

Supplementary Methods and Figures

S1 Cohort and ingestion diagnostics

S1.1 Per-participant rib-side count distribution

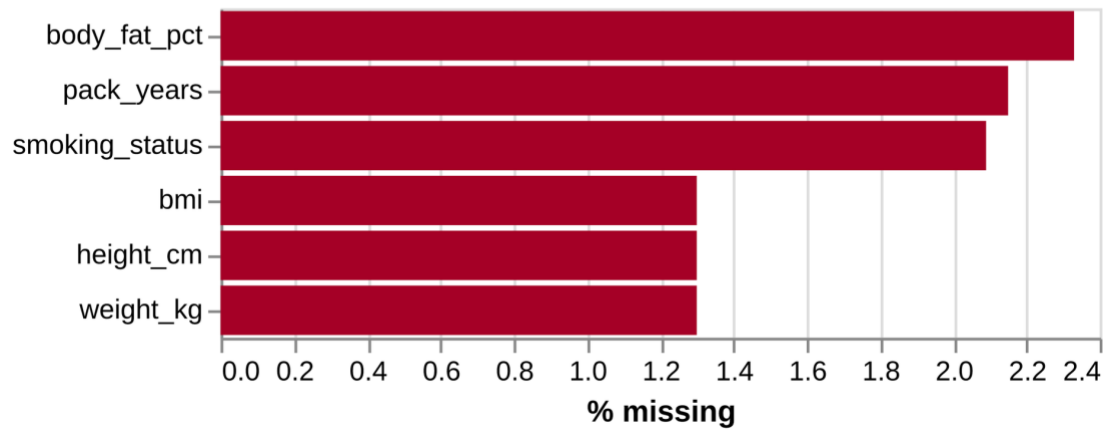
Supplementary Table 1: Distribution of per-participant rib-side counts among the candidate cohort. The inclusion criterion ([Methods §2.5](#)) requires exactly 24 rib sides; participants with different counts are excluded. Low counts (≤ 23) are consistent with transitional thoracolumbar vertebrae, missing segmentations or merged segmentations; high counts (≥ 25) with supernumerary cervical or lumbar ribs.

Rib sides per participant	n participants	%
8	1	0.0%
12	1	0.0%
17	1	0.0%
18	7	0.0%
19	6	0.0%
20	38	0.1%
21	28	0.1%
22	551	1.8%
23	795	2.6%
24	27,502	91.0%
25	693	2.3%
26	595	2.0%

S1.2 Per-variable missingness across the joined cohort

Missingness per variable

% of patients with missing value



Supplementary Figure 1: **Missingness across baseline metadata variables.** Computed on the inner-joined cohort ($n = 30,218$) prior to quality-control exclusions.

S1.3 Predictor collinearity

Correlation matrix

Patient level - Spearman ρ - lower triangle

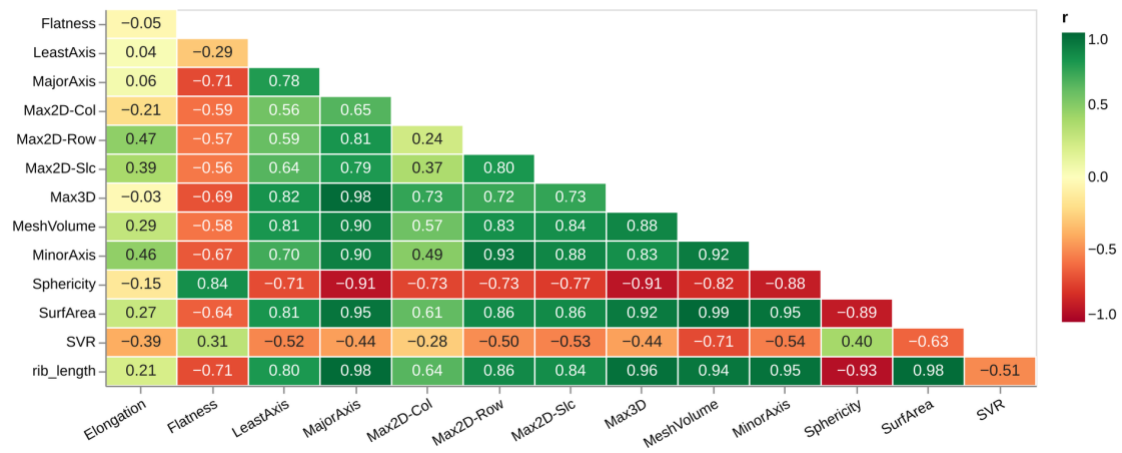


Supplementary Figure 2: **Pairwise Pearson correlations between continuous metadata predictors.** Body mass, BMI, and body-fat percentage cluster tightly, motivating the multivariable adjusted regression design of [Methods §2.8.2](#) rather than relying on bivariate associations.

S1.4 Descriptor collinearity

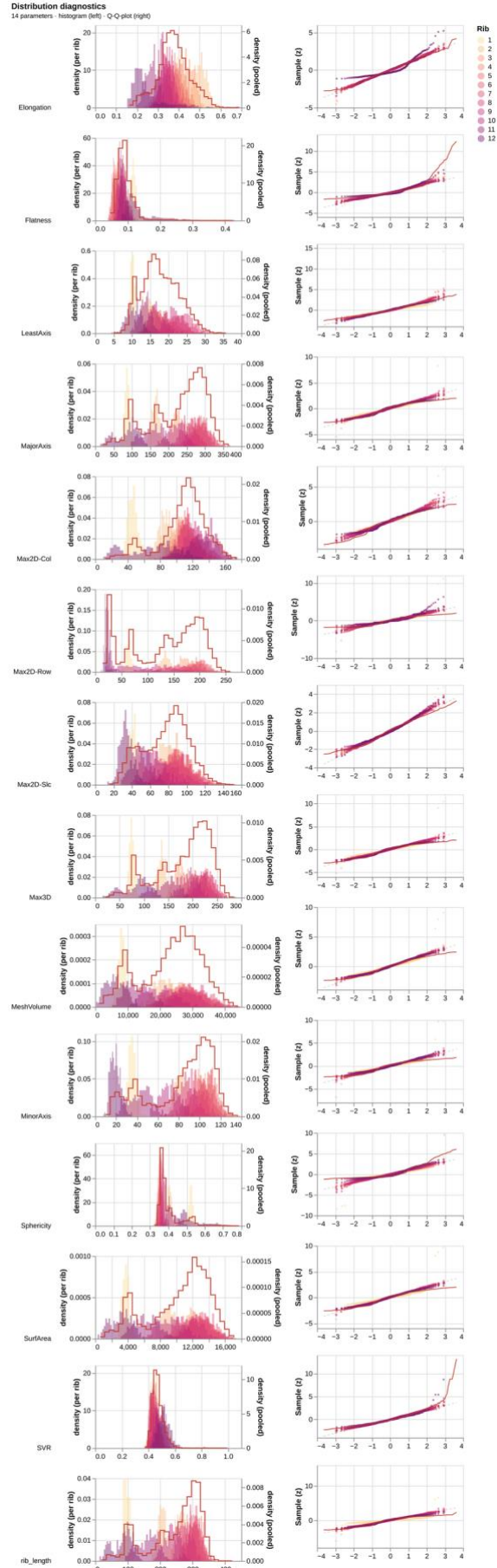
Shape parameter correlations

Rib level · Pearson r · lower triangle



Supplementary Figure 3: **Pairwise Pearson correlations between per-rib shape descriptors.** Computed on participant-level per-rib means. Voxel volume was dropped at the analysis stage ([Methods §2.6](#)) due to conceptual identity (and perfect collinearity) with mesh volume.

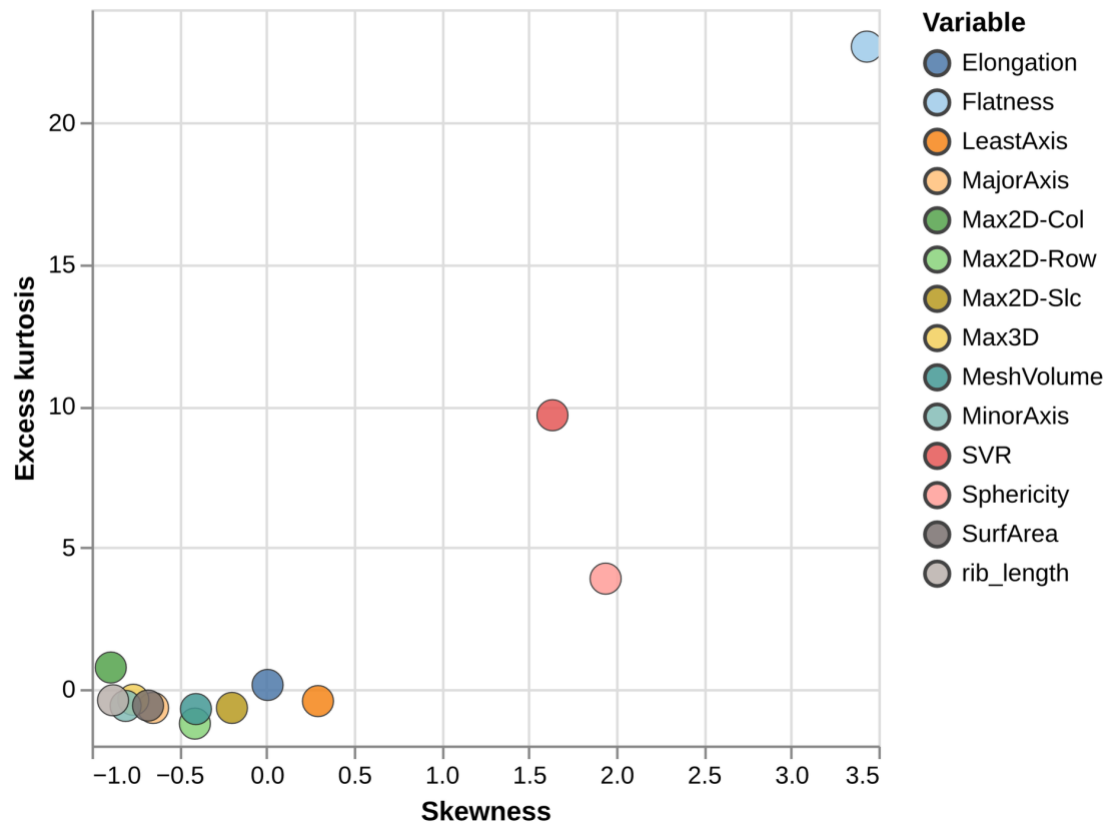
S1.5 Distribution and normality of shape descriptors



Supplementary Figure 4: Per-descriptor histograms with Q-Q overlays. Q-Q axes are z-standardised so that the y =

Normality diagnostics

Skewness $\times$ kurtosis

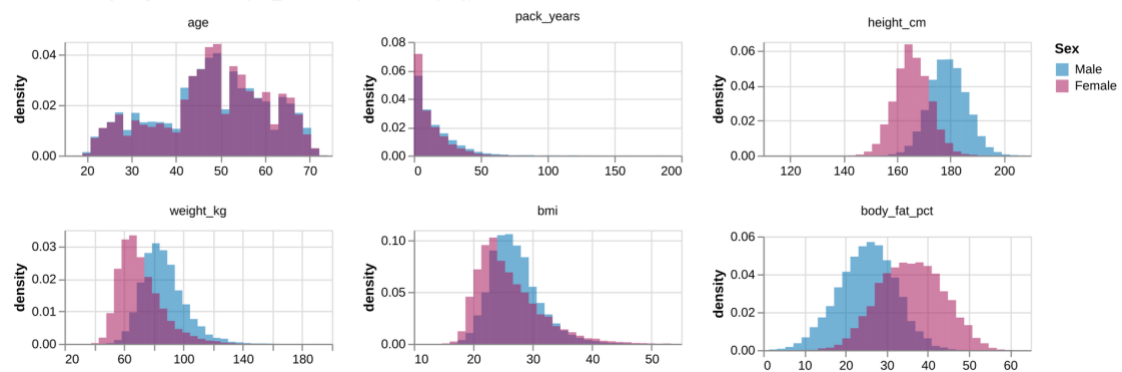


Supplementary Figure 5: *Shapiro–Wilk normality summary across shape descriptors.*

