Online Resource 1

Supplementary methods, figures and tables

Acute kidney injury is associated with less reversal of urinary metabolomic changes after coronary artery bypass grafting

Journal: Metabolomics

Donglin Xu[†], Zengkun Liu[†], Haixuan Zhao, Ang Li, Qiduo Chen, Jing Liu and Hourong Sun[*]

Hourong Sun; Department of Cardiovascular Surgery, Qilu Hospital of Shandong University, 107 Wenhua Xi Road, Jinan, Shandong 250012, China; Email: sunhourong2016@126.com; Tel.: +86-185-6008-6790

Donglin Xu and Zengkun Liu contributed equally.

Supplementary methods

S1.1. Study design and patient-level analysis

We reanalyzed public, de-identified molecular data. Patients, rather than samples or cells, were the units in all permutation tests, bootstrap procedures, delete-one analyses and regression models. The longitudinal source contained 159 on-pump CABG patients and 622 available samples. Requiring exactly one sample at each of four timepoints retained 145 patients (50 AKI and 95 non-AKI), 580 samples and 4,048 shared features. Fourteen incomplete series were excluded (5 AKI and 9 non-AKI).

S1.2. Source AKI labels

The source publication classified 55 of 159 patients as AKI and 104 as non-AKI. Its public article record, supporting information and processed matrices did not provide the measurements, operational criteria and diagnostic timing needed to independently reconstruct the labels. We retained those labels rather than assigning new AKI definitions or stages. The independent 55-patient urine dataset also supplied AKI/non-AKI labels; its severity labels were used as reported by that source.

S1.3. Preprocessing for the within-patient comparisons

The return-to-baseline calculations used T1, T3 and T4 (preoperative, approximately 20 h and approximately 44 h). T2, at approximately 6 h, was used in the trajectory analysis. Preprocessing for the within-patient comparisons used the full 622-sample source matrix. For each feature containing zeros, half the minimum positive value was added to all of that feature's values; features without zeros received no pseudocount. Values were $\log_2$-transformed, centered by their median across all 622 samples and divided by MAD×1.4826. If that scale was non-finite or non-positive, IQR/1.349, sample standard deviation and 1 were used in succession. No non-finite entries were present (0 of 2,517,856 cells), so no imputation was required. This cohort-wide preprocessing was distinct from the training-only preprocessing used for the trajectory.

S1.4. Residual ratio and reversal fidelity

For each patient, let a=T3−T1 denote the earlier change, r=T4−T3 the subsequent change, and q=T4−T1 the remaining difference from baseline, using the scaled feature vectors. Residual ratio was $\|q\|_2/\|a\|_2$. Reversal fidelity was the cosine between r and −a. The code added 10^{-12} to the denominator of each expression for numerical stability. A residual ratio below 1 indicates that the final profile is closer to baseline than the intermediate profile. A larger reversal fidelity indicates a later change more closely aligned with the direction back to baseline.

Effects were AKI-minus-non-AKI median differences. Cliff's delta summarized between-group rank separation. Tests used 100,000 patient-label permutations of the rank-sum statistic, with the upper tail for residual ratio and the lower tail for reversal fidelity. Benjamini–Hochberg correction covered these two tests only. Confidence intervals used 20,000 stratified patient bootstraps of the calculated scores. Every patient was also omitted once. The positive residual-ratio difference occurred in 99.995% of bootstrap draws and all 145 delete-one analyses; the negative reversal-fidelity difference occurred in all bootstrap draws and all 145 delete-one analyses. These summaries come from the same patients and samples.

S1.5. Leave-one-patient-out trajectory and time coordinate

All four samples from one complete patient were held out in each of 145 repetitions. For each feature in the primary analysis, half its minimum positive training-sample value was added to every training and held-out value before $\log_2$ transformation. Feature medians and robust scales were estimated using only the other 144 complete patients. The scale fallback sequence was MAD×1.4826, IQR/1.349, sample standard deviation and 1.

Within each training patient, the T1 feature vector was subtracted from the vectors at every timepoint. Means across training patients were calculated at 0, 6, 20 and 44 h, and neighboring means were joined with line segments. Each held-out profile was also expressed relative to that patient's own T1. The nearest point across the three segments was found by Euclidean projection, restricting each segment coefficient to [0, 1]. The minimum distance was off-trajectory distance. The segment coefficient was used to interpolate a molecular time between the segment endpoints. Molecular time cannot extend outside 0–44 h, and the baseline time and distance are both zero by construction.

