Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Perfusion indices cannot be reproduced from publicly available perioperative waveform data: a reproducibility and usability assessment of VitalDB

Wenyue Xiang
Affiliated Qingyuan Hospital, Guangzhou Medical University, Qingyuan People's Hospital
Zhiming Huang
Affiliated Qingyuan Hospital, Guangzhou Medical University, Qingyuan People's Hospital
Weixian Xie
2026691021@gzhmu.edu.cn

Affiliated Qingyuan Hospital, Guangzhou Medical University, Qingyuan People's Hospital

Research Article

Keywords: perfusion index, photoplethysmography, acute kidney injury, intraoperative monitoring, reproducibility, VitalDB, pulse pressure variation

Posted Date: November 9th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11404280/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Perfusion indices cannot be reproduced from publicly available perioperative waveform data: a reproducibility and usability assessment of VitalDB

Wenyue Xiang[1], Zhiming Huang[1], Weixian Xie[1]*

[1] *Department of Anesthesiology, Affiliated Qingyuan Hospital, Guangzhou Medical University, Qingyuan People's Hospital, Qingyuan 511518, Guangdong, China*

**Corresponding author: Weixian Xie, No. 35 Yinquan North Road, Qingcheng District, Qingyuan 511518, Guangdong, China. E-mail: 2026691021@gzhmu.edu.cn. Telephone: +86-18826610702. ORCID: 0009-0005-5285-8716.*

Author e-mail addresses: Wenyue Xiang, xwy8759@163.com (ORCID 0009-0004-6423-4815); Zhiming Huang, hzm7501@126.com (ORCID 0009-0002-6496-4450); Weixian Xie, 2026691021@gzhmu.edu.cn (ORCID 0009-0005-5285-8716).

Word counts: Abstract 346; main text (Background–Conclusions) 4610; whole manuscript 5897.

Article type: Research article.

Number of figures/tables: 4 figures, 4 tables, 1 supplementary file.

Abstract

Background. The peripheral perfusion index (PI) is a monitor-derived plethysmographic measure reported to predict postoperative acute kidney injury (AKI). VitalDB, the largest public perioperative waveform database, originates from the same institution as the original PI–AKI study; we asked whether it is reproducible from the public release.

Methods. Using the public VitalDB release (6,388 cases), we attempted four sequential reproduction steps: identification of a monitor-derived PI/PVI numeric channel; derivation of PI from the raw plethysmographic waveform with assessment of cross-case scale comparability; construction of a within-case relative PI (rel-PI); and computation of pulse pressure variation (PPV). The outcome was KDIGO AKI. Associations were evaluated in the waveform core set (n = 3,518) with threshold and multivariable analyses; explicit hypotension (mean arterial pressure < 65 mmHg) served as a positive control.

Results. No monitor-derived PI/PVI numeric channel exists in the public release. Waveform-derived PI ranged 6.4–115.9% across cases (86% of 3,507 cases above the physiological upper bound ≈25%): the display scale is auto-ranged and not comparable across cases. The relative index was measurable and reproducible, and was not attributable to noise or automatic gain artefact (split-half reliability r = 0.967), but was unrelated to AKI. The primary exposure (a persistent ≥ 2-min episode with a ≥ 50% fall in the relative perfusion signal at normal arterial pressure) showed no association (crude OR 0.94, 95% CI 0.76–1.17; adjusted OR 0.98, p = 0.86); a nominally significant crude association at ≥ 30 min (OR 1.40, p = 0.026) was attenuated after adjustment (OR 1.31, p = 0.148); no association was seen without explicit hypotension (OR 0.84, 95% CI 0.62–1.14). In an underpowered pilot (241 cases), PPV showed no association (OR 1.17, 95% CI 0.52–2.63). The positive control remained robust (adjusted OR 1.62, 95% CI 1.21–2.16, p = 0.0011).

Conclusions. A monitor-derived perfusion index cannot be reproduced from public perioperative waveform data: the monitor channel is absent and the released waveform scale is not comparable across cases. The within-case surrogate is measurable and reproducible but a different construct and does not reproduce the reported association. These constraints reflect open waveform archive design, not VitalDB alone.

Keywords. perfusion index; photoplethysmography; acute kidney injury; intraoperative monitoring; reproducibility; VitalDB; pulse pressure variation

Background

Postoperative acute kidney injury (AKI) complicates a substantial proportion of non-cardiac surgeries and is independently associated with short- and long-term morbidity and mortality, even when the injury is mild and creatinine later recovers [1]. Because AKI is frequently silent and is driven by potentially modifiable intraoperative physiology, considerable effort has been directed at identifying intraoperative haemodynamic signals that precede organ injury and could inform management.

The most extensively studied such signal is arterial blood pressure. Observational cohorts have consistently reported graded associations between intraoperative hypotension and postoperative AKI, whether hypotension is defined by absolute thresholds, most commonly a mean arterial pressure (MAP) below 65 mmHg, or by reduction from a preoperative baseline [2,3]. Absolute and relative definitions perform comparably, and risk rises with the duration of exposure [3]. Systemic pressure, however, is a global and comparatively blunt descriptor of organ perfusion: at any given MAP, regional and microcirculatory perfusion vary with vascular tone, volume status, and cardiac output. Pressure and flow can therefore be dissociated, which motivates interest in measures that reflect perfusion more directly at the tissue level.

The peripheral perfusion index (PI) is one such measure. Derived from the pulse oximetry photoplethysmographic signal, PI is the ratio of the pulsatile (AC) to the non-pulsatile (DC) component of light absorption and quantifies the strength of the peripheral pulse relative to static tissue [4,5]. It is non-invasive, continuous, and available on most modern patient monitors. Because it reflects the interaction between central haemodynamics (stroke volume) and peripheral vascular tone, a low PI has been proposed as an early marker of haemodynamic compromise [4]. Intraoperative PI has been associated with severe postoperative complications and mortality in high-risk surgical patients [6]. Most relevant here, a single-centre study based on the Seoul National University Hospital (SNUH) database reported that intraoperative PI was associated with postoperative AKI, defining the exposure as monitor-recorded PI below 0.5 for more than 10% of surgery time and proposing a Youden-optimal absolute threshold of 3.13 [7]. This exposure, a monitor-derived PI level, is attractive because it is available on the monitor rather than reconstructed retrospectively by the researcher. It should be distinguished from the outset from a PI that a researcher re-derives offline from a stored plethysmographic waveform (a waveform-derived PI, wd-PI): the two are computed from related but not identical quantities, and need not be numerically equivalent.

Whether such an exposure can be reproduced from openly released data is increasingly important. Perioperative waveform datasets have proliferated, and VitalDB, released by the same institution as the PI–AKI study and, with more than 6,000 cases and 500-Hz waveforms, the most widely used open perioperative biosignal resource, provides an appropriate setting in which to test reproducibility [8,10]. Open datasets exist so that exposures, models, and associations can be independently re-examined [10]. Reproducing a monitor-computed exposure, however, requires more than access to the raw signals: the release must contain the variable the original study actually analysed, not merely a subset of the underlying waveforms. Concretely, it presupposes two conditions: that the corresponding numeric channel is present in the public release, and that any waveform from which the exposure is re-derived is on a comparable and interpretable scale.

We therefore conducted a pre-specified sequence of reproduction attempts on the public VitalDB release, asking whether the published monitoring-based perfusion exposure, and its association with AKI, can be