S1.6 Predictor distributions by sex

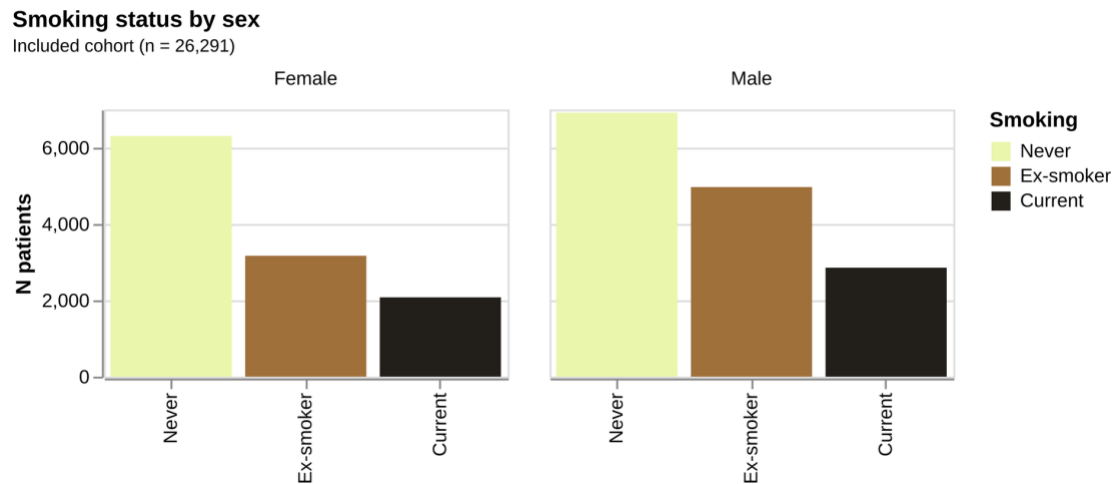
Metadata by sex

Patient-level density histograms · n=26,291 · pack_years zero bin (never-smokers) dropped



Supplementary Figure 6: *Distribution of each continuous metadata predictor stratified by biological sex.*

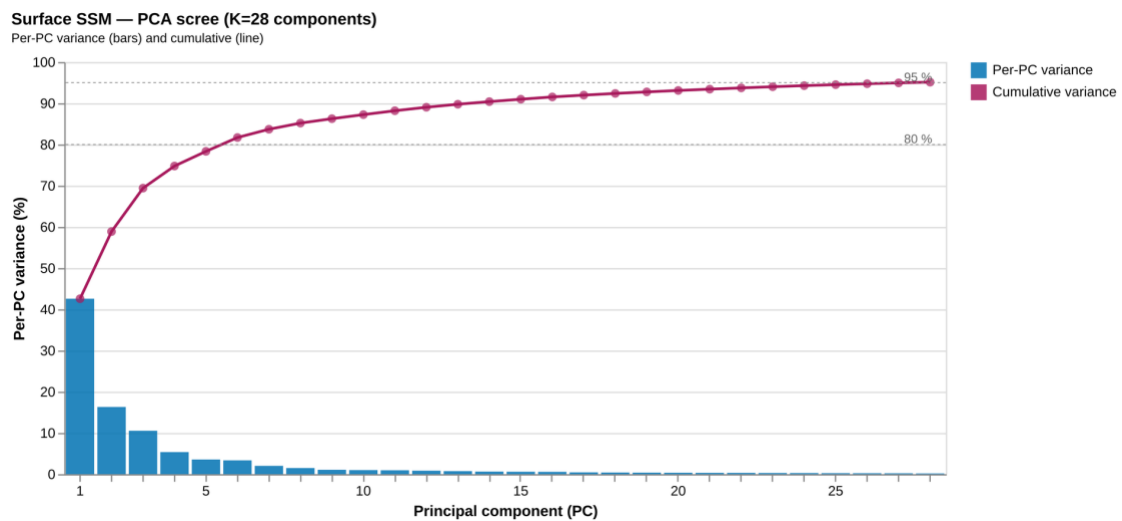
Companion view to [Table 1](#).



Supplementary Figure 7: *Smoking-status composition stratified by biological sex.* Companion view to [Table 1](#).

S2 Shape-model variance and convergence

S2.1 Scree plot

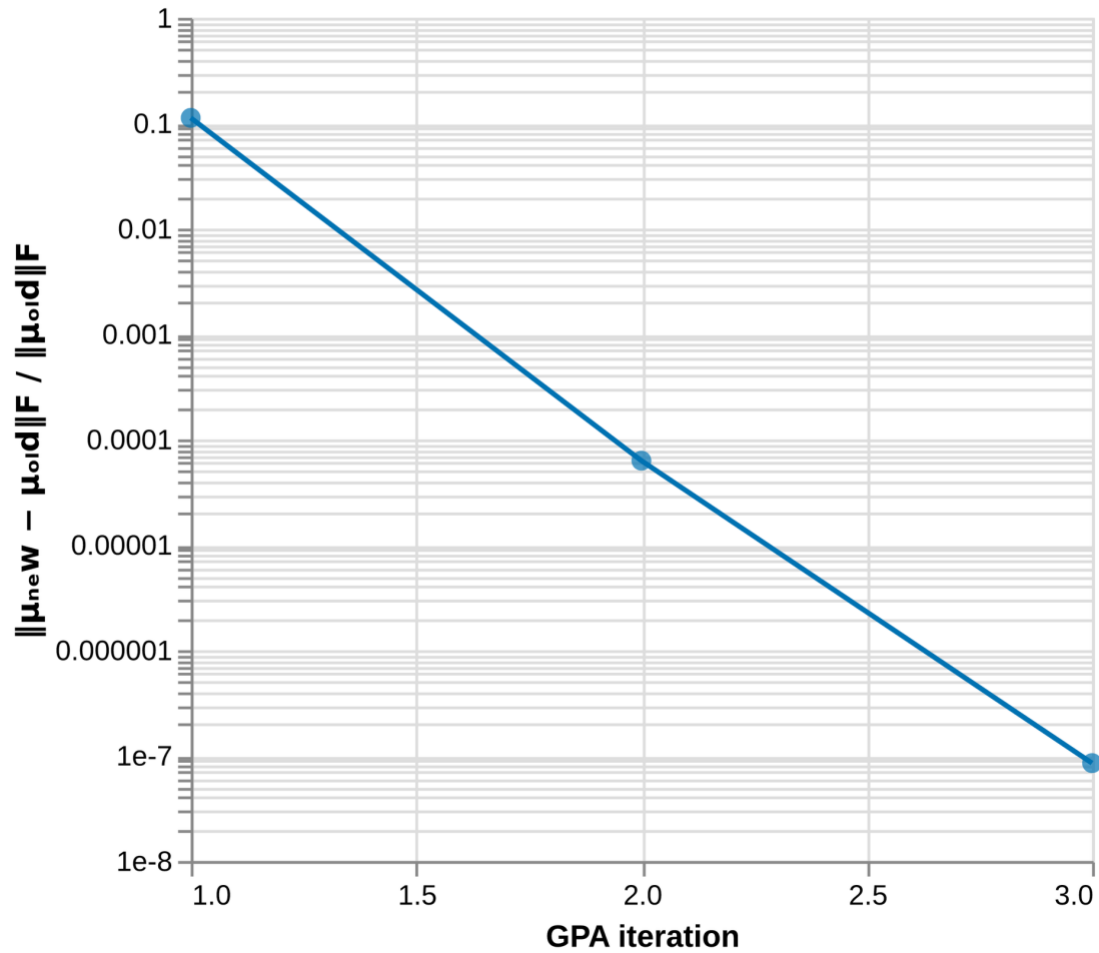


Supplementary Figure 8: *Per-mode and cumulative explained variance.* Variance retained at the truncation threshold (95%) is annotated; 28 modes were carried forward into all downstream PC-score analyses.

S2.2 Generalized Procrustes Analysis convergence

GPA convergence

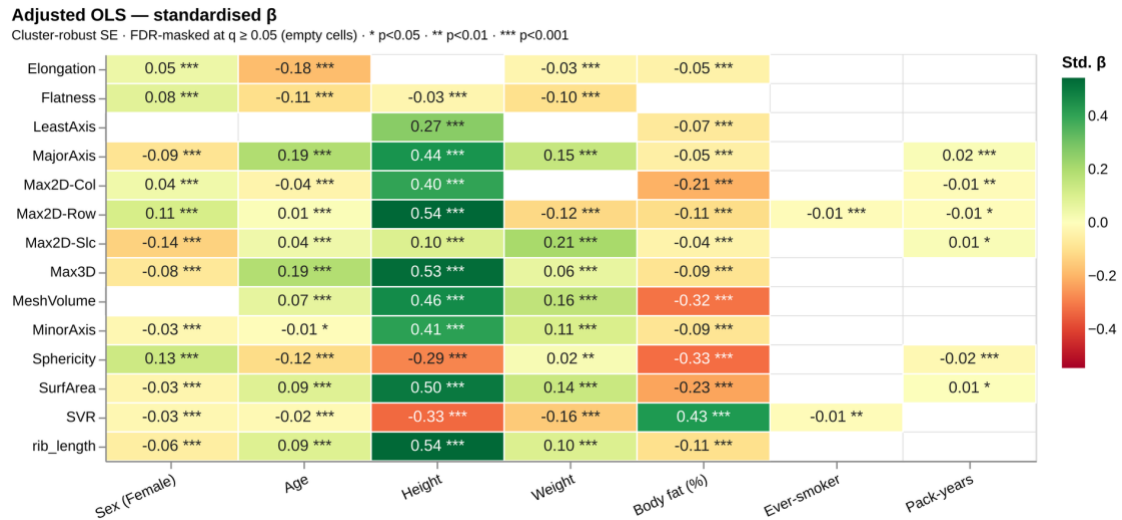
Relative mean-shape change per iteration (log y)



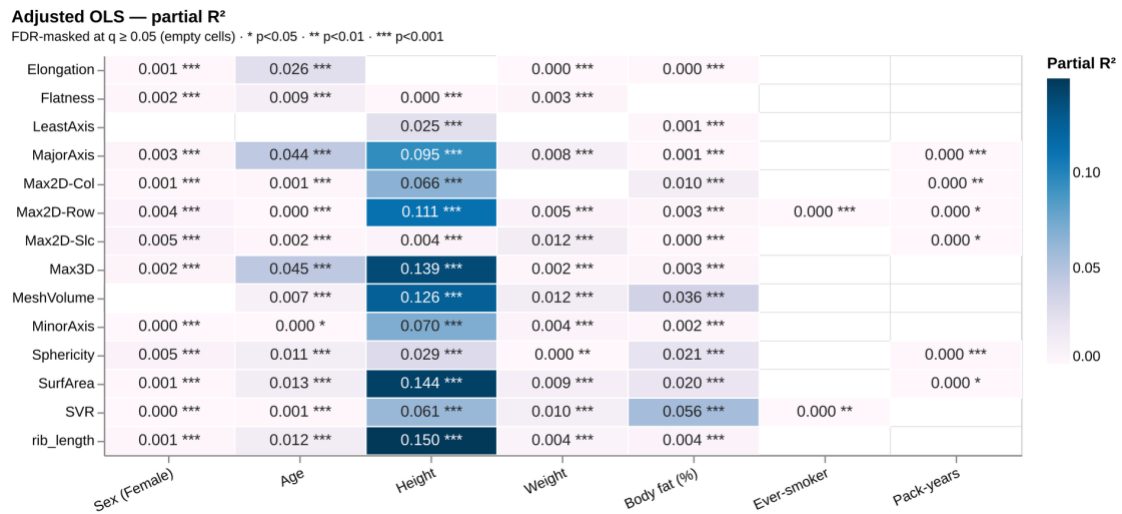
Supplementary Figure 9: **GPA convergence trace.** Mean-shape change per iteration of the Generalized Procrustes Analysis described in [Methods §2.7.3](#), with the convergence threshold marked.