AKI labels were not used in constructing the trajectory or coordinates. The training patients included both AKI and non-AKI participants; the path was therefore a cohort-average reference, not a normal-recovery curve. The procedure was repeated using centered, L2-normalized within-sample average ranks and using sample-median centering after log transformation and before feature-wise robust scaling. The trajectory-distance contrasts used 100,000 upper-tail patient-label permutations and 10,000 stratified bootstraps of fixed patient scores. Hedges g intervals also used 10,000 patient bootstrap draws. Delete-one analyses omitted one score at a time rather than rebuilding every trajectory.

S1.6. Time-position and magnitude sensitivity analyses

Lag at T4 was defined as 44 minus the assigned molecular time. A positive AKI-minus-non-AKI contrast would indicate that AKI profiles were assigned earlier positions on the fitted path. The primary mean difference was −0.691 h (one-sided permutation p = 0.7158; 10,000-bootstrap 95% CI −2.811 to 1.590). All 145 delete-one estimates were negative in each profile representation. This analysis did not support greater lag in AKI, although the fitted time coordinate was bounded and was not a measured recovery duration.

State_norm in the code is the L2 norm of the held-out T4−T1 vector. The magnitude sensitivity used the already calculated leave-one-patient-out distance and State_norm scores. For each patient, log1p(off-trajectory distance) was regressed on log1p(State_norm) using the other 144 patient scores. The held-out residual was the observed log1p distance minus its predicted value. Group labels were added after residual calculation. Thus, the regression step left out the scored patient; it did not rebuild the other patients' original trajectory fits inside this second step.

The primary residual contrast was small and inconclusive (Hedges g = 0.265; one-sided permutation p = 0.064; 95% stratified-bootstrap CI −0.004 to 0.032). All 145 delete-one effects were positive, but the bootstrap interval crossed zero. This sensitivity does not establish an AKI-associated distance difference independent of the overall size of the change from baseline.

S1.7. Exploratory joint display of the two distance measures

Residual ratio and T4 off-trajectory distance were median-centered and robustly scaled for the joint display. Their Pearson correlation was 0.4643 and Spearman correlation was 0.5721. The minor-axis variance fraction was 0.2056 and the covariance condition number was 3.86. Splitting both coordinates at zero on this scaled display formed four regions containing 51, 21, 21 and 52 patients, with 0, 0, 15 and 35 AKI cases, respectively. These thresholds correspond to the cohort medians; they are not clinical cutoffs.

The global 2×4 association used 100,000 patient-label permutations ($p = 1.00\times10^{-5}$; Cramér's V=0.721). A joint logistic model included both scaled coordinates. The residual-ratio coefficient had OR 0.716 (95% CI 0.430 to 1.192; p = 0.199), and the off-trajectory coefficient had OR 6.16 (95% CI 3.20 to 11.85; $p = 5.06\times10^{-8}$). This model adjusted each coordinate only for the other coordinate, not for State_norm or clinical confounders. The display does not establish clinical subtypes, validated prediction or a magnitude-independent association.

S1.8. Independent early postoperative urine cohort

The source "fill0.5min" worksheet contained 55 patients (29 AKI and 26 non-AKI), each sampled within 4 h after surgery, and 1,384 features. Quality checks excluded features with more than 20% non-finite values or zero variance.

Remaining non-finite values would have been replaced by the feature median, followed by removal of any feature constant after imputation. In the actual analyzed matrix, all 76,120 values were finite and no feature was removed.

For the primary analysis, each feature was median-centered across the cohort and scaled by MAD×1.4826, with IQR/1.349, sample standard deviation and 1 as successive fallbacks. The sample-centered analysis first subtracted each patient's across-feature median and then applied the same feature scaling. The rank/unit analysis used centered, L2-normalized within-patient average ranks.

For patient i, normative displacement was $D_i=\sqrt{\{mean_j[((z_{ij}-median_ref,_j)/scale_ref,_j)^2]\}}$, where z is the preprocessed feature value. AKI patients were compared with all 26 non-AKI patients; each non-AKI patient was compared with the other 25. Reference scales used MAD×1.4826, then IQR/1.349 and sample standard deviation. A scale still non-finite or non-positive was set to 1 for AKI scores, whereas that feature was omitted from the affected leave-one-out non-AKI score. Rank/unit scores used scale_ref=1, without a second reference-scale division. No AKI patient entered the observed non-AKI reference.

The primary AKI contrast used 100,000 upper-tail patient-label permutations. Reference membership and scores were fully recalculated after each permutation. The 10,000 stratified patient-bootstrap intervals instead used fixed scores and did not rebuild the reference; they are therefore conditional on those calculated scores. This cohort tested early differences from a non-AKI reference, not change from an individual preoperative sample.

