Online Supplementary File S1

Preoperative Patient-Reported Global Physical Health and Postoperative Acute Kidney Injury in the All of Us Research Program: A Cohort Study

Kutay Çelik, MD

Supplementary methods, Tables S1-S4, and Figure S1

Supplementary methods

Data source and operative cohort

The study used the national United States All of Us Research Program Controlled Tier v9 curated data repository (C2025Q4R6) in the Researcher Workbench. The exact calendar span of index operations and data collection was unavailable; the exposure and outcome windows are specified below. An episode had to occur on or after both official All of Us consent dates. The unit of analysis was one operation per adult participant.

Candidate operations were identified from the published fixed 565-code Operative Stress Score (OSS) codelist. CPT codes 33025, 35820, and 35840 were removed, leaving a 562-code core. No additional obstetric or transplant category filter was applied. A candidate episode was a participant-day containing at least one qualifying code, assigned the maximum same-day OSS. The first eligible episode was selected before survey availability was considered. The primary path used standardized Observational Medical Outcomes Partnership concepts; a supporting analysis used source-system procedure concepts. Same-day anesthesia billing was not required. The adapted episode algorithm was checked across the two coding paths but was not validated against manual chart review.

Patient-reported exposure

The exposure was the official four-item PROMIS Global Physical Health (GPH) T-score derived from general physical health, everyday physical activity, average pain over seven days, and fatigue. A score required exactly one valid response to each component on the same date. Incomplete, contradictory, unresolved nonnumeric, and coded-skip responses were not scored. The most recent complete score 7 to 90 days before the operation was selected; same-day and post-index surveys were excluded. Higher scores indicate better physical health. GPH was modeled linearly per 10-point increase.

The question concept identifiers were 1585723 (physical health), 1585741 (everyday activities), 1585747 (pain), and 1585748 (fatigue). Physical health responses from excellent to poor, activity responses from completely to not at all, and fatigue responses from none to very severe were scored 5 to 1 in order. Integer pain responses of 0, 1-3, 4-6, 7-9, and 10 were scored 5, 4, 3, 2, and 1,

respectively. The four scores were summed. Raw totals 4 through 20 mapped, in order, to T-scores 16.2, 19.9, 23.5, 26.7, 29.6, 32.4, 34.9, 37.4, 39.8, 42.3, 44.9, 47.7, 50.8, 54.1, 57.7, 61.9, and 67.7.

Creatinine outcome and eligibility

Serum creatinine measurements used concept identifiers 3016723 and 3020564. Accepted source-unit forms were mg/dL, 258797006, and umol/L; values reported in micromoles per liter were divided by 88.4 to obtain mg/dL. Valid values were 0.2 or greater and less than 25 mg/dL, equivalent to 17.68 or greater and less than 2,210 micromoles per liter. A non-null, non-midnight datetime was treated as precise.

The reference was the most recent valid serum creatinine in the 60-day preoperative window. Precisely timed reference and postoperative values had to occur strictly before and after the operation. When precise operation or laboratory time was unavailable, the operation date was excluded from both windows; the reference window was days -60 through -1 and the postoperative window was days +1 through +7. Postoperative measurements were censored at a subsequent qualifying operation.

The primary endpoint used an operational serum-creatinine definition based on Kidney Disease: Improving Global Outcomes thresholds. AKI was present when postoperative creatinine was at least 0.3 mg/dL above the preoperative reference during the first 48 postoperative hours or at least 1.5 times the reference within seven days. The absolute-rise window was anchored to surgery and did not require the reference and postoperative measurements to be within 48 hours of one another. An absolute creatinine of at least 4.0 mg/dL was classified as stage 3 only after an acute-rise criterion was met. Urine output and renal replacement therapy were unavailable and not imputed.

Strict renal eligibility required a computable 2021 Chronic Kidney Disease Epidemiology Collaboration creatinine eGFR of at least 15 mL/min/1.73 m^2 after rounding to the nearest whole number. Sex concepts unsupported by the equation were excluded rather than assigned a binary coefficient. Records missing a primary model field were excluded. The analyzed cohort was complete for required model fields; variable-specific missing-data counts before exclusion were unavailable.