S3 Descriptor-level associations

S3.1 Adjusted OLS effect maps



(a) Panel A – standardised β .



(b) Panel B – Frisch–Waugh–Lovell partial R^2 .

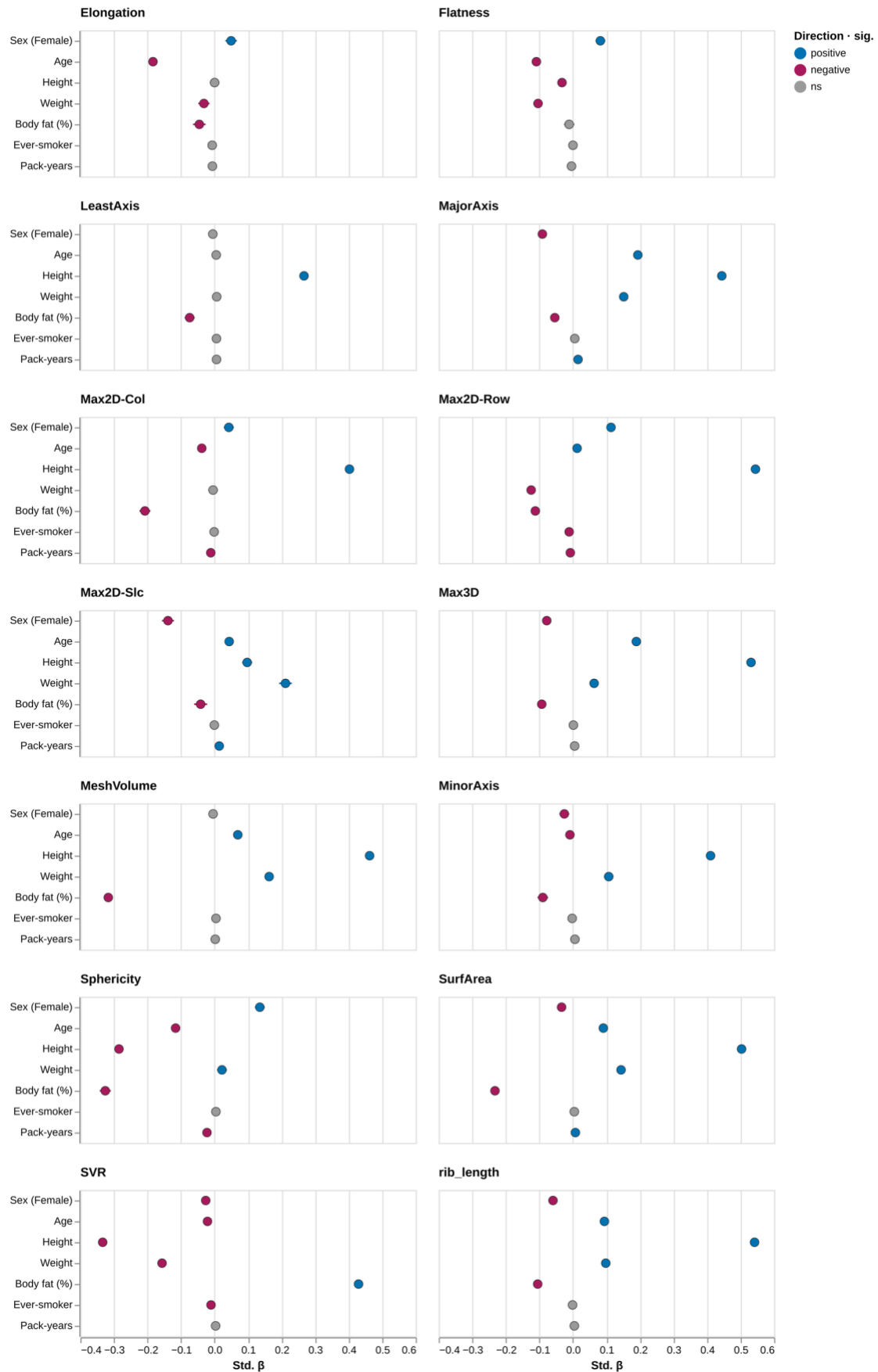
Supplementary Figure 10: **Adjusted (multivariable) OLS effect maps.** (A) Standardised β and (B) Frisch–Waugh–Lovell partial R^2 per (shape descriptor $\times$ predictor) pair from the single adjusted model (Methods §2.8.2) – the mutually-adjusted counterpart to the [main-text targeted maps](#). FDR-masked at $q \geq 0.05$; cluster-robust standard

errors at the participant level.

S3.2 Adjusted OLS standardised β forest plot

Adjusted OLS — standardised β with 95 % CI

Cluster-robust SE · BH-FDR



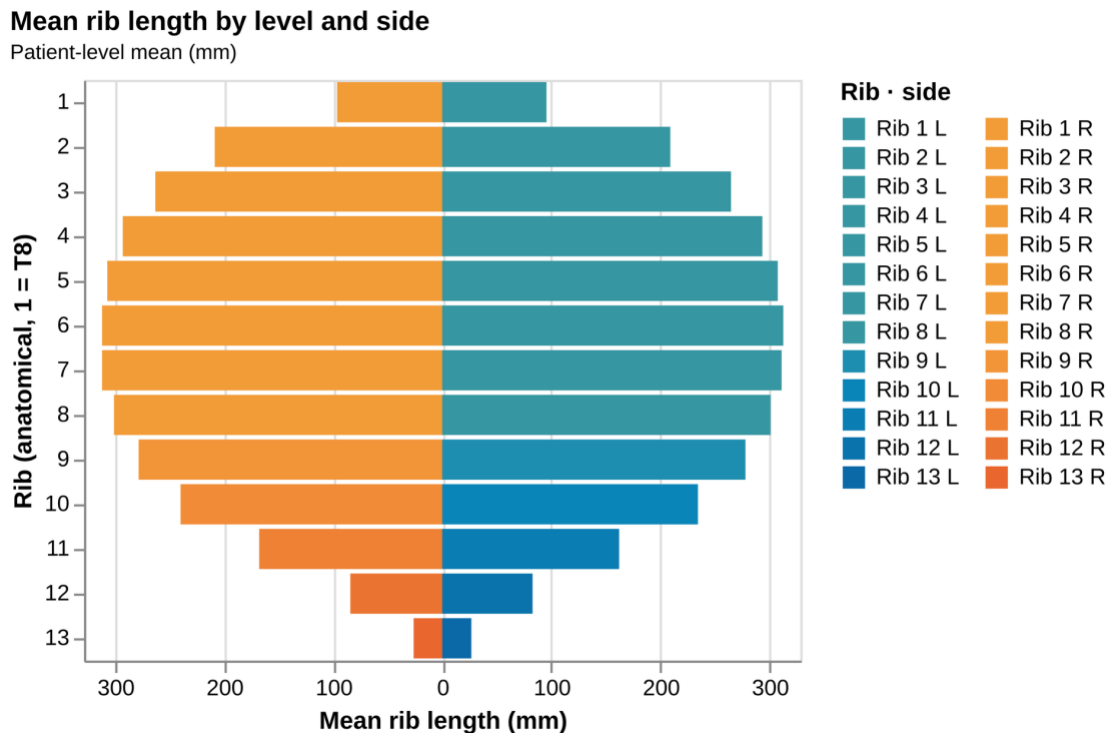
Supplementary Figure 11: Adjusted OLS standardised β forest plot. 95% confidence intervals on the standardised slope per (predictor, descriptor) pair from the adjusted model.

S3.3 Unadjusted (marginal) standardised β heatmap



Supplementary Figure 12: *Unadjusted (marginal) standardised β heatmap*. Per-predictor marginal slope for each (shape descriptor $\times$ predictor) pair, including BMI (unadjusted layer only) – the total-association counterpart to the adjusted and targeted maps.

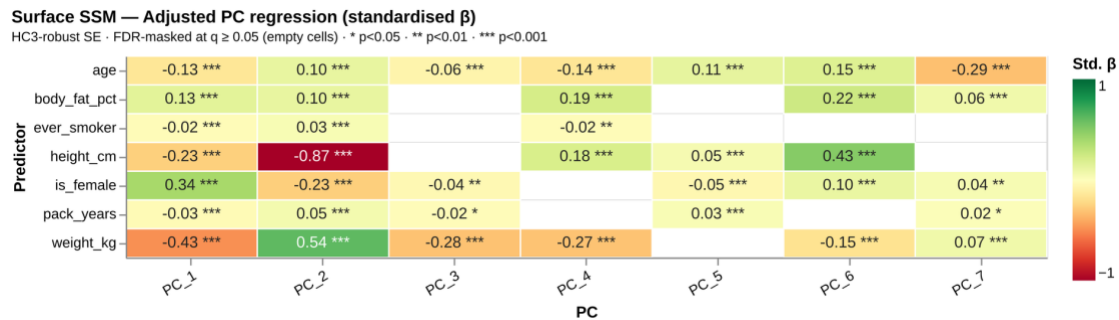
S3.4 Per-descriptor rib-position maps



Supplementary Figure 13: Mean rib length by anatomical level and side.