S1.9. Source-defined severe-versus-mild AKI comparison

The 29 AKI patients were compared using their source-defined severe (n=13) and mild (n=16) labels and their fixed primary-analysis displacement scores. The effect was the severe-minus-mild mean difference. An upper-tail test used 100,000 permutations of severity labels among the AKI patients. The 95% percentile interval used 10,000 bootstraps, resampling separately within severity groups. Source labels were not independently re-adjudicated. The result was secondary and inconclusive (difference 4.265; Hedges g = 0.614; p = 0.0543; 95% CI −0.533 to 9.431).

S1.10. Patient-level kidney-tissue comparisons

GSE183276 metadata contained 45 patients: 12 AKI, 18 reference and 15 CKD. Eligible cells were source-labeled epithelial cells in proximal, distal, collecting and intermediate tubules. The five source state.l2 labels were degenerative, adaptive - epi, transitioning, cycling and reference. There were 79,336 eligible cells across all 45 patients and 39,756 cells in the 30-patient AKI-versus-reference comparison. Each patient had at least 50 eligible cells.

Counts were aggregated within patients before fractions were calculated. The primary contrast compared the degenerative fraction between 12 AKI and 18 reference kidneys, using 100,000 one-sided patient-label permutations and 10,000 stratified patient bootstraps. The four other state contrasts were two-sided and corrected together by Benjamini–Hochberg. These fractions describe the composition of sampled tubular epithelial cells. The tissue cohort was not restricted to cardiac-surgery-associated AKI and was not paired with the urine cohorts.

S1.11. Scope of the measurements

The 4,048 longitudinal measurements are referred to as molecular features because high-confidence structural identities are limited. The main comparisons describe changes in the measured urinary profiles, not recovery of kidney function or a specified metabolic pathway. The separate urine dataset contains only one early sample, and the kidney atlas provides unpaired tissue observations. Neither dataset reproduces the longitudinal return-to-baseline analysis.

