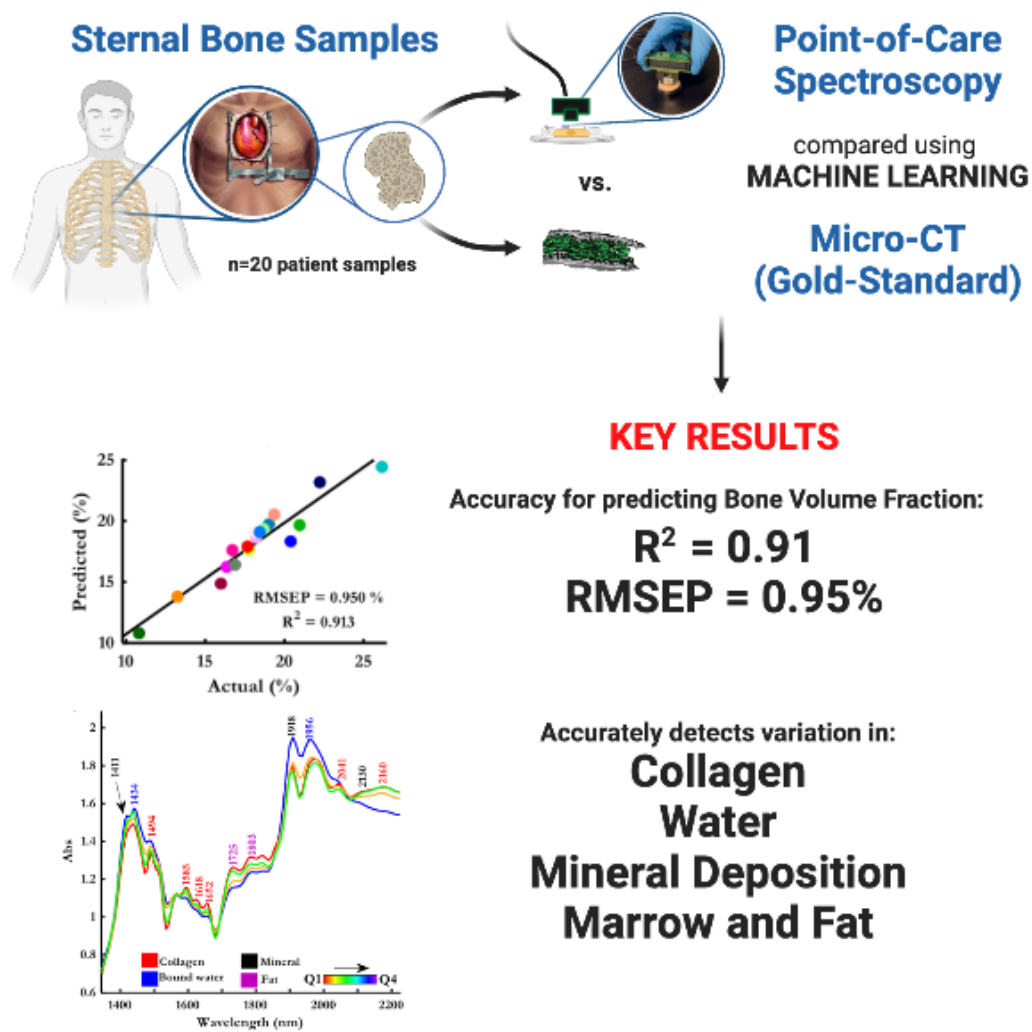# Bone & Joint Open
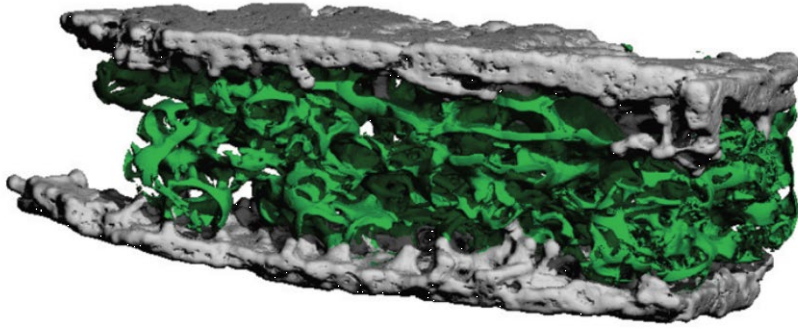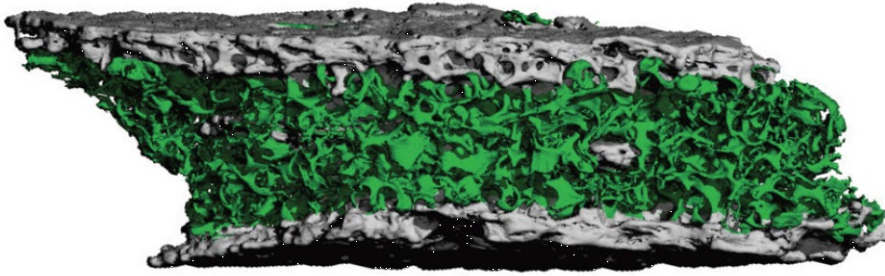
Supplementary Material