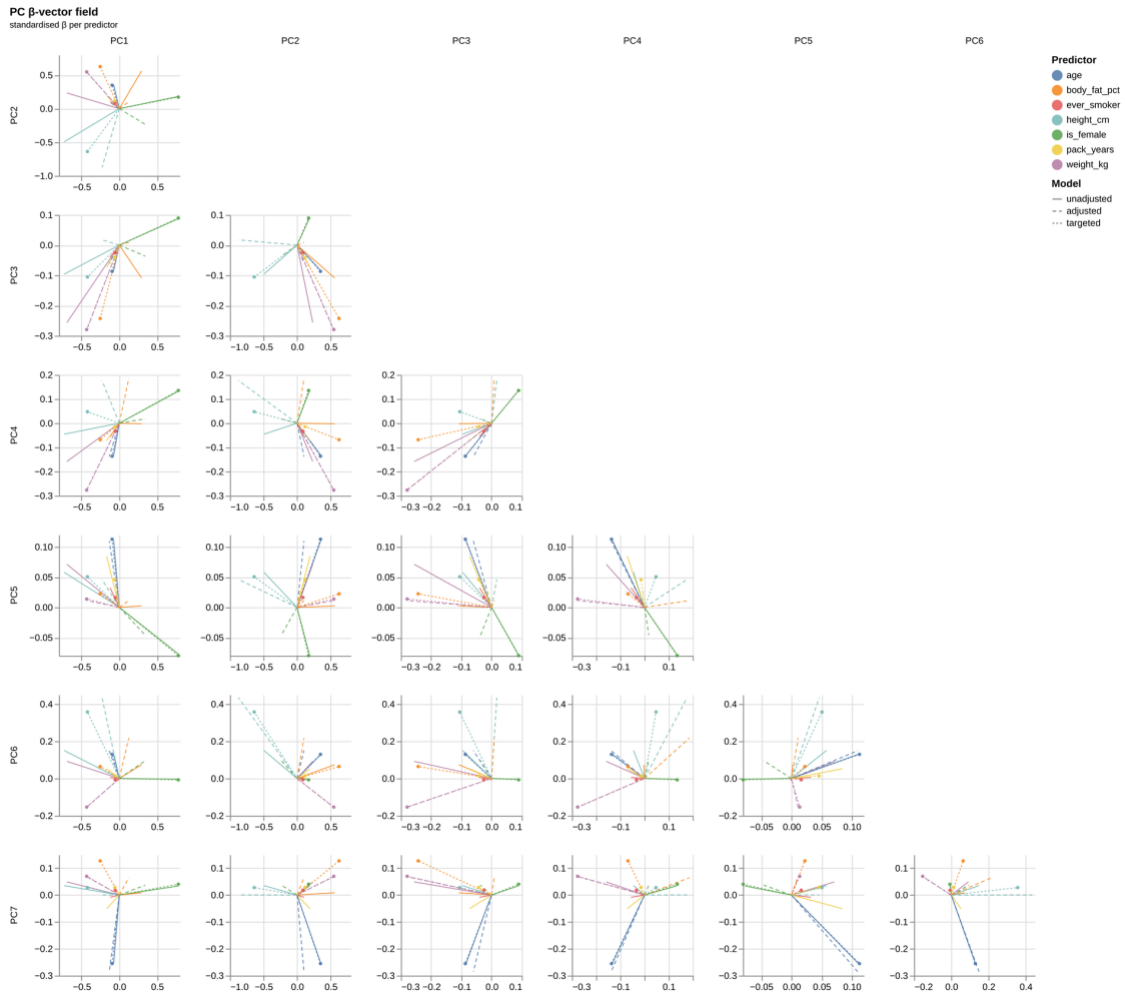
S4 PC score associations

S4.1 Adjusted PC association map



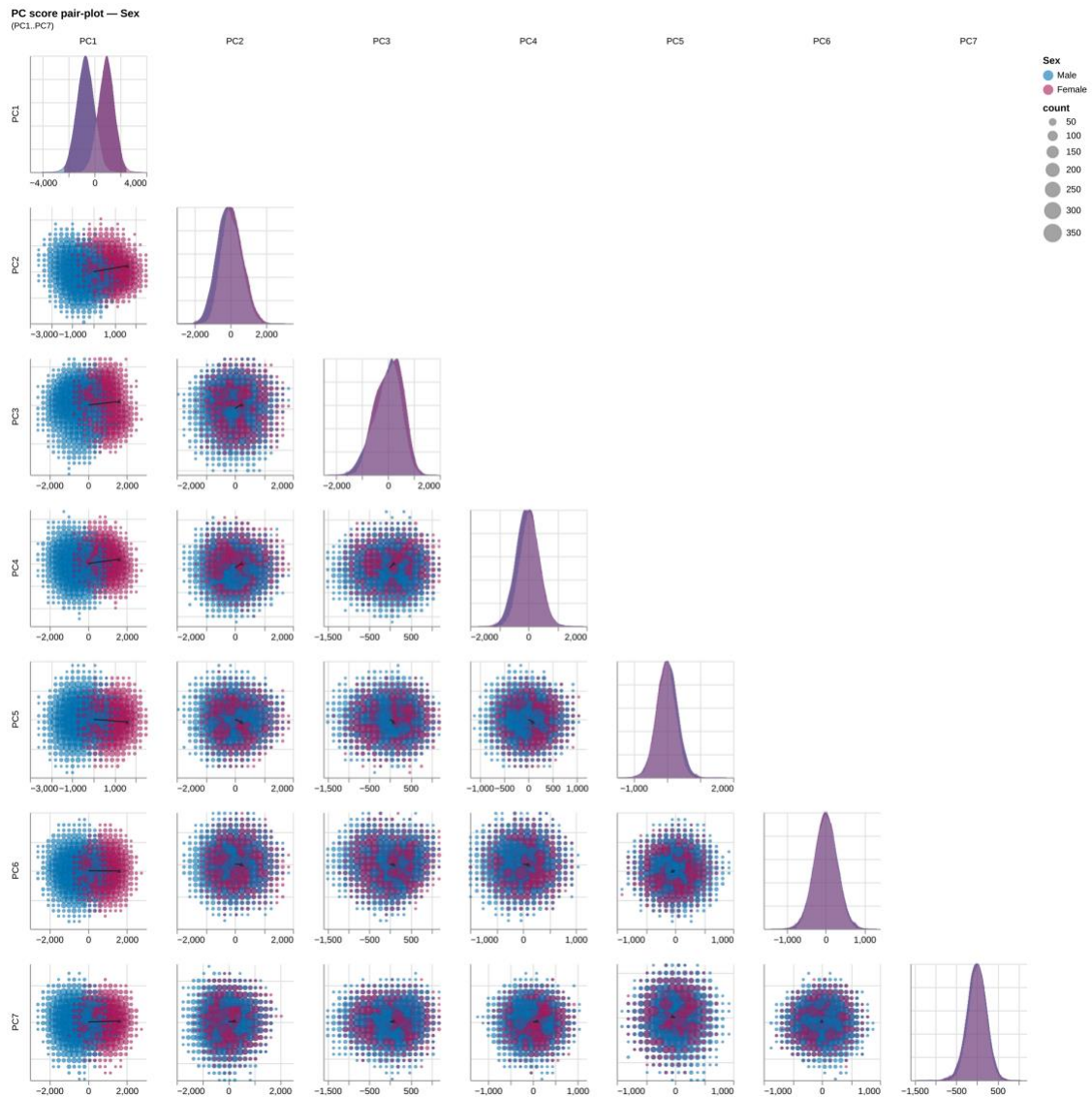
Supplementary Figure 14: Adjusted (multivariable) associations with the leading PC scores. Standardised β from the per-PC adjusted (HC3) OLS against the seven covariates – the mutually-adjusted counterpart to the [main-text targeted map](#).

S4.2 PC β -vector field

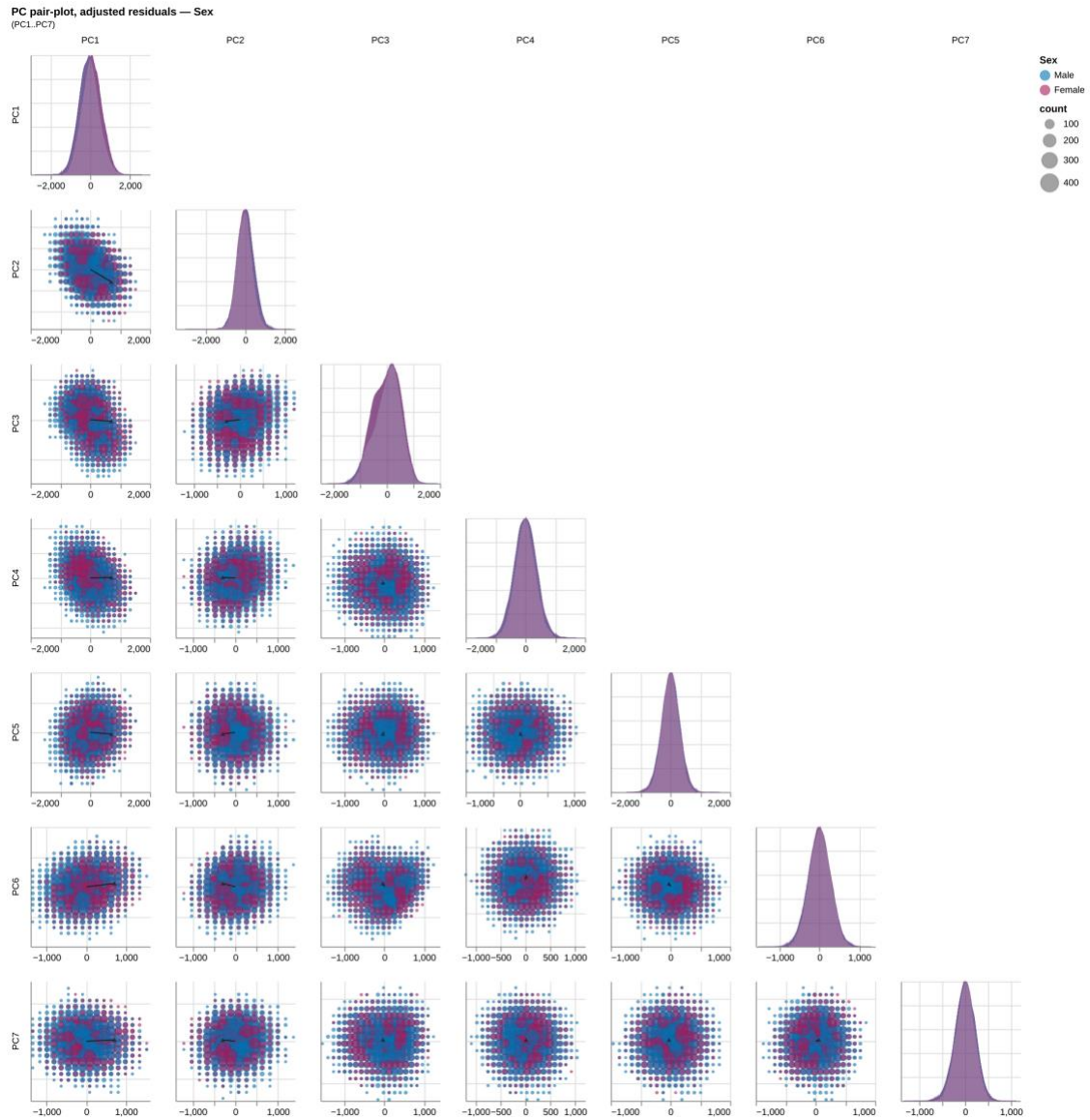


Supplementary Figure 15: **Per-PC β -vector field.** Each predictor's standardised β as an arrow in PC space; solid arrows are the unadjusted (marginal) effect, dashed arrows the adjusted effect. Complements the magnitude ranking of the heatmaps with effect direction.

S4.3 Marginal and adjusted PC sex pair-plots

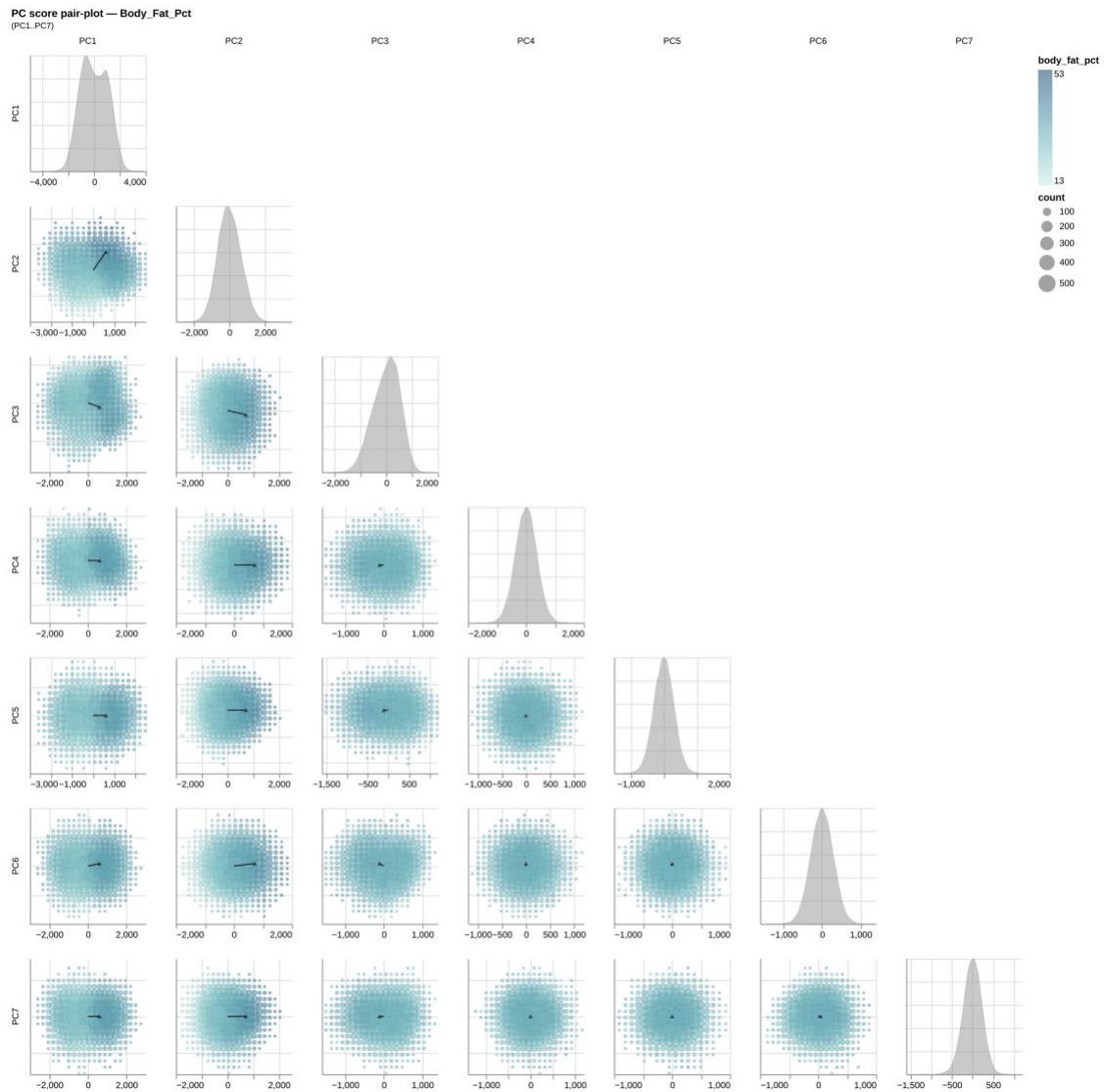


Supplementary Figure 16: *PC score pair-plot by sex (marginal)*. The unadjusted by-sex pair-plot; companion to the targeted view in the main text.