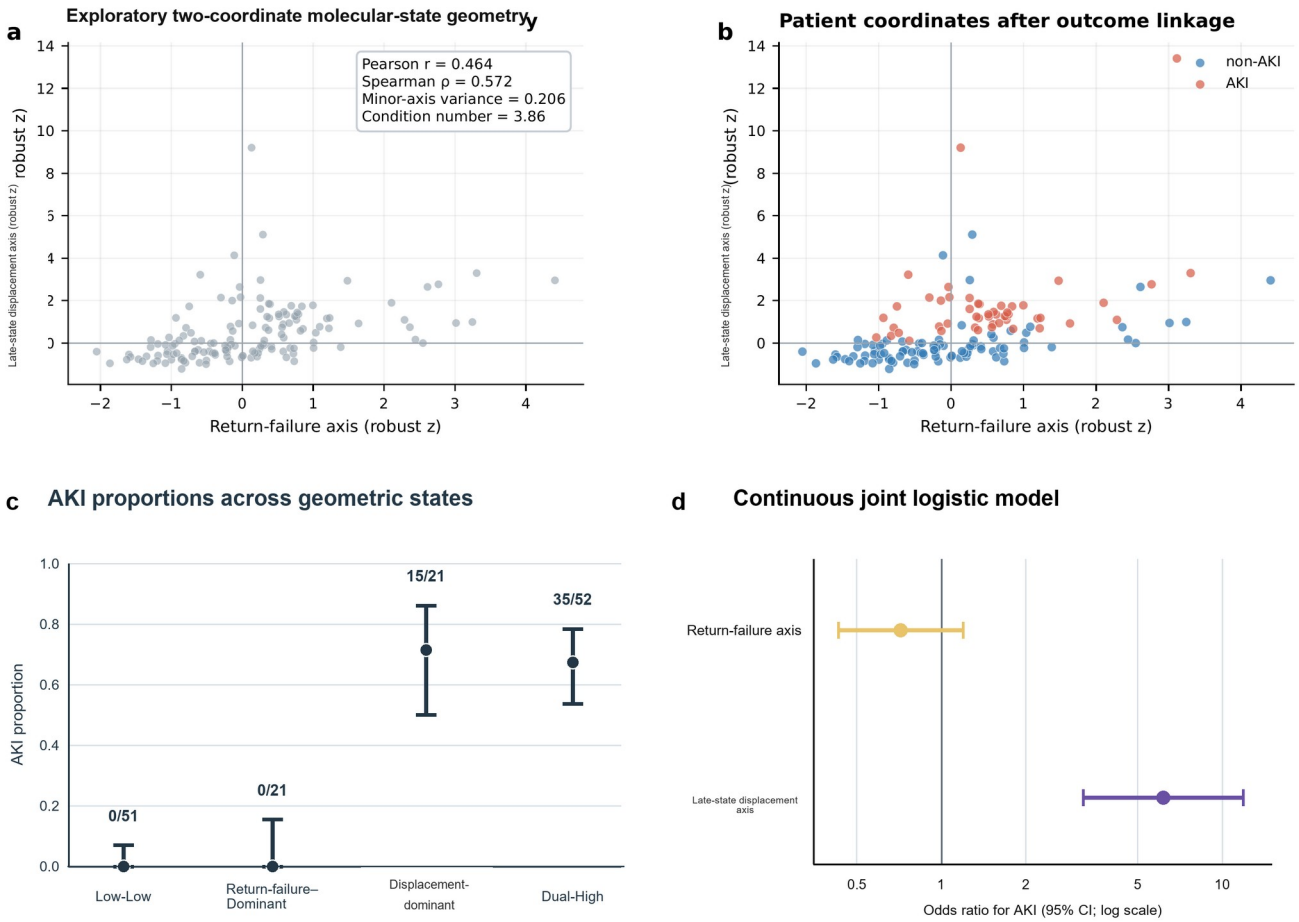


Fig. S1 Exploratory joint display of residual ratio and 44-h distance from the trajectory. a, Median-centered, robustly scaled coordinates before displaying AKI labels (Pearson r=0.464; Spearman ρ = 0.572). b, The same patients colored by source AKI labels. c, AKI counts and proportions in four regions defined by the two cohort medians. d, Odds ratios from a logistic model including both continuous coordinates. “Return-failure axis” in the original panel denotes residual ratio, and “late-state displacement axis” denotes off-trajectory distance. The region names are descriptive plotting labels, not established patient subtypes. The model did not account for overall T4-to-T1 change magnitude or clinical confounders

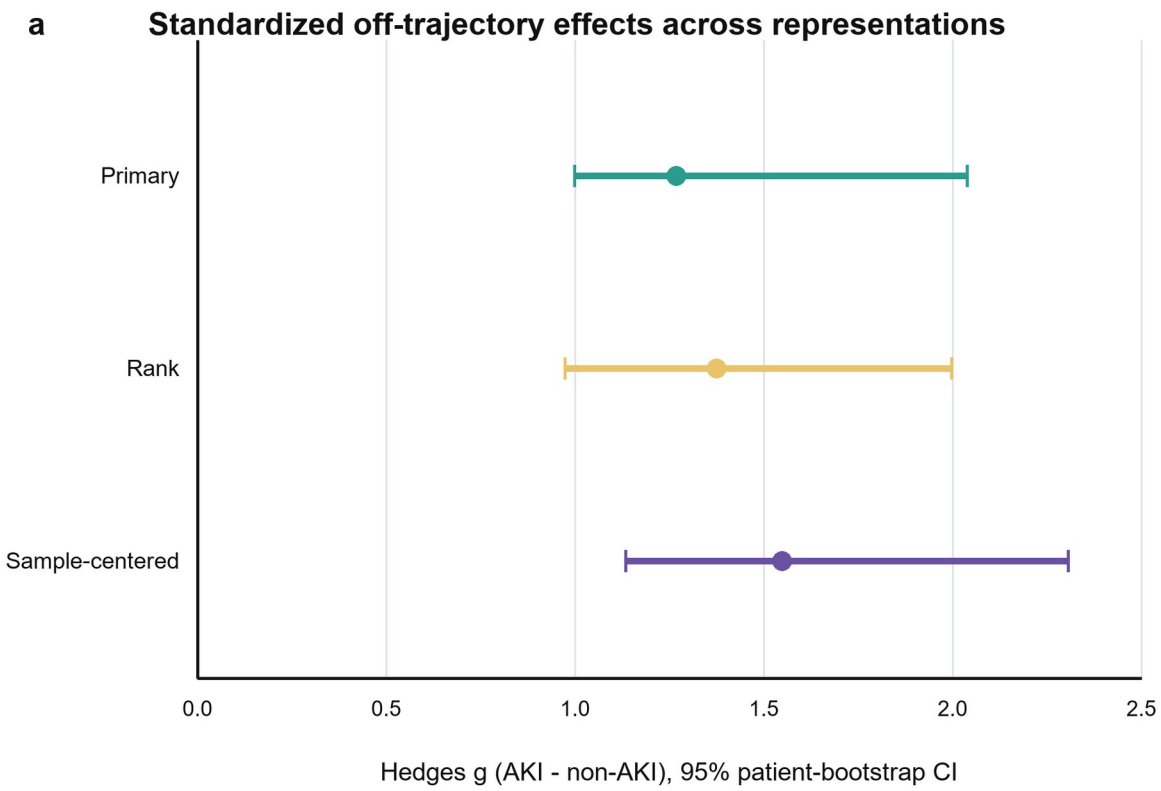


Fig. S2 Sensitivity of the 44-h distance comparison to data preprocessing. Hedges g for the AKI-minus-non-AKI off-trajectory comparison using the primary, within-sample rank and sample-median-centered profiles. Error bars show 95% patient-bootstrap intervals. The underlying patients and outcome contrast are the same in each analysis

