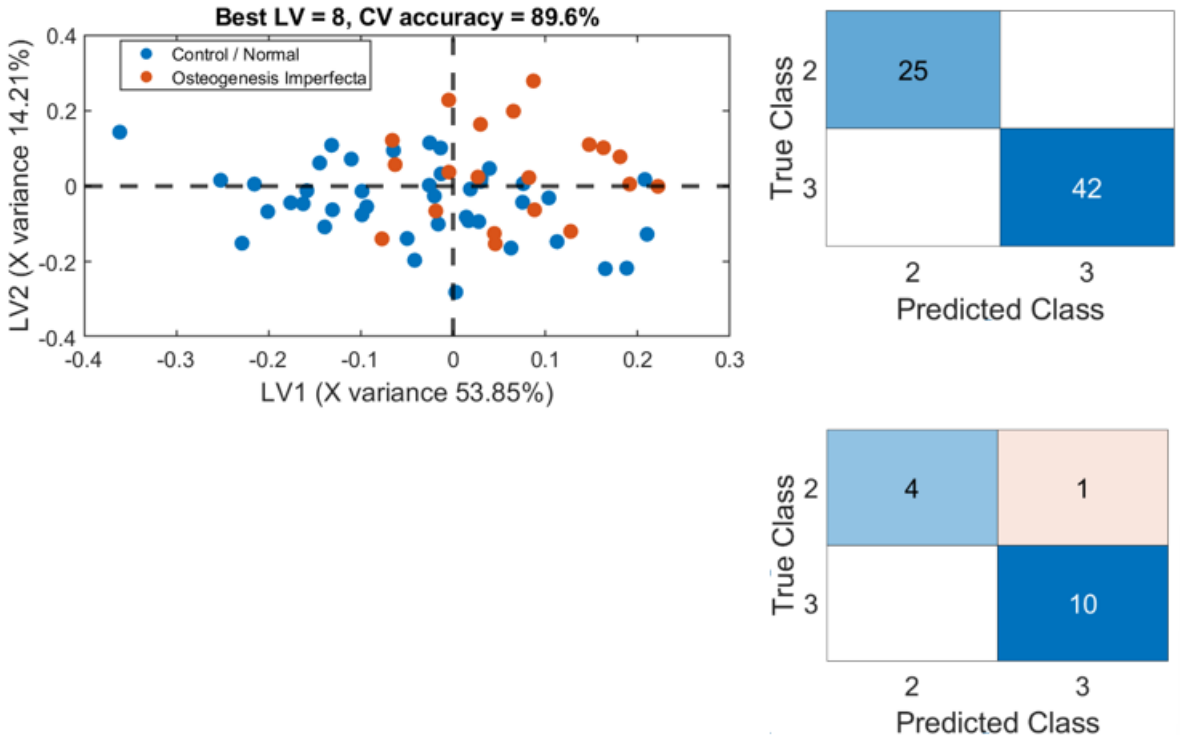

Figure 4. PLS-DA classification of Osteogenesis Imperfecta (OI) versus Control/Normal at the selected SORS offset. Left: LV1–LV2 training score plot using the optimal 8 latent variables, with a cross-validation accuracy of 89.6%. Right: corresponding training and independent-test confusion matrices, showing class-specific prediction performance for OI (class 2) and Control/Normal (class 3).

Disease	Offset mm	Train Control	Train Disease	Test Control	Test Disease	Best LV	CV Accuracy	Train Accuracy	Test Accuracy
OI	0	42	25	10	5	6	84.9	98.5	53.3
OI	2	8	5	2	1	2	93.3	92.3	66.7
OI	5	8	5	2	1	2	80	84.6	66.7
OI	**10**	**42**	**25**	**10**	**5**	**8**	**89.6**	**100**	**93.3**
OP	0	42	26	10	4	5	91.3	100	85.7
OP	2	8	7	2	1	3	66.7	100	66.7
OP	5	8	7	2	1	2	46.7	73.3	100
OP	**10**	**42**	**26**	**10**	**4**	**6**	**89.7**	**100**	**92.9**
OA	0	42	71	10	20	5	77.9	95.6	70
OA	2	8	12	2	4	1	55	65	83.3
OA	5	8	12	2	4	2	55	70	66.7
OA	**10**	**42**	**71**	**10**	**20**	**6**	**79.6**	**97.3**	**83.3**

Table 1. PLS-DA classification performance for osteogenesis imperfecta (OI), osteoporosis (OP), and osteoarthritis (OA) versus normal controls across SORS offsets (0, 2, 5, and 10 mm).

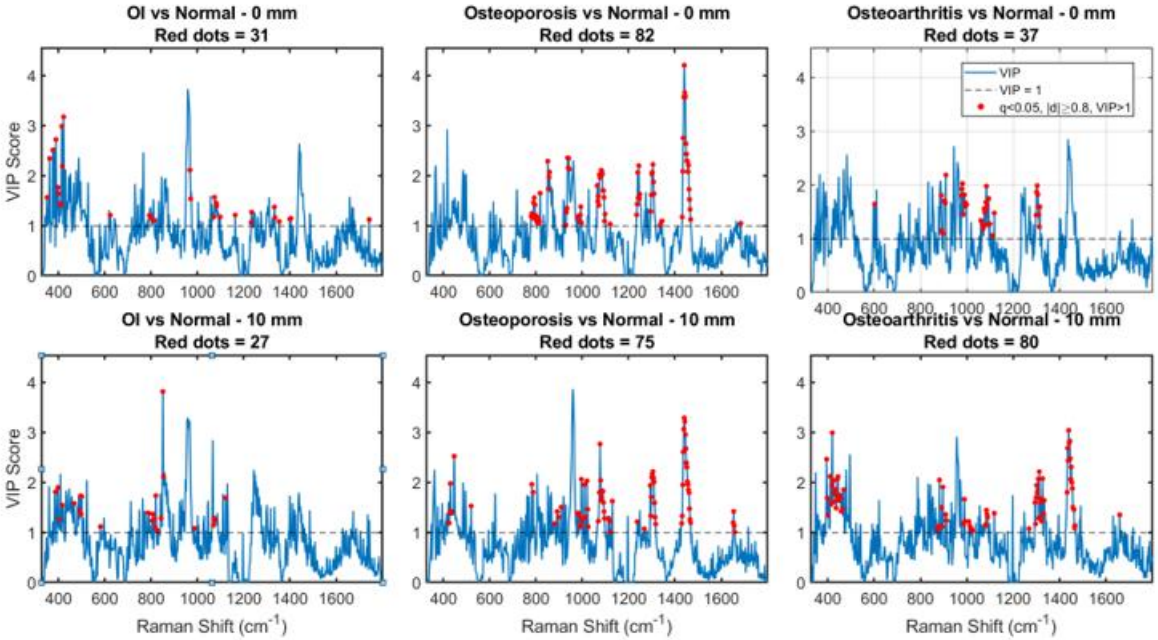

Figure 5. VIP profiles for OI, osteoporosis, and osteoarthritis versus normal controls at 0 and 10 mm SORS offsets. Red dots indicate features meeting VIP > 1, |d| ≥ 0.8, and FDR q < 0.05.