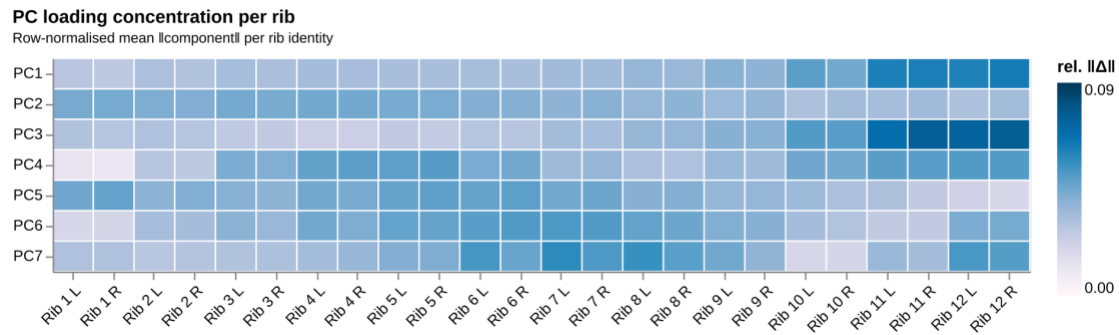
*Supplementary Figure 17: **PC score pair-plot by sex (adjusted).** PC scores residualised on all covariates except sex; the PC1 separation shrinks relative to the marginal and targeted views, visualising the body-composition mediation of the sex difference.*

S4.4 PC scores pair-plot by body fat



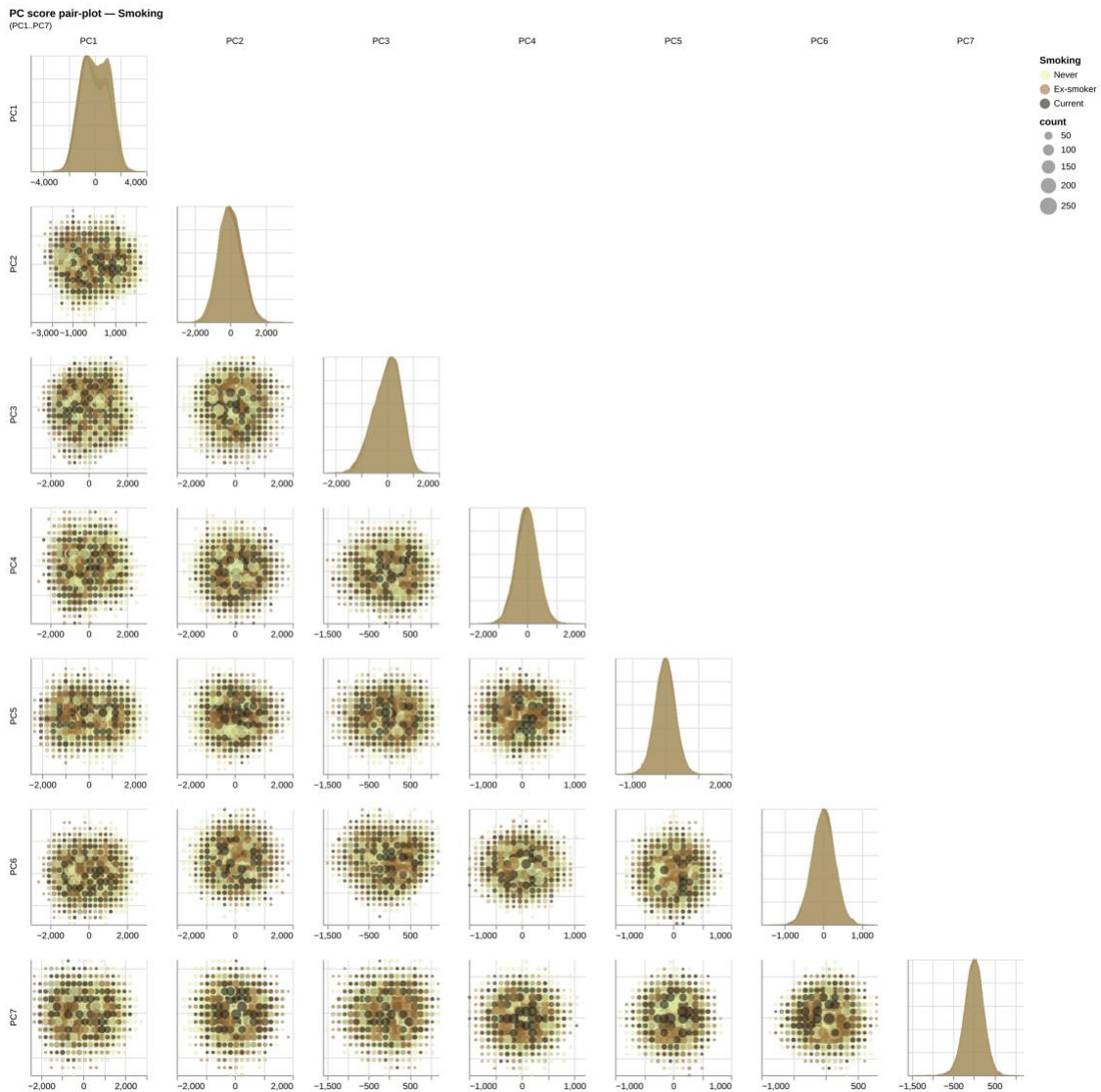
Supplementary Figure 18: PC score pair-plot coloured by body-fat percentage. Off-diagonal bins coloured by mean body-fat percentage, showing the body-composition gradient aligning with the leading modes.

S4.5 Per-rib PC loadings



Supplementary Figure 19: **Per-rib loadings of the leading principal components.** Loadings of each PC onto each of the 24 rib positions.

S4.6 PC scores pair-plot by smoking status



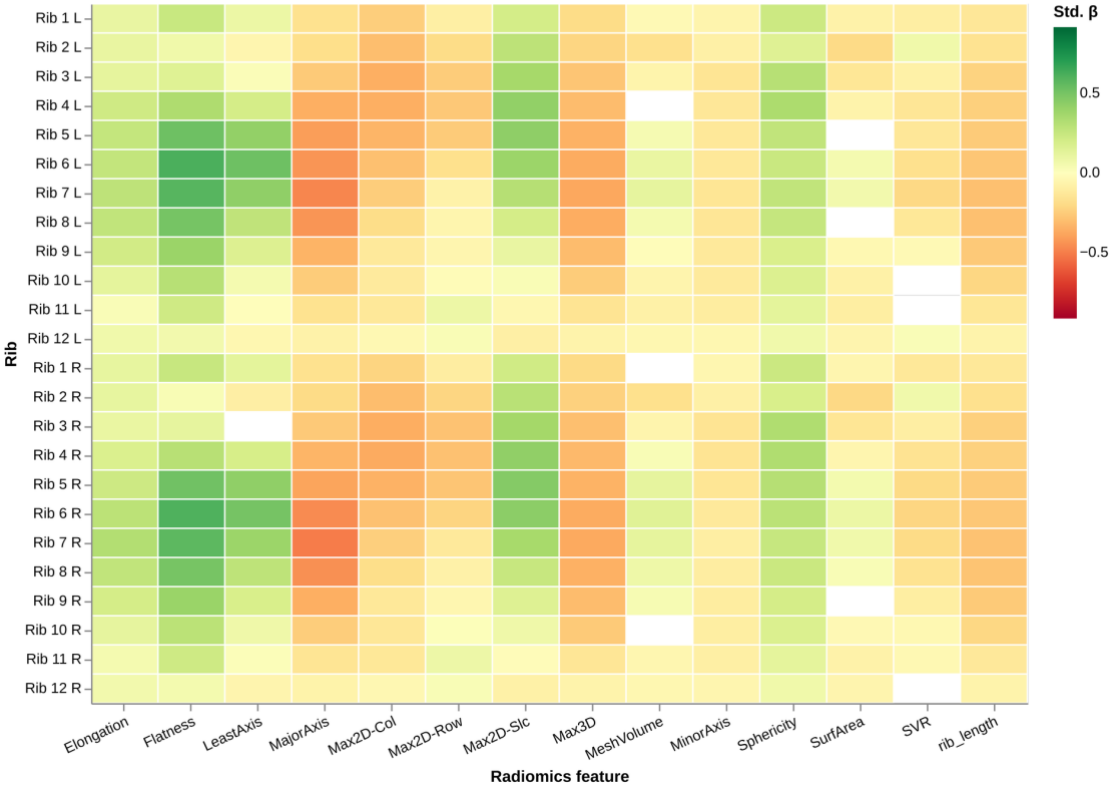
Supplementary Figure 20: PC score pair-plot stratified by smoking status. Companion view to the [by-sex pair-plot](#) in the main text. Smoking is shown here as the descriptive three-level variable (Never / Ex-smoker / Current); the association models use the binary ever-smoker contrast. Marginal density panels show the per-PC distribution per smoking-status group.