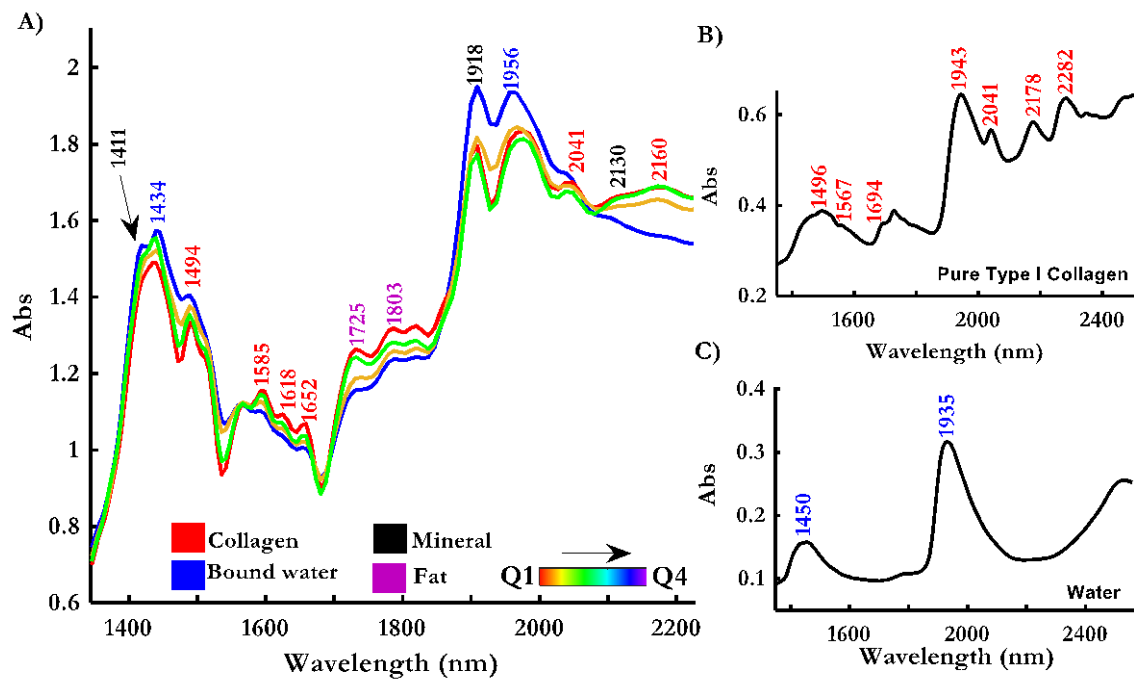**Fig. a.** Graphical abstract of key findings. RMSEP, root mean square error of prediction.