Primary model and absolute-scale summaries

All eligible participants in the data release were included; no formal sample-size calculation was performed. The primary logistic model contained seven slope degrees of freedom: GPH per 10 points, age per 10 years, the equation-supported sex concept, baseline eGFR per 10 mL/min/1.73 m², OSS as a linear ordinal term, inpatient versus other setting, and centered calendar year. Data-contributing organization fields were used as provenance groups rather than hospital identities. Episodes linked to more than one organization were assigned to one multisource category. The covariance plan used finite-sample-corrected clustering when at least 30 provenance groups and 20 event-bearing groups were available; otherwise, HC3 covariance was intended. The primary model met the clustering rule.

The primary estimand was the adjusted conditional odds ratio (OR) for AKI per 10-point higher GPH. A crude association estimate with precision was unavailable. Model-standardized risks were calculated by setting GPH to the observed first and third quartiles and averaging fitted risks over the cohort's observed covariate distribution. Confidence intervals for the two risks and their difference used 4,000 draws from a multivariate normal distribution defined by the fitted coefficients and covariance; the 2.5th and 97.5th percentiles formed the interval limits.

Supporting analyses

Initial supporting analyses examined source-system procedure codes; highest and lowest eligible baseline creatinine; a three-knot restricted cubic spline for GPH with knots at the 10th, 50th, and 90th analytic percentiles; precise-time and date-only operation subsets; three survey-lag windows; exclusion of multisource episodes; and anonymized leave-one-provenance-source-out refits. Adjusted lag-window models required at least 100 AKI events. A permissive endpoint construction intended to retain eGFR-unavailable records produced no interpretable estimate because its design matrix was rank deficient.

A later exploratory renal-proxy model included GPH, age, explicit male, female, and unsupported sex-concept terms, centered log baseline creatinine, OSS, inpatient setting, and centered calendar year. The formal operation-time interaction added precision status and a GPH-by-precision term to the strict primary model; timing-specific ORs were derived from that clustered model.

The paired date-window comparison reran precise-time records with the date-only creatinine windows and refit the primary model among participants eligible under both rules. The separate precise-time and date-only fits and both paired date-window fits used uncorrected heteroskedasticity-robust covariance (HC0). HC3 covariance had been intended when cluster support was insufficient, but statsmodels 0.14.6 implemented HC0 for these historical fits. They were not retrospectively recomputed with HC3.

The observed-48-hour analysis retained the original complete-case cohort, covariates, provenance-clustered covariance, and 1.5-fold reference criterion. The absolute branch instead required an observed rise of at least 0.3 mg/dL between measurements no more than 48 hours apart, ending within the original postoperative window. When both times were precise and non-midnight, the interval had to be strictly positive and no more than 48 elapsed hours. If either time was unavailable, dates had to be exactly one day apart; untimed same-day and two-day pairs did not establish a confirmed interval. The earlier value had to fall within the original admissible baseline or uncensored postoperative window, and the later value within the uncensored postoperative window. Participants without a qualifying pair remained in the cohort and could still qualify through the relative-rise branch. No observed pair therefore meant no observed absolute-rise evidence under this rule, not absence of clinical AKI.

Analysis chronology and observation weighting

The primary model and initial supporting analyses were defined before the primary estimate. The renal-proxy model, formal operation-time interaction, paired date-window comparison, and observed-48-hour analysis were developed later and are exploratory. None replaced the primary model or converted the study into independent confirmation.

Observation weighting was explored but is not used as corroborating evidence. The initial event effective sample size, a measure of information retained under unequal weights, was 77.54 against a planned minimum of 80. A later model changed the observation predictors and eligible population and removed that minimum, with event effective sample size 76.04. The changed reporting rule did not provide an independent precision basis for relying on the weighted estimate. The primary analysis did not use these weights.