S5 Anatomical interpretation of principal modes

S5.1 PC-anchored anatomical maps (PC4 onwards)

PC × radiomics — PC_4

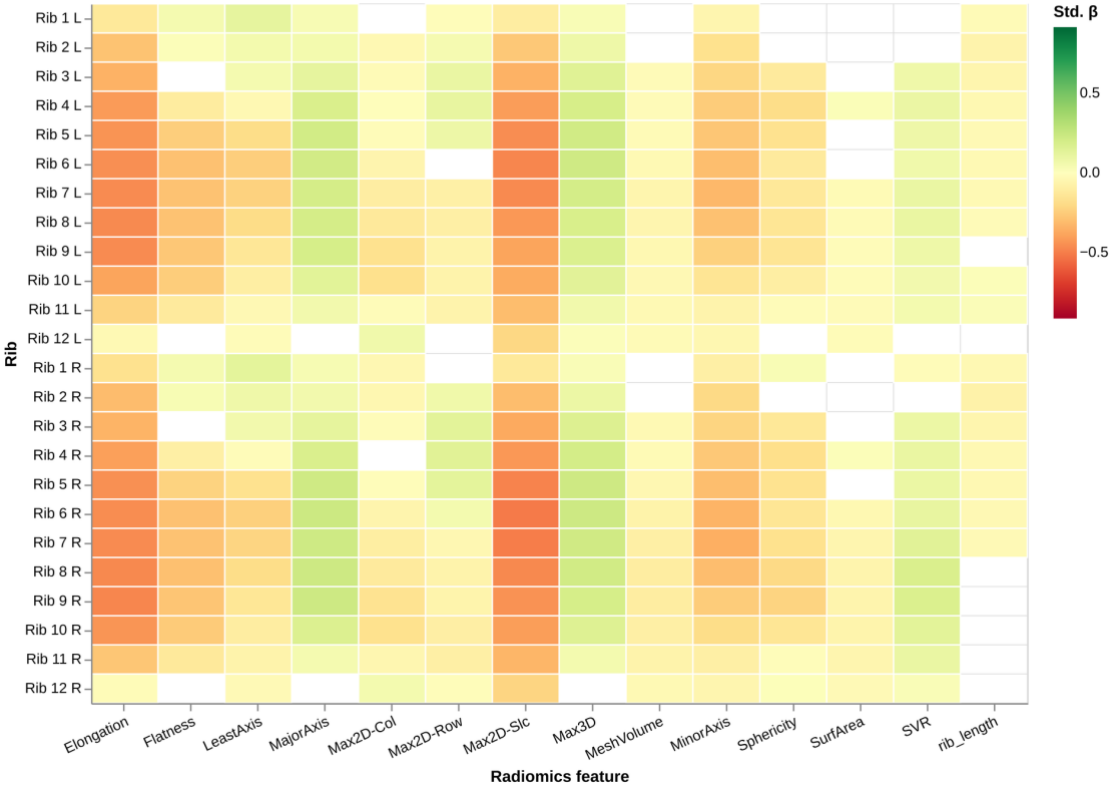
Standardised β across 24 ribs × 14 features - FDR-masked at $q \geq 0.05$ (empty cells)



(a) PC4

PC × radiomics — PC_5

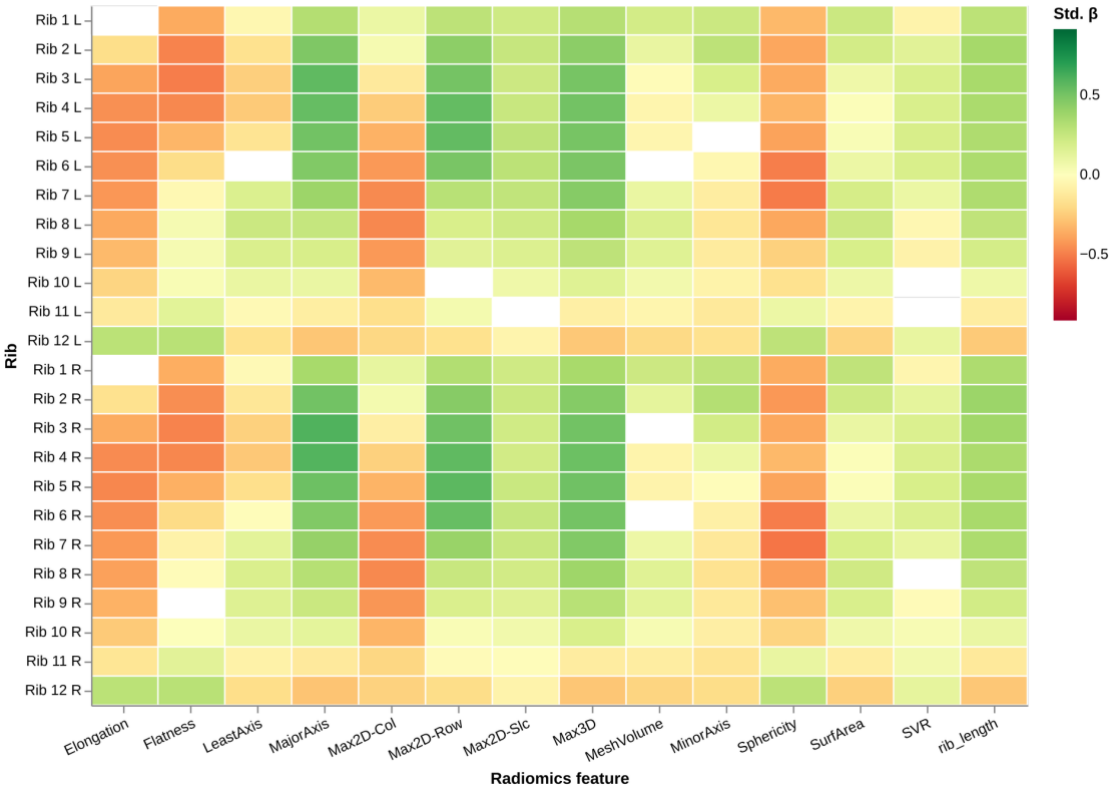
Standardised β across 24 ribs × 14 features - FDR-masked at $q \geq 0.05$ (empty cells)



(b) PC5

PC × radiomics — PC_6

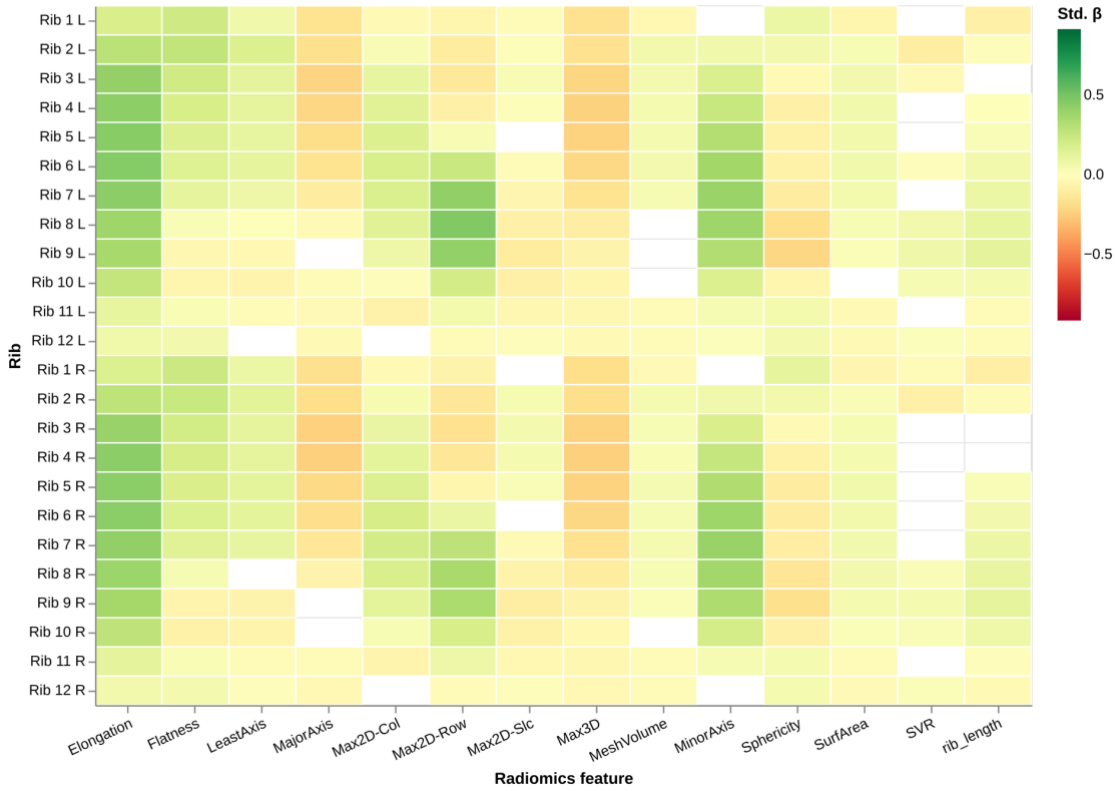
Standardised β across 24 ribs × 14 features - FDR-masked at $q \geq 0.05$ (empty cells)



(c) PC6

PC × radiomics — PC_7

Standardised β across 24 ribs × 14 features · FDR-masked at $q \geq 0.05$ (empty cells)



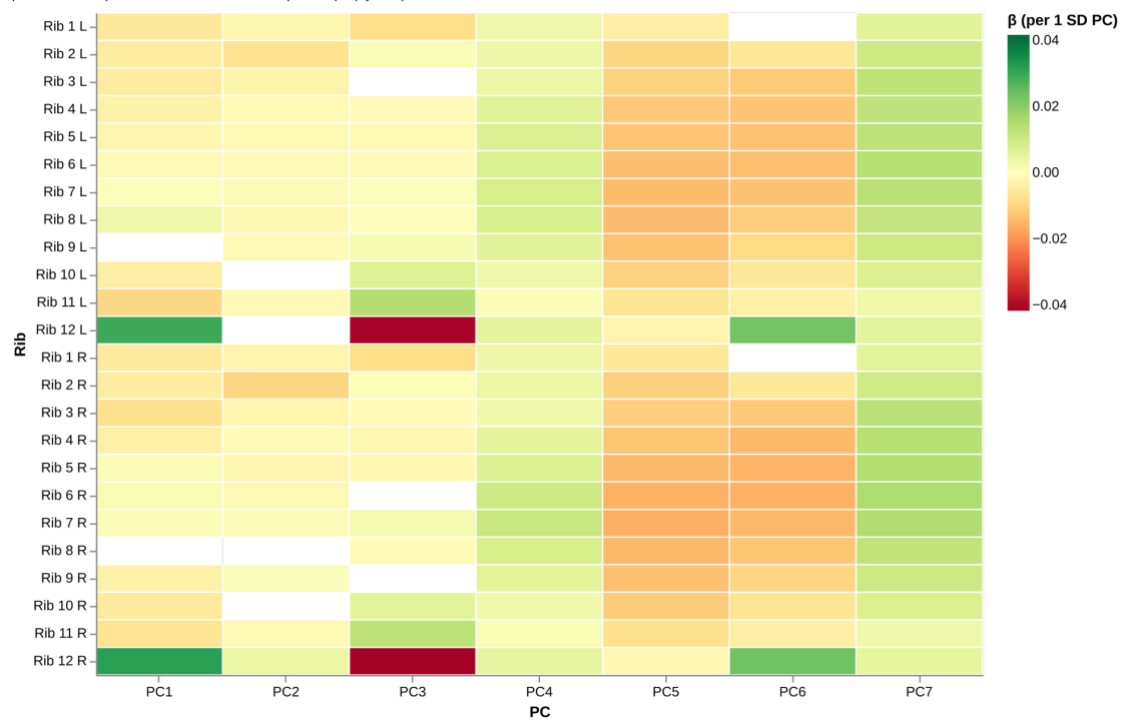
(d) PC7

Supplementary Figure 21: **Per-PC anatomical cross-walk (PC4 onwards)**. Standardised slope of every (rib position × shape descriptor) pair on each PC score (cohort-z-scored), rendered as 24-row × 14-column heatmaps. Print shows PC4–PC7; the interactive supplement browses all 28 modes. PC1–PC3 are in the [main-text anatomical maps](#).