A)

B)

Fig. b. Micro-CT analysis of bone samples. 3D reconstructions of micro-CT data of two extreme samples with a) low and b) high bone volume density. Cortex is visualized in grey and trabeculae in green.

Bone component	Observed Bands (nm)	Band assignment and intensities across quartiles
Fibrous Collagen	1494	N-H stretch first overtone
	1585, 1618, 1652	N-H stretch + C-H stretch
	2041, 2160	Symmetric N–H stretch and C–H deformation combination respectively
Lipid/Fat	1725, 1803	C–H stretch first overtone (CH_2)
Water	1434, 1954	O-H First and Second overtone
Mineralized component	1411, 1918	O–H first overtone and P-OH combination of bending and stretching modes
	2130	O-H Bending mode

Fig. c. Spectral differences in sternal bone of varying bone volume fraction and their band assignments. a) Average near-infrared (NIR) spectra of human sternal bone recorded by a miniature NIR spectrometer with major bands labelled. The spectra revealed in vivo bone composition, mostly bound water, inorganic mineral content approximated as hydroxyapatite ($Ca_{10}(PO_4)_6(OH)_2$), organic component (collagen), and bone marrow (fat/lipid). The spectra were presented based on the quartile ranking of the bone volume fraction (BV/TV) parameters. Red spectrum is Q1, green is Q2, yellow is Q3, and blue Q4. b) NIR spectrum of type I collagen with bands assigned. c) NIR spectrum of free water, showing two major modes at 1,450 and 1,935 nm. The assigned bands in a), b), and c) were colour-coded to show the relationship between the bands assigned in the in vivo spectra, and bands arising from the pure standard. Abs, absorbance; Q, quartile of bone volume fraction.
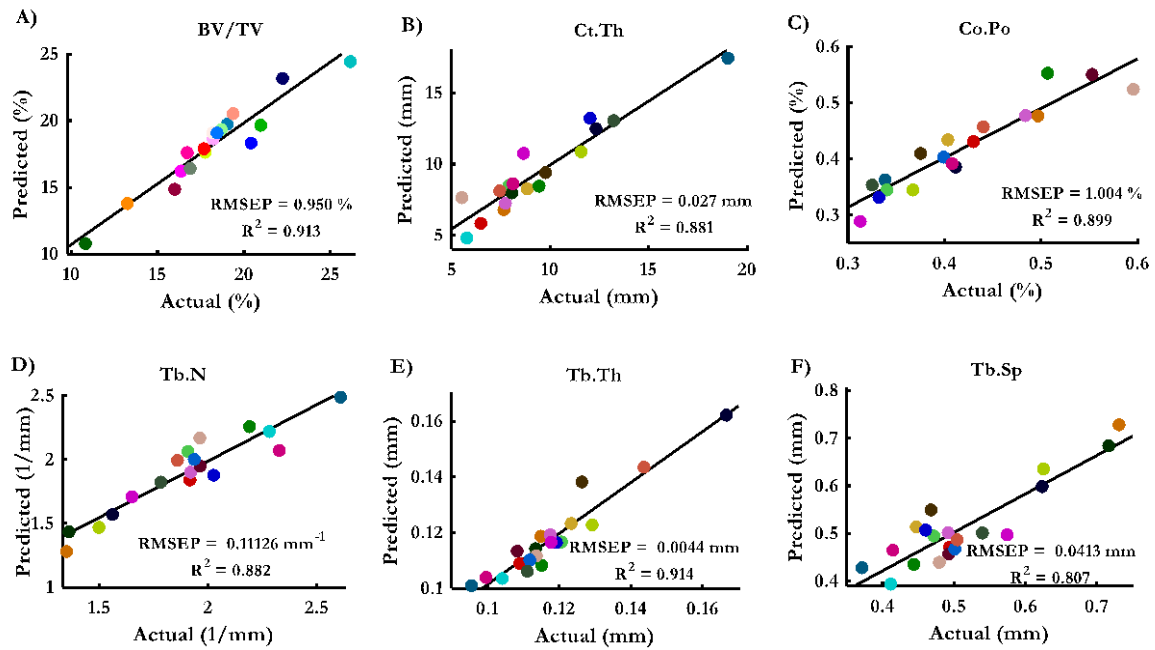
Fig. d. Partial least square regression analysis showing correlation between average near-infrared (NIR) spectra of a human bone sample captured by miniature NIR spectrometer on the trabecular surface with micro-CT parameters. The score plot shows the relationship between the measured spectra and a) bone volume fraction (BV/TV), b) cortical thickness (Ct.Th), c) cortical porosity (Ct.Po), d) trabecular number (Tr.N), e) trabecular thickness (Tb.Th), and f) trabecular spacing (Tb.Sp). RMSEP, root mean square error of prediction.
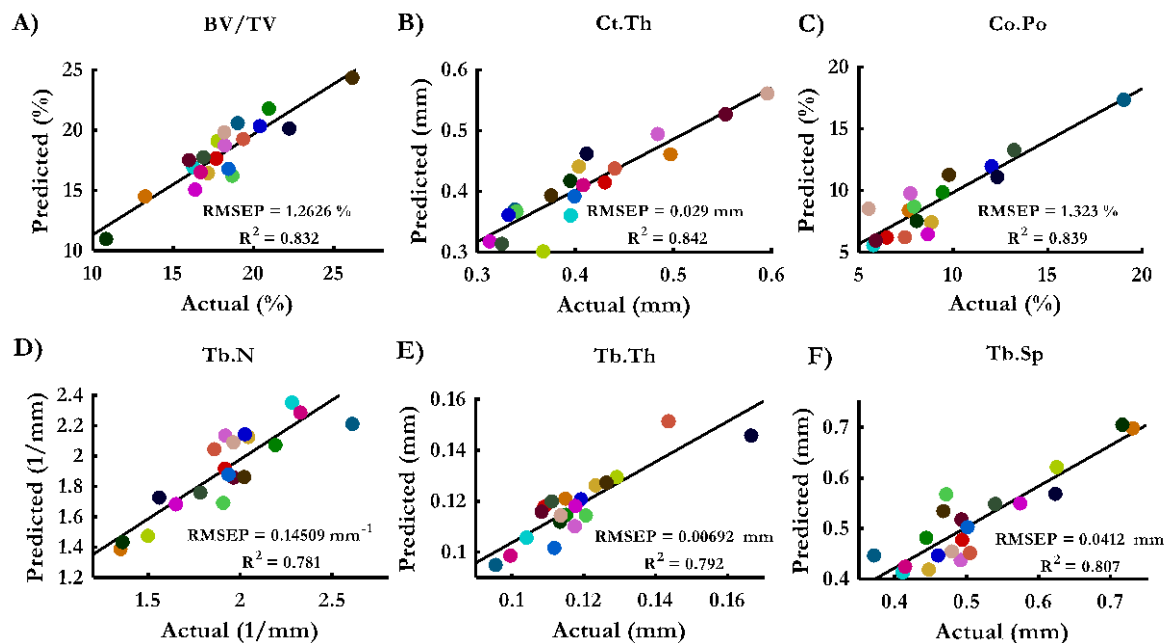