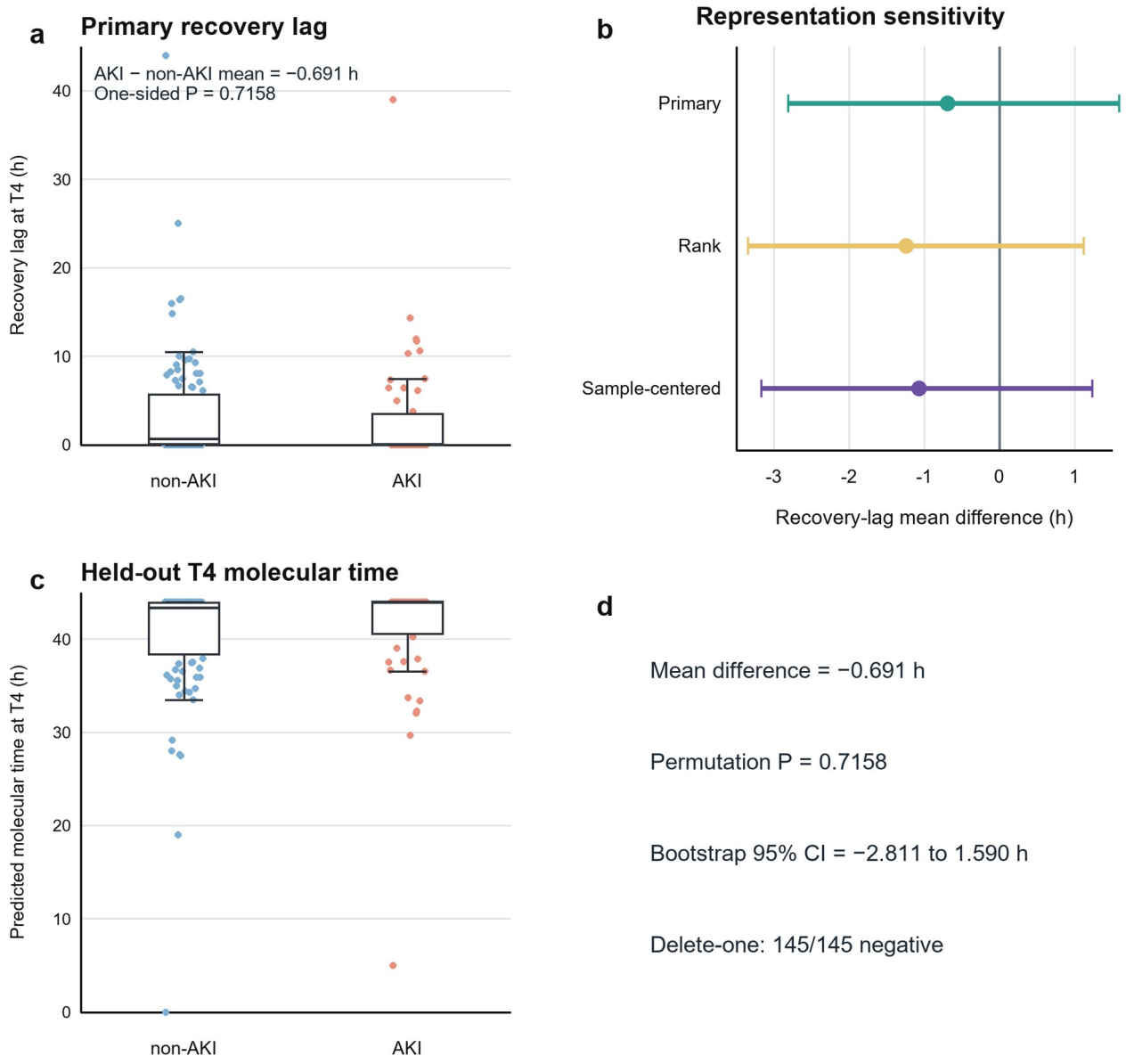


Fig. S3 The fitted time coordinate did not show greater lag in AKI. a, Primary lag values, defined as 44 h minus the time assigned to the nearest point on the path. b, Mean group differences under the three preprocessing approaches. c, Assigned molecular times at T4. d, Primary estimate and uncertainty: AKI-minus-non-AKI mean lag difference −0.691 h; one-sided patient-label permutation p = 0.7158; 95% interval from 10,000 stratified patient bootstraps −2.811 to 1.590 h. All 145 delete-one estimates were negative in each preprocessing approach. The “recovery lag” label refers to this bounded fitted coordinate, not an observed clinical recovery duration