S5.2 Descriptor-anchored anatomical maps

PC × radiomics — Elongation

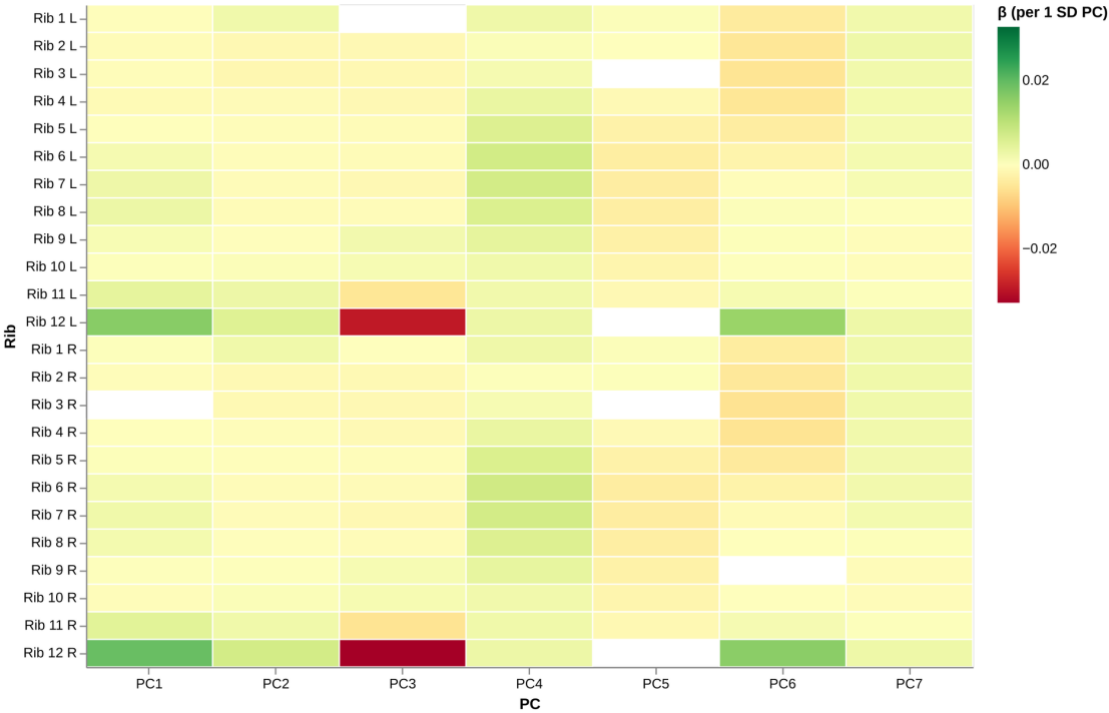
β in native units per 1 SD PC · FDR-masked at q ≥ 0.05 (empty cells)



(a) Elongation

PC × radiomics — Flatness

β in native units per 1 SD PC · FDR-masked at $q \geq 0.05$ (empty cells)



(b) Flatness

PC × radiomics — LeastAxis

β in mm per 1 SD PC · FDR-masked at $q \geq 0.05$ (empty cells)



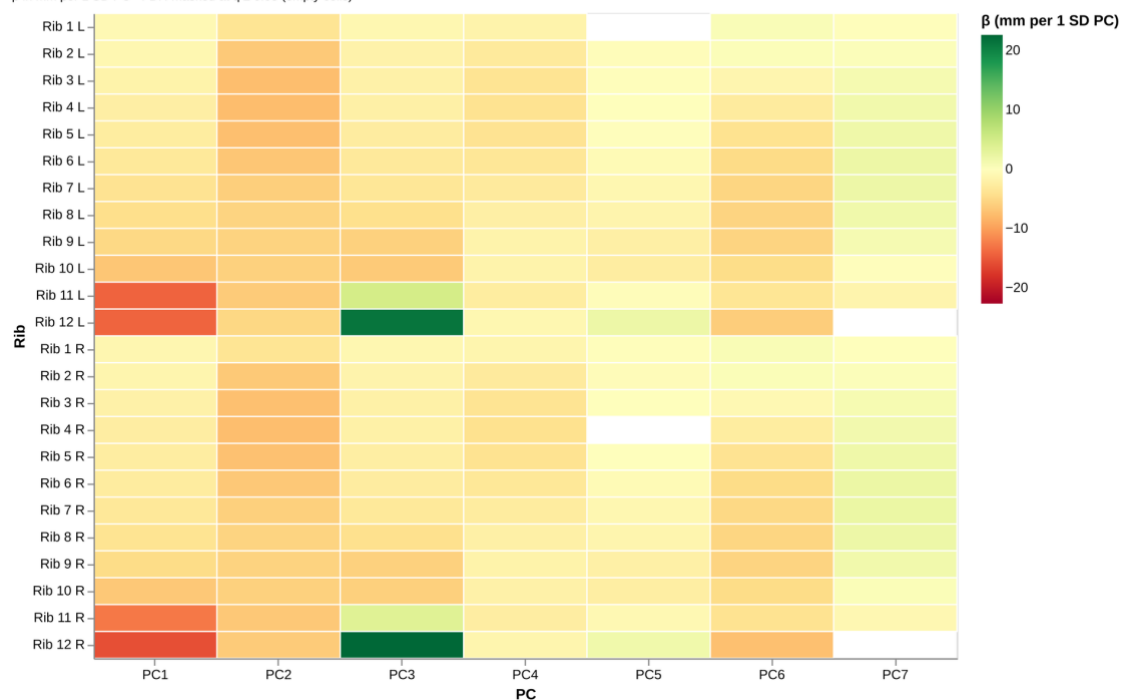
(c) *LeastAxis*



(d) *MajorAxis*

PC × radiomics — Max2D-Col

β in mm per 1 SD PC - FDR-masked at q ≥ 0.05 (empty cells)



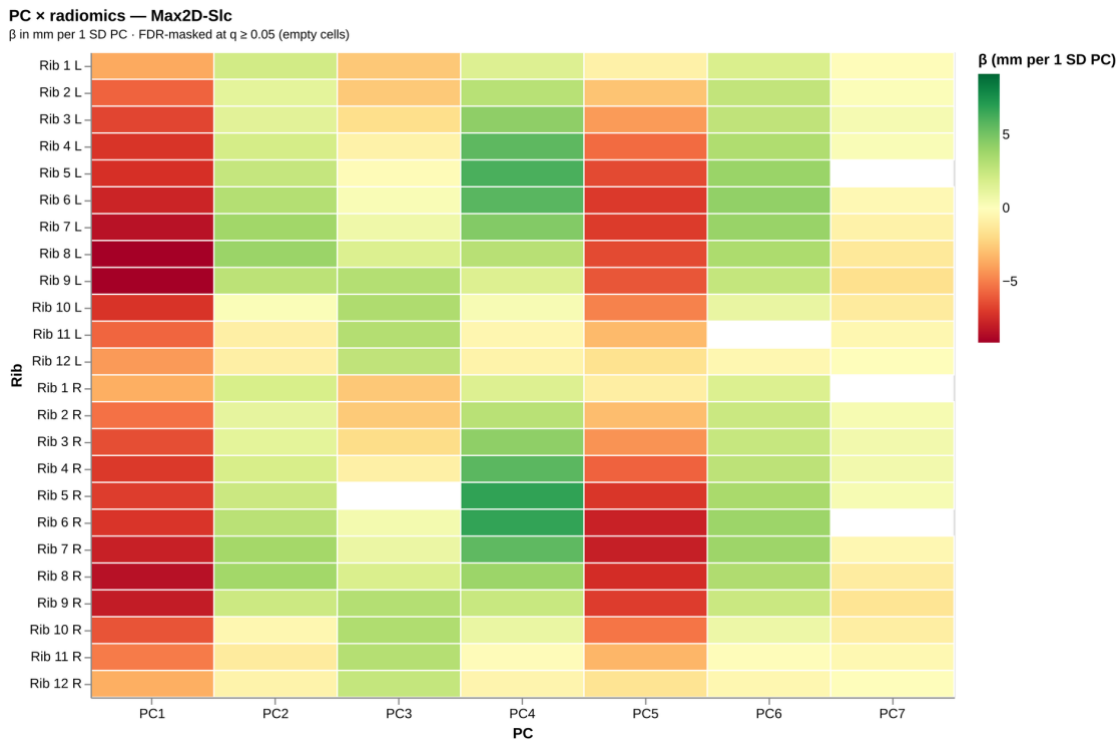
(e) *Max2D-Col*

PC × radiomics — Max2D-Row

β in mm per 1 SD PC - FDR-masked at q ≥ 0.05 (empty cells)



(f) *Max2D-Row*



(g) *Max2D-Slc*

PC × radiomics — Max3D

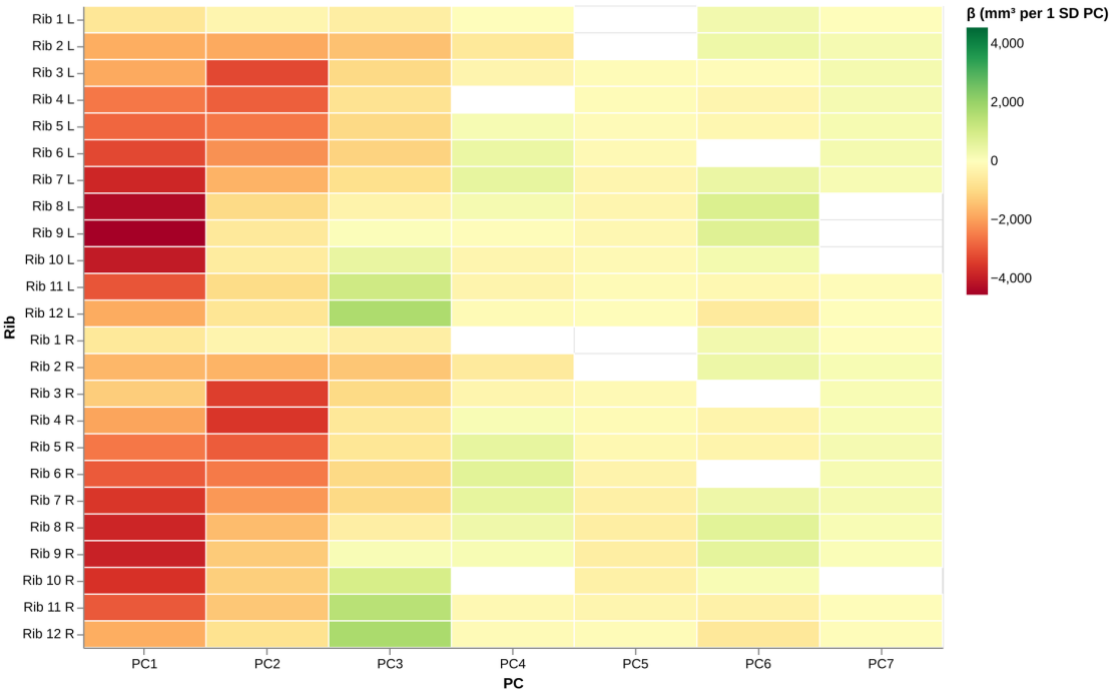
β in mm per 1 SD PC · FDR-masked at $q \geq 0.05$ (empty cells)



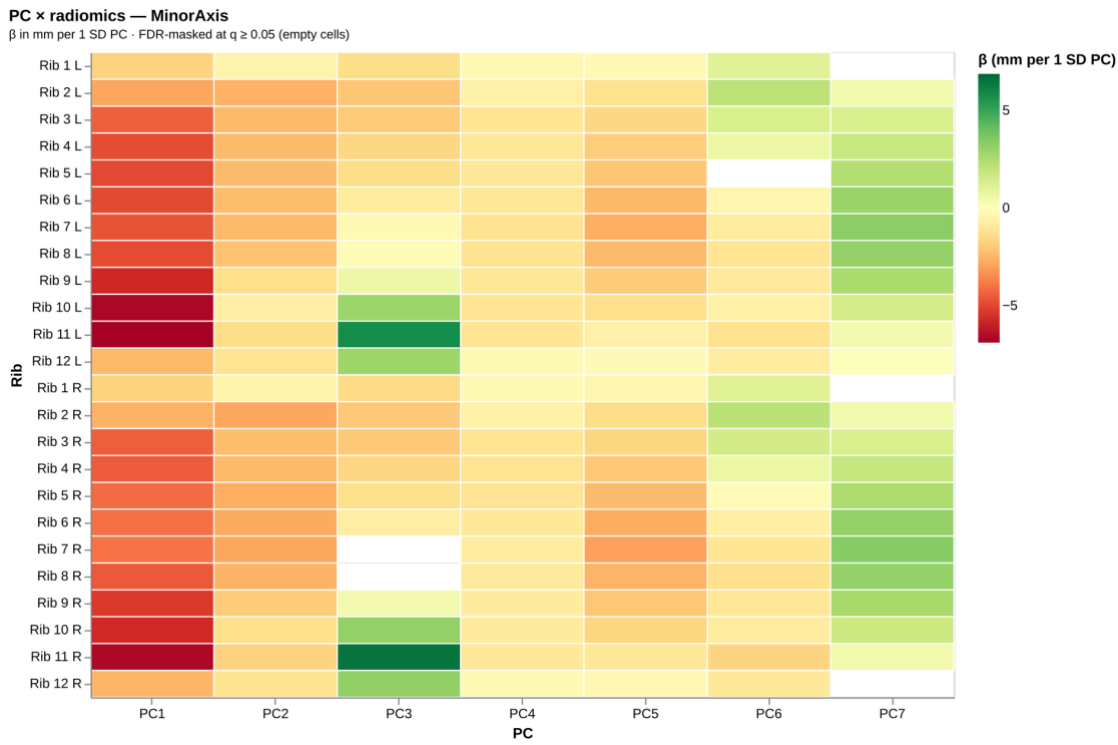
(h) Max3D

PC × radiomics — MeshVolume

β in mm³ per 1 SD PC · FDR-masked at $q \geq 0.05$ (empty cells)