Fig e. Partial least square regression analysis showing correlation between average near-infrared (NIR) spectra of a human bone sample acquired from the outer cortical surface, with micro-CT parameters. The score plots show the relationship between the measured spectra and a) bone volume fraction (BV/TV), b) cortical thickness (Ct.Th), c)

cortical porosity (Ct.Po), d) trabecular number (Tb.N), e) trabecular thickness (Tb.Th), and f) trabecular spacing (Tb.Sp). RMSEP, root mean square error of prediction.
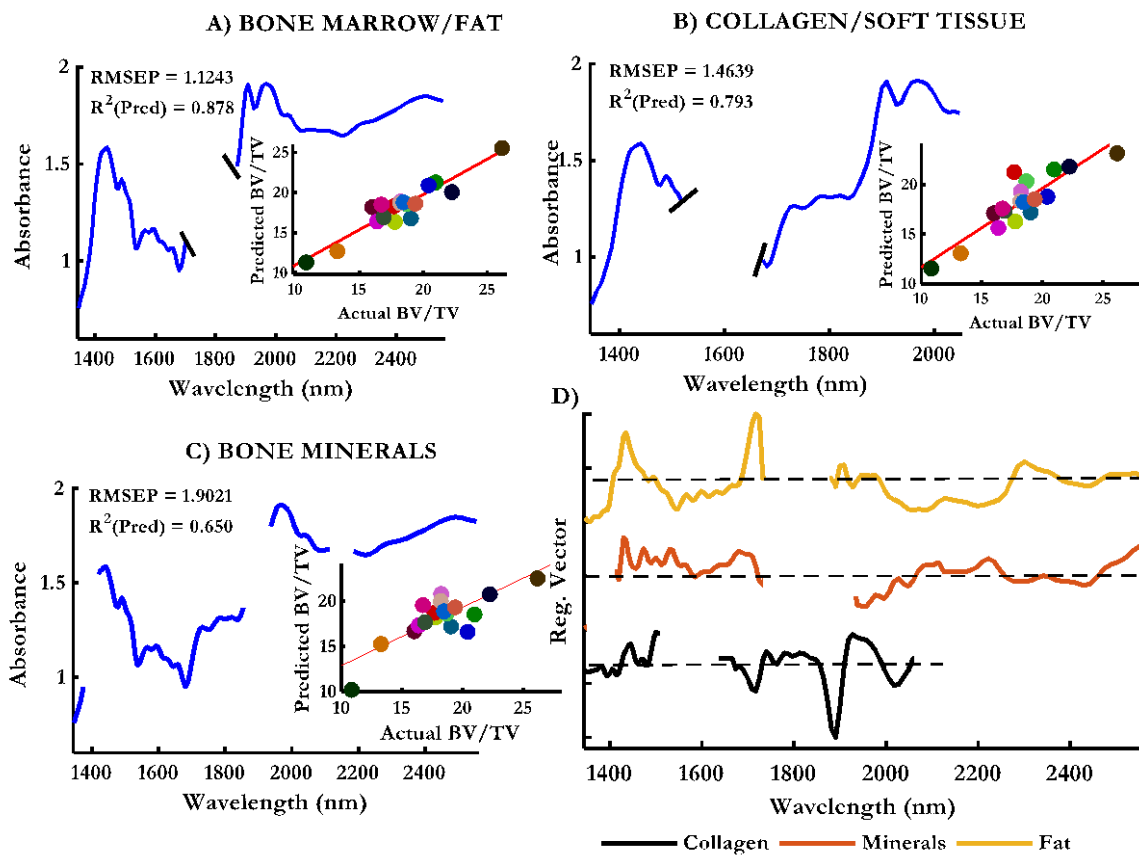
Fig. f. Results depicting the effect of the various compositions of bone on the prediction capabilities and accuracies. Partial least square (PLS) regression score plot after eliminating a) lipid/fat modes at 1,700 to 1,850 nm. b) Collagen modes at wavelength range