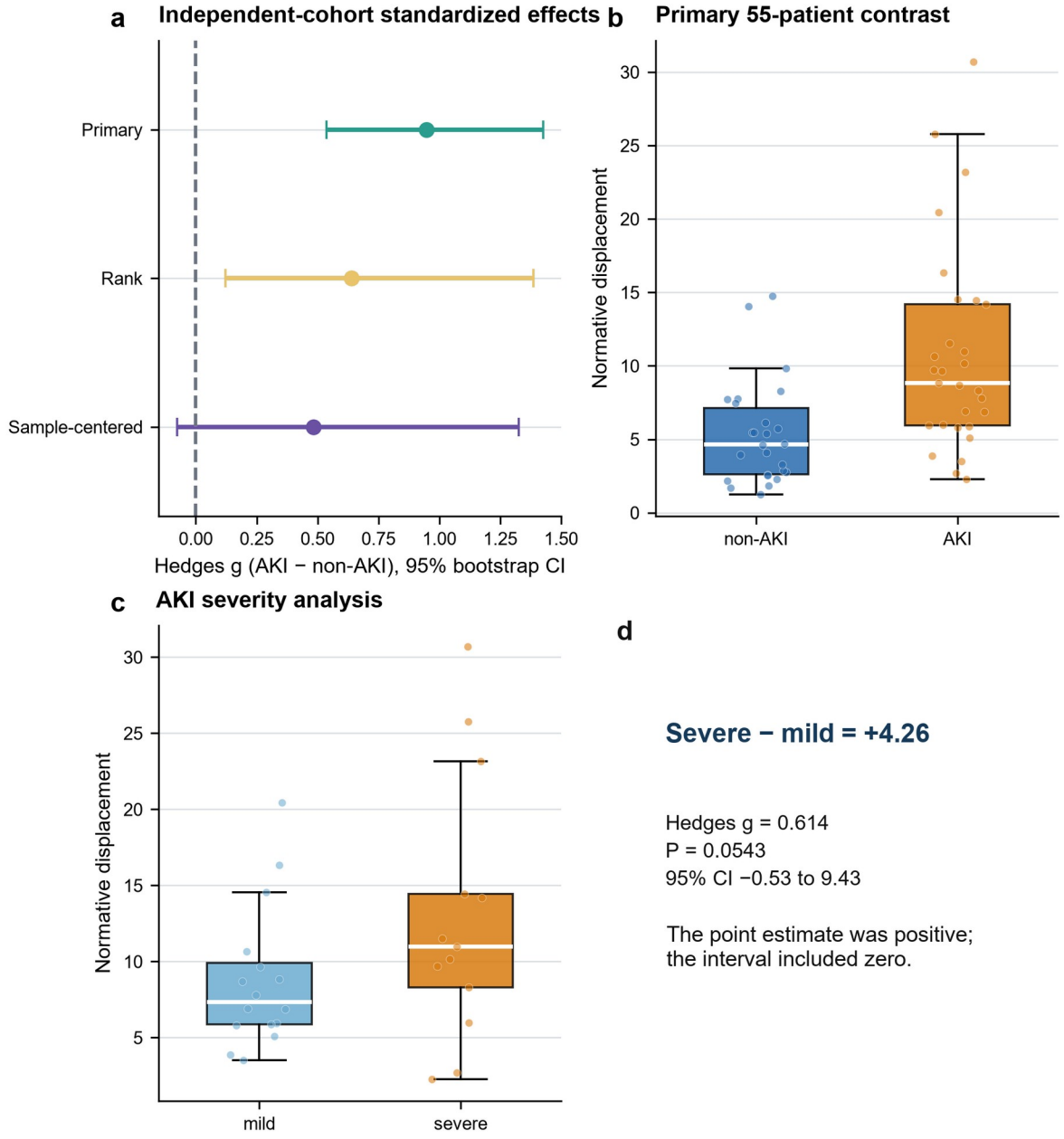


Fig. S4 Independent urine-cohort comparisons and source-defined AKI severity. a, Standardized AKI-minus-non-AKI effects under the three preprocessing approaches. Point estimates were positive, but uncertainty differed across approaches. b, Individual primary-analysis distances from the non-AKI reference in 55 patients. c,d, Secondary comparison of 13 source-labeled severe and 16 mild AKI patients. These labels correspond to stages 2–3 and stage 1 as reported by the source, not independently reassigned stages. The severe-minus-mild mean difference was 4.265 (one-sided p = 0.0543; fixed-score bootstrap 95% CI −0.533 to 9.431). The cohort had one urine sample per patient and cannot test return toward baseline