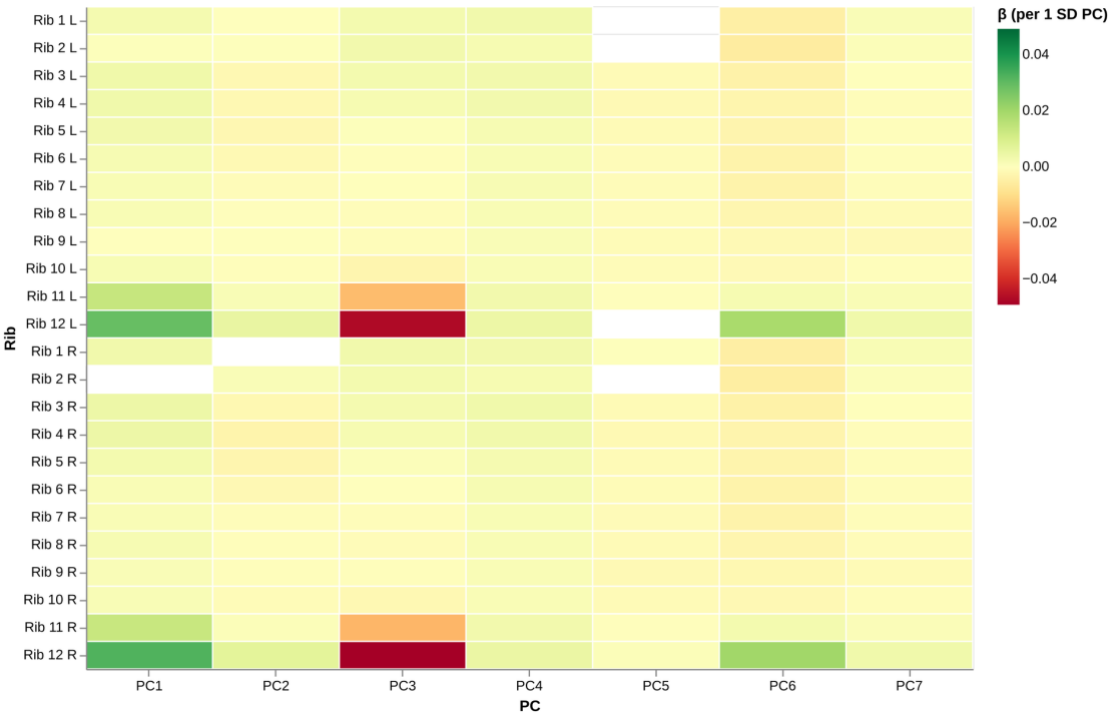
(i) *MeshVolume*



(j) *MinorAxis*

PC × radiomics — Sphericity

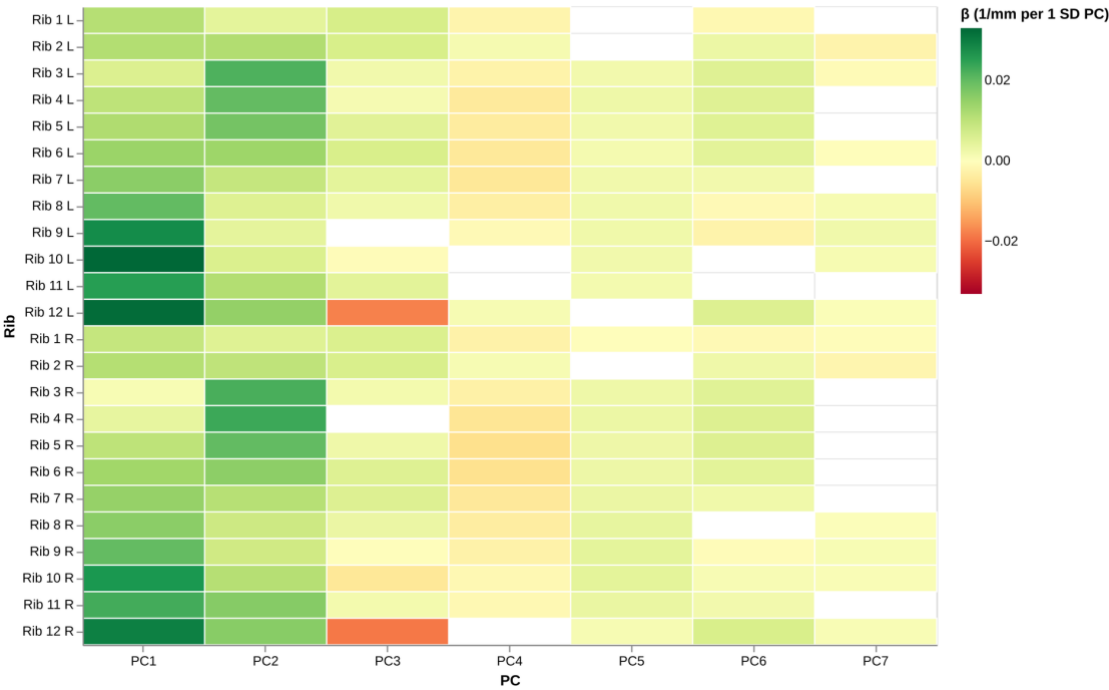
β in native units per 1 SD PC · FDR-masked at $q \geq 0.05$ (empty cells)



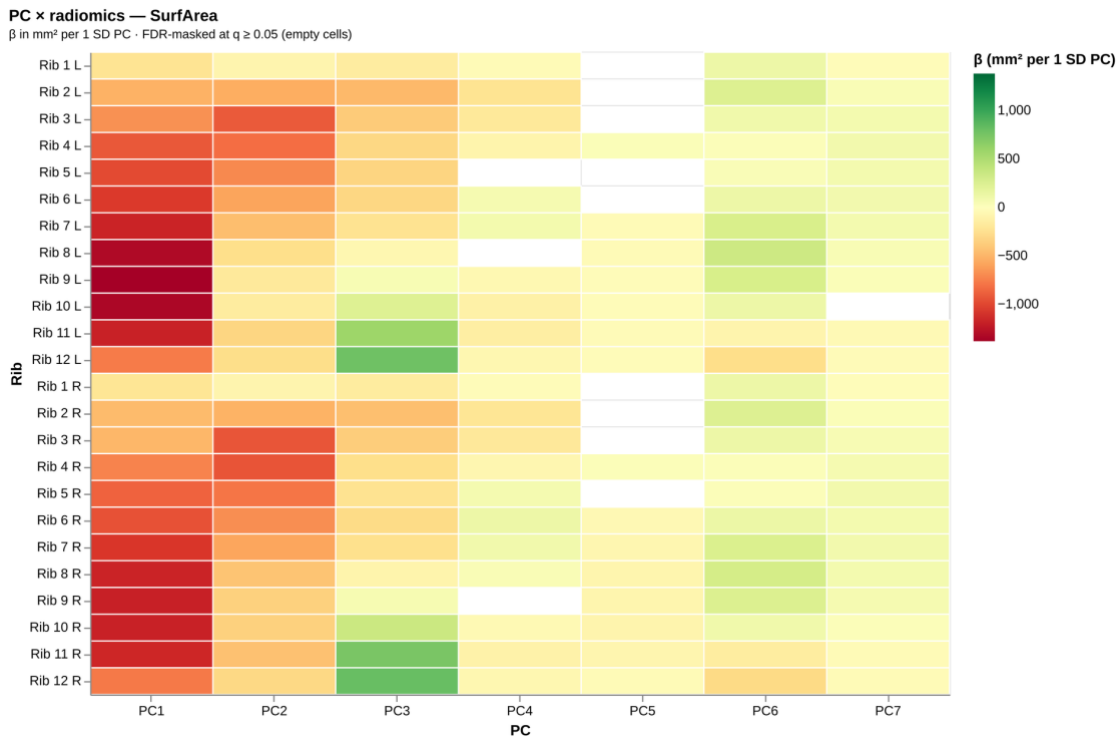
(k) Sphericity

PC × radiomics — SVR

β in 1/mm per 1 SD PC · FDR-masked at $q \geq 0.05$ (empty cells)



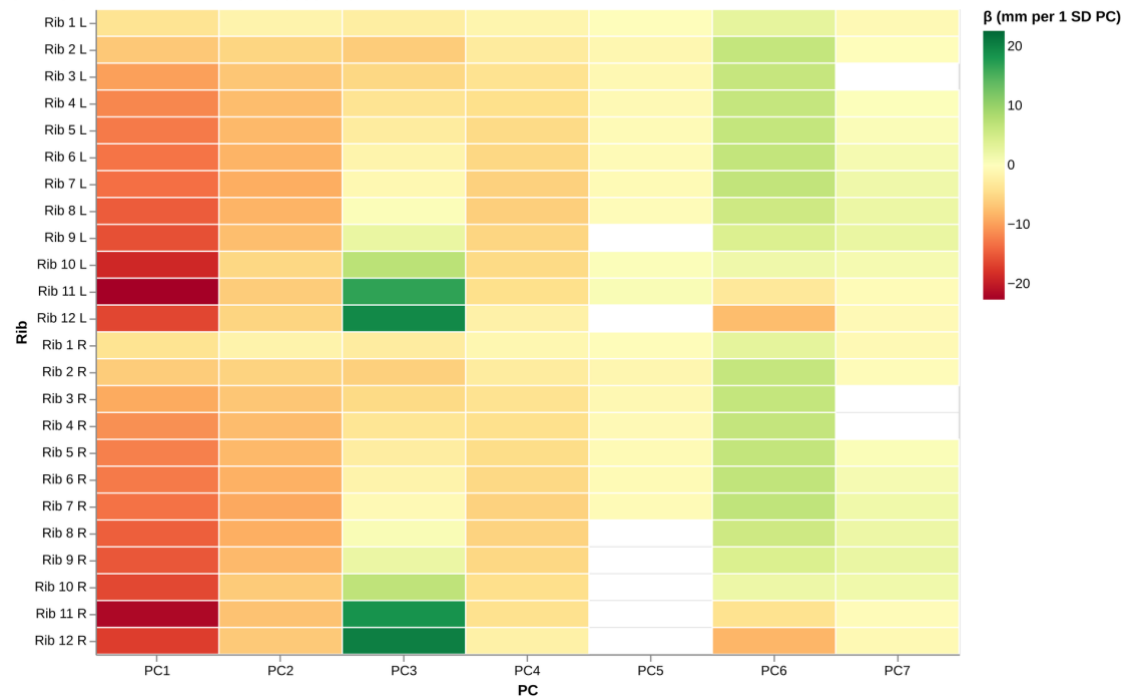
(l) SVR



(m) SurfArea

PC × radiomics — rib_length

β in mm per 1 SD PC : FDR-masked at q ≥ 0.05 (empty cells)



(n) rib_length

Supplementary Figure 22: **Per-descriptor anatomical cross-walk**. Standardised slope of one shape descriptor at every rib position on each PC score, one panel per descriptor. Print shows all 14 descriptors; the interactive supplement browses the same set.