Interpretation for future evaluation

Across the primary and supporting analyses, recent patient-reported physical health was associated with postoperative AKI alongside measured clinical factors. Because GPH records the patient's perspective, the finding supports evaluating whether a brief preoperative assessment could help clinicians recognize vulnerability and inform proportionate postoperative vigilance as one input to judgment. This clinical-use hypothesis requires prospective testing of incremental prediction and patient benefit.

Disclosure control

Counts from 1 through 20 were neither reported nor made inferable. Small cohort totals are shown as 20-wide intervals derived from retained lower-bound buckets. No participant row, identifier, date, timestamp, source label, residual, or fitted value left the Researcher Workbench.

GPH indicates PROMIS Global Physical Health; OSS, Operative Stress Score; eGFR, estimated glomerular filtration rate; AKI, acute kidney injury; OR, odds ratio; HC0, uncorrected heteroskedasticity-robust covariance estimator; and HC3, leverage-adjusted heteroskedasticity-robust covariance estimator.

Supplementary results

Supplementary Table S1. Primary absolute-scale estimates and supporting association analyses

Scientific question	Analysis	Estimate or retained result	Interpretation
Principal association	Primary model	OR 0.714 (95% CI, 0.579 to 0.882)	Primary adjusted estimate
Absolute scale	AKI risk at GPH 34.9	10.2% (95% CI, 7.86% to 13.5%)	Model-standardized cohort average
Absolute scale	AKI risk at GPH 47.7	6.95% (95% CI, 5.55% to 9.06%)	Model-standardized cohort average
Absolute scale	GPH 47.7 versus 34.9 risk difference	-3.28 percentage points (95% CI, -5.78 to -1.19)	Associational contrast, not treatment effect
Procedure coding	Source-system procedure codes	OR 0.751 (95% CI, 0.620 to 0.910)	Direction retained
Baseline reference	Highest eligible baseline creatinine	OR 0.743 (95% CI, 0.555 to 0.995)	Construction changed magnitude
Baseline reference	Lowest eligible baseline creatinine	OR 0.650 (95% CI, 0.565 to 0.748)	Construction changed magnitude
Renal adjustment	Paired-creatinine renal-proxy model	OR 0.733 (95% CI, 0.599 to 0.897)	Later exploratory sensitivity
Exposure form	Three-knot GPH spline	Nonlinear coefficient -0.027 (95% CI, -1.493 to 1.440)	No supported departure from linearity
Source influence	Excluding multisource episodes	OR 0.704 (95% CI, 0.565 to 0.877)	Direction retained
Source influence	Leave-one-source-out refits	34 of 34 valid; GPH coefficient range -0.3765 to -0.2738; largest upper CI limit -0.0543	No single anonymized source accounted for the direction
Operation-time subset	Precise-time records, separate fit	HC0 OR 0.570 (95% CI, 0.416 to 0.781)	Initial supporting fit; not the formal interaction estimate
Operation-time subset	Date-only records, separate fit	HC0 OR 0.947 (95% CI, 0.705 to 1.273)	Initial supporting fit; not the formal interaction estimate

Values are adjusted conditional ORs per 10-point higher GPH unless identified as standardized risks, a risk difference, or another diagnostic. The permissive endpoint construction yielded no interpretable estimate because its design matrix was rank deficient. Adjusted survey-lag models were not reported because each contained fewer than the prespecified 100 AKI events. The operation-time interaction and observed-48-hour analysis appear in Figure 1 and the timing tables rather than being

duplicated here. No observation-weighted estimate is used as supporting evidence.

Supplementary Table S2. Creatinine timing and recorded measurement availability

Recorded measurement	Primary cohort	GPH at or below 42.3	GPH above 42.3
Baseline-reference lag, median [Q1, Q3], days	8 [1, 20]	8 [1, 20]	8 [2, 20]
Distinct postoperative measurement days, median [Q1, Q3]	Not retained	2 [1, 4]	1 [1, 3]
First postoperative measurement day, median [Q1, Q3]	1 [1, 1]	1 [1, 1]	1 [1, 1]
Last postoperative measurement day, median [Q1, Q3]	2 [1, 5]	3 [1, 6]	2 [1, 5]

The split at GPH 42.3 was the retained cohort median, not a clinical threshold. No subgroup model or inferential comparison was performed. These summaries describe recorded availability rather than verified test-ordering frequency or complete surveillance. Suppressed event-overlap and precise-sample support cells were not reconstructed or reported.