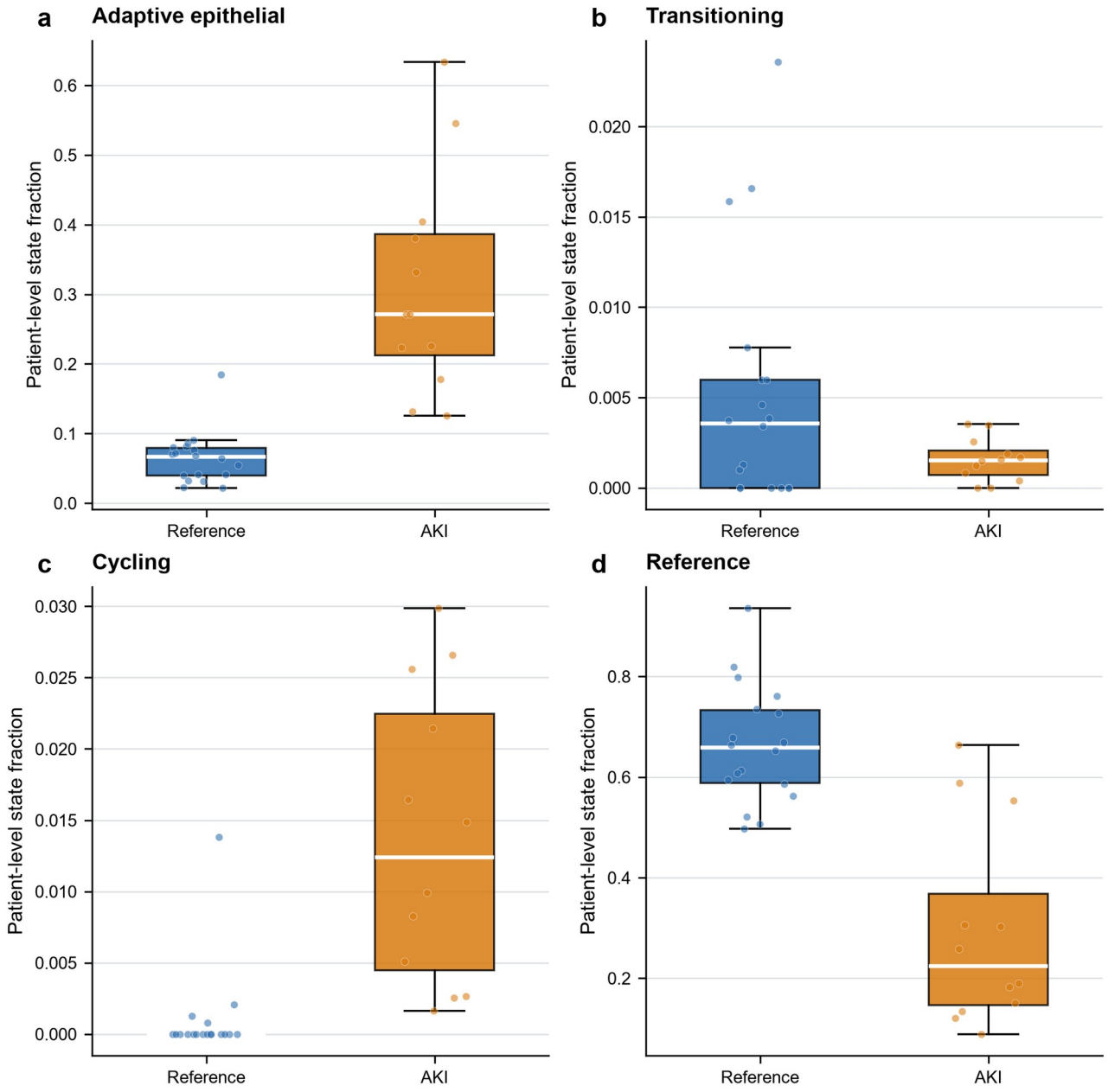


Fig. S5 Other source-defined tubular-cell states in the kidney atlas. Patient-level fractions of cells labeled adaptive epithelial, transitioning, cycling and reference in 12 AKI and 18 reference kidneys. Fractions were calculated after aggregating eligible cells within each patient. These four secondary comparisons were tested two-sided and corrected together by Benjamini–Hochberg; estimates are in Table S4. They describe sampled cell-state composition rather than mechanisms explaining the urinary measurements

Supplementary tables

Table S1. Within-patient comparisons with the preoperative profile.