of 1,490 to 1,650 nm and 2,050 to 2,350 nm. c) Bone mineral absorption at region 1,411 nm and 2,130 nm. d) The loading plot associated with the PLS models a), b), and c). BV/TV, bone volume fraction; RMSEP, root mean square error of prediction.
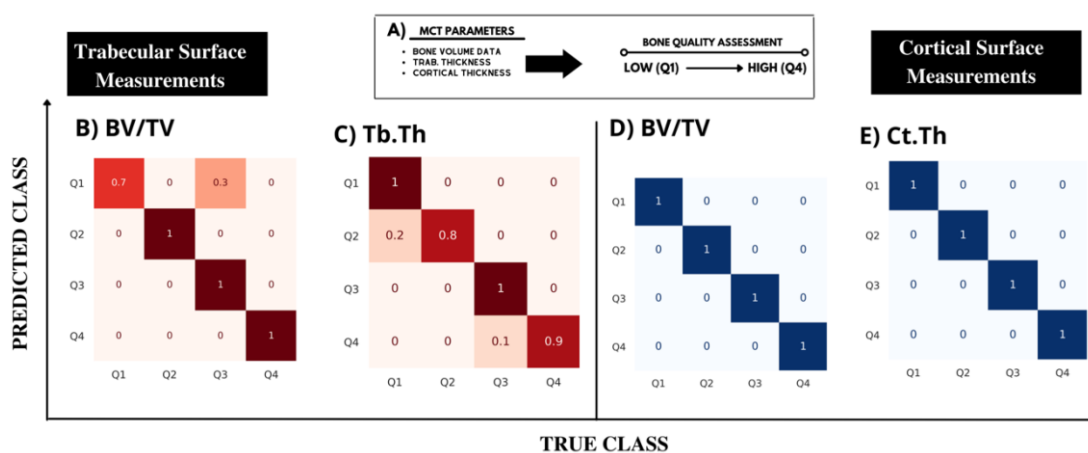
Fig. g. Support vector machine classification analysis. The normalized confusion matrices from each model showed the ability of the near-infrared spectra (NIRS) technique to classify bone into different classes. a) The classes used in these models are the micro-CT (MCT) parameters ranked into quartiles (Q1, Q2, Q3, and Q4). The true class represents the MCT measurements, and predicted class represents the NIR predictions. Models are separated for measurements taken at the b) and c) trabecular and d) and e) cortical surface.