AKI indicates acute kidney injury; CI, confidence interval; GPH, PROMIS Global Physical Health; HC0, uncorrected heteroskedasticity-robust covariance estimator; OR, odds ratio; and Q1 and Q3, first and third quartiles.

Additional operation-time evidence

Supplementary Table S3. Paired date-window comparison

Precise-time records were reanalyzed using the date-only rule, which excludes the operation day from both the reference and postoperative windows. Among participants eligible under both rules, the cohort was held constant and only the measurement rule differed.

Analysis	N range	Event range	Adjusted OR per 10-point higher GPH	95% CI
Primary model, whole cohort, native rule	1,400-1,419	120-139	0.714	0.579 to 0.882
Precise time, native rule	840-859	60-79	0.570	0.416 to 0.781
Date only, native rule	560-579	40-59	0.947	0.705 to 1.273
Precise time eligible under both rules, native rule	740-759	60-79	0.544	0.393 to 0.752
Precise time eligible under both rules, date-only rule	740-759	40-59	0.555	0.402 to 0.766

Measurement change	Value
Precise-time records losing endpoint eligibility under the date-only rule	12.2%
Discordant AKI classifications divided by the native event count	33.3%
Median baseline-reference lag under the native rule	8 days

Measurement change	Value
Median baseline-reference lag under the date-only rule	11 days
Median change in baseline creatinine	0.000 mg/dL

The discordance percentage uses the native event count as its denominator and includes both native events lost and new events gained; it is not the percentage of all paired records whose classification changed. The whole-cohort primary model used provenance-clustered covariance. The separate precision-subset fits and both paired fits used HC0 despite the planned HC3 fallback and were not retrospectively recomputed. Imposing date-only windows did not reproduce the attenuation observed between the original timing groups and did not identify its mechanism.

Supplementary Table S4. Characteristics by operation-time precision

Characteristic	Precise time	Date only
Age, years	57.10 (14.78)	58.06 (15.17)
PROMIS Global Physical Health T-score	41.45 (9.00)	43.00 (9.70)
eGFR, mL/min/1.73 m²	84.96 (25.55)	84.20 (26.16)
Baseline creatinine, mg/dL	0.97 (0.46)	0.96 (0.43)
Operative Stress Score	2.23 (0.85)	2.36 (0.86)
Survey-to-operation lag, days	43.63 (24.19)	42.67 (24.65)
Baseline-reference lag, days	12.80 (14.78)	13.76 (13.50)
First postoperative measurement day	1.15 (1.45)	1.63 (1.48)
Last postoperative measurement day	3.03 (2.28)	3.22 (2.23)
Seven-day AKI	8.3%	8.8%
Inpatient operation	43.2%	60.1%

Continuous values are mean (SD). Counts underlying percentages are disclosure-controlled 20-wide intervals, and percentages were withheld when a contributing cell was 20 or smaller. No inferential comparison was performed. The groups describe record precision and are not clinical subgroups.

Supplementary Figure S1. Cohort construction

Flow from the first eligible noncardiac operation identified with standardized procedure concepts to a complete official GPH score 7 to 90 days before the operation, sufficient creatinine data for the primary endpoint, and seven-day creatinine-defined AKI. Counts are disclosure-controlled 20-wide intervals.

GPH indicates PROMIS Global Physical Health; eGFR, estimated glomerular filtration rate; AKI, acute kidney injury; OR, odds ratio; CI, confidence interval; HC0, uncorrected heteroskedasticity-robust covariance estimator; and HC3, leverage-adjusted heteroskedasticity-robust covariance estimator.