Metric	AKI median	non-AKI median	Median difference	Cliff delta	Permutation P	BH Q	Bootstrap 95% CI	Delete-one direction
Residual ratio	1.0349	0.9802	+0.05475	0.4236	1.00×10^{-5}	1.00×10^{-5}	0.03339 to 0.07903	145/145
Reversal fidelity	0.3206	0.4134	−0.09274	−0.5217	1.00×10^{-5}	1.00×10^{-5}	−0.12056 to −0.05345	145/145

Note. Residual ratio and reversal fidelity use T1, T3 and T4 (preoperative, approximately 20 h and approximately 44 h). Effects are AKI-minus-non-AKI median differences; intervals use 20,000 stratified patient bootstraps. p values are one-sided rank-sum permutations. BH Q covers these two tests only. The delete-one column gives the number of analyses retaining the reported direction out of 145.

Table S2. The 44-h distance comparison under different preprocessing approaches.

Representation	AKI mean	non-AKI mean	Raw difference	Hedges g	Permutation P	Raw bootstrap 95% CI	Hedges g bootstrap 95% CI
Primary	98.524	73.544	24.980	1.268	1.00×10^{-5}	17.869 to 34.032	0.999 to 2.039
Within-sample rank	0.8469	0.7225	0.1244	1.375	1.00×10^{-5}	0.0942 to 0.1585	0.974 to 1.998
Sample-median-centered	95.845	73.560	22.285	1.548	1.00×10^{-5}	17.351 to 28.040	1.134 to 2.307

Note. Raw differences are AKI-minus-non-AKI mean differences at T4; the Hedges g intervals are standardized-effect intervals. p values are one-sided. These are alternative representations of the same patients, not separate cohorts. The magnitude-adjusted primary sensitivity was inconclusive.

Table S3. Independent early postoperative urine-cohort estimates.

Contrast	n	Effect	Hedges g	P	Bootstrap 95% CI
AKI vs non-AKI displacement	29 vs 26	+5.384	0.946	0.0168	2.619 to 8.218
Severe vs mild AKI	13 vs 16	+4.265	0.614	0.0543	−0.533 to 9.431

Note. The urine cohort contained one postoperative sample per patient. The AKI/non-AKI permutation test rebuilt the reference and scores at each reassignment. Bootstrap intervals used fixed scores. The severe-versus-mild comparison used source-defined labels and was secondary; both p values are one-sided.

Table S4. Patient-level tubular-cell state comparisons in the kidney atlas.

State	AKI mean	Reference mean	Difference	Hedges g	Permutation P	Bootstrap 95% CI	BH-adjusted FDR/Q
degenerative	0.3795	0.2669	+0.1126	0.874	0.01162	0.0248 to 0.2024	—
adaptive epithelial	0.3103	0.0644	+0.2459	2.315	1.00×10^{-5}	0.1628 to 0.3371	2.00×10^{-5}
transitioning	0.0016	0.0052	−0.0036	−0.661	0.07291	−0.0069 to −0.0008	0.07291
cycling	0.0137	0.0010	+0.0127	1.800	4.00×10^{-5}	0.0071 to 0.0185	5.33×10^{-5}
reference	0.2950	0.6625	−0.3676	−2.315	1.00×10^{-5}	−0.4815 to −0.2435	2.00×10^{-5}

Note. Fractions were calculated within patients before comparison. The degenerative-state comparison was one-sided and was not part of the four-test secondary false-discovery-rate family. The other four tests were two-sided. Bootstrap intervals describe patient-resampling uncertainty and are not inversions of the permutation tests.

Table S5. Public data inputs and archive identifiers.

Data record	File or identifier	Official record	Role in this analysis
Longitudinal processed urine	Table S1: pr5c00113_si_002.xlsx (AKI); Table S2: pr5c00113_si_003.xlsx (non-AKI)	ACS Supporting Information	Processed matrices analyzed
Raw mass-spectrometry archive	PXD063754; iProX IPX0009727000	ProteomeCentral PXD063754	Archive provenance only; raw files not analyzed
Independent processed urine	Supplementary Material 2: 11306_2026_2423_MOESM2_ESM.xlsx	Springer article and SI	Processed matrix analyzed
Kidney-tissue metadata	GSE183276; patient-level metadata file used	NCBI GEO GSE183276	Patient-level state.l2 metadata analyzed

Records were accessed on 7 October 2026. Processed Supporting Information files and tissue metadata were the analysis inputs. The PXD/iProX record identifies the raw archive only; raw mass-spectrometry files were not analyzed.