Fig. h. Iterative steps to generate automated diagnosis for clinical instruments that use machine learning. Our instrument achieves the steps of collating a) source data, entering them into a b) code model, used to c) train, evaluate, and tune a diagnostic model, which can d) be deployed to e) generate predicted diagnoses, supplementing clinical best practice. Future work will need to f) monitor ongoing predictions, with g) ongoing updates to models and versions. Collectively, this can be used to create a model for automated diagnoses.

Influence of bound water on bone quality assessment

Figure i below shows the 'before' and 'after' of water peak subtraction of the BV/TV PLS model. All bands were assigned to their respective molecules based on previous work, The raw spectrum (Figure ia) shows the two intense water peaks at 1,434 nm and 1,956 nm, while Figure ib shows the PLS regression plot for the BV/TV model exhibiting a high R^2 value. After excluding the bound water regions, the correlation dropped slightly from 0.915 to 0.814 (Figures ic and id) and resulted in increased RMSEP values (0.95% to 1.5659%). The corresponding loading plots (Figures ie and if) highlight the main peaks that influenced the model. The PLS-R regression vector for the model inclusive of water modes is shown in Figure f indicating that all the three components (water, fat, and collagen) are primarily responsible for the prediction. The protein component is characterized by bands at 1,541, 1,665, 2,091, 2,209, and 2,287 nm, while fat has an intense band at 1,724 nm and another slightly less intense band at 1,872 nm. Furthermore, the presence of minerals is characterized by the strong peak at 1,411 nm, and a moderately intense band at 1,927 nm. The shoulder peak at 1,445 nm indicates the presence of the bound water. However, after adjusting for water modes, most of these bands' positions shifted to a lower wavelength.
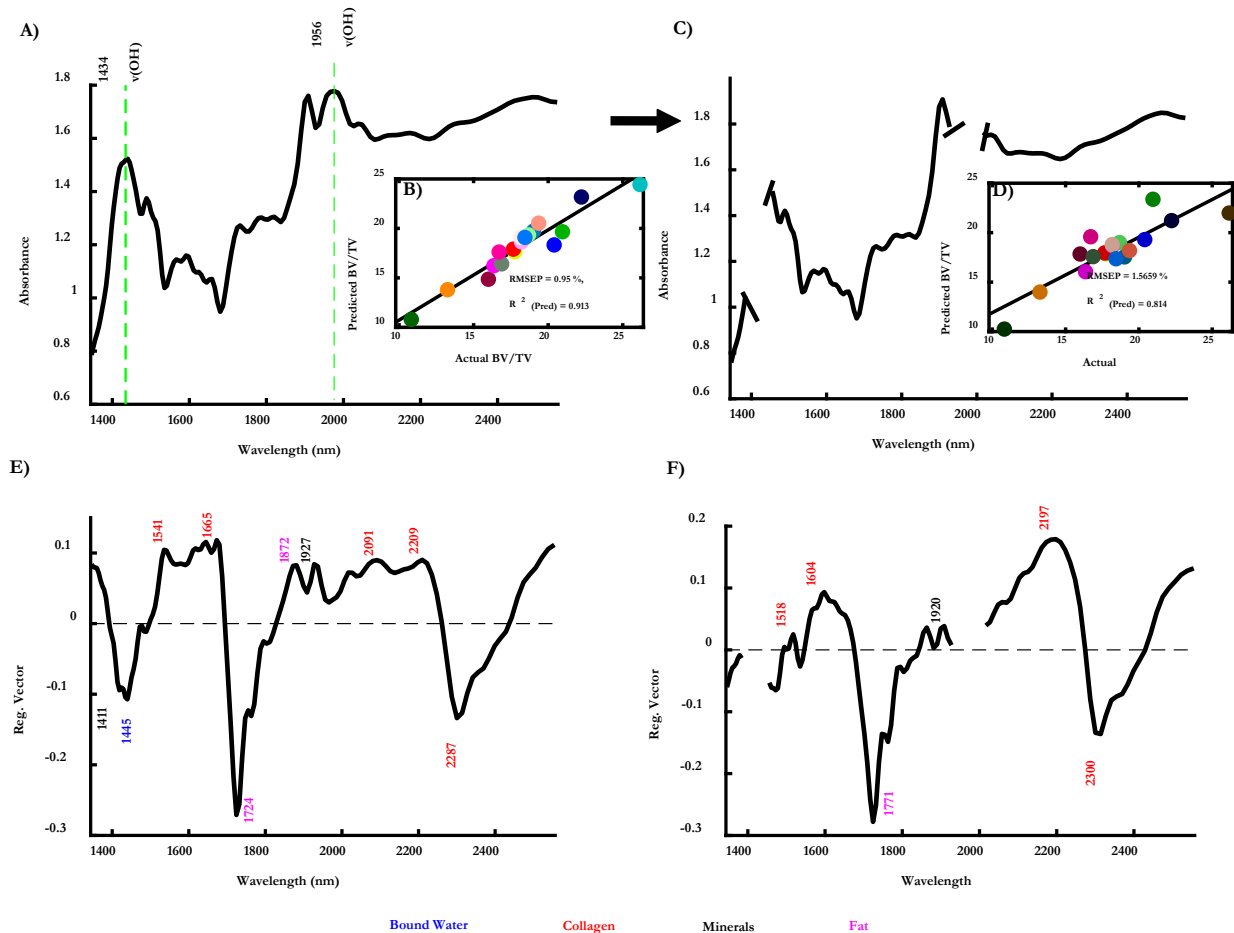
Fig. i. Results demonstrating the effect of water absorption in quantifying bone volume data (BV/TV). a) Raw near-infrared (NIR) spectra used for partial least square (PLS) modelling before adjusting for water absorption. b) The resulting PLS score obtained from the PLS model built, showing high correlation coefficient (0.913%) and low root mean square error progression (RMSEP). c) The raw NIR spectra with major water modes eliminated. d) The corresponding PLS regression (PLS-R) score plot. The R^2 values have reduced to 0.814% and the RMSEP has increased (1.5659%). e) and f) The loading plots from both models, revealing the characteristic bands driving the models.

Discussion of chemical composition

NIR spectra presented in this work show major bands from the key chemical components of the bone matrix that can be used to classify bone integrity.[1-3] After careful band assignment and comparison with the literature,[1-5] the following deductions were made.

First, the two bound water bands at 1,434 nm and 1,956 nm were similar to those previously reported.[2] However, while the band at 1,434 nm is related to mineral bound water and is a marker for bone strength,[1] the 1,956 nm is connected to the collagen-bound water, which is representative of bone toughness.[1,2] The intensities of the bands decreased from Q4 to Q1, which indicates weaker bones have less bound water content and vice versa. This is consistent with what Ailavajhala et al[3] who reported dehydration studies performed on human cadaveric bone using an environmentally controlled chamber and NIR imaging.

Second, the use of NIRS for the determination of bone mineralized components are generally believed to be weak due to the non-existent absorption from the inorganic moieties in the NIRS compared to Raman and infrared.[6] More recent studies,[1,3] however, have studied the distribution of water associated with bone mineral at the microscopic level using NIRS imaging and assigned specifics peaks to the phosphate hydroxide (P-OH) vibrations emerging from the moieties of hydroxyapatite. The bands assigned to mineral matrix here are consistent with those earlier reports showing a decrease across the quartiles (Q4 to Q1). This is anticipated as bones with higher BV/TV parameters have a higher mineral content and lower Ct.Po.

Furthermore, most NIRS bone studies always have the bone marrow (fat content) removed,[2,3] except for Rajapakse et al[1] who carried out the first study, evaluating NIR spectra imaging of bone with marrow. Although they were mostly interested in bone water assessment, they indicated that there was not a significant correlation between the NIRS intensity of the water absorbance and the fat absorbance. To the best of our knowledge, this is the first study reporting on single point spectral analysis of bone including marrow, requiring much less preparation. In the region 1,700 nm to 1,850 nm where the CH_2/CH_3 modes correspond to lipids, there was a negative relationship with the spectra presented in Figure ba where the bones with higher BV/TV were associated with lesser lipid/fat content and vice versa.

References

1. **Rajapakse CS, Padalkar MV, Yang HJ, Ispiryan M, Pleshko N.** Non-destructive NIR spectral imaging assessment of bone water: Comparison to MRI measurements. *Bone.* 2017;103:116–124.
2. **Ailavajhala R, Querido W, Rajapakse CS, Pleshko N.** Near infrared spectroscopic assessment of loosely and tightly bound cortical bone water. *Analyst.* 2020;145(10):3713–3724.
3. **Ailavajhala R, Oswald J, Rajapakse CS, Pleshko N.** Environmentally-controlled near infrared spectroscopic imaging of bone water. *Sci Rep.* 2019;9(1):10199.
4. **Adegoke JA, Kochan K, Heraud P, Wood BR.** A near-infrared "matchbox size" spectrometer to detect and quantify malaria parasitemia. *Anal Chem.* 2021;93(13):5451–5458.
5. **Afara IO, Florea C, Olumegbon IA, et al.** Characterizing human subchondral bone properties using near-infrared (NIR) spectroscopy. *Sci Rep.* 2018;8(1):9733.
6. **Spizzirri PG, Cochrane NJ, Prawer S, Reynolds EC.** A comparative study of carbonate determination in human teeth using Raman spectroscopy. *Caries Res.* 2012;46(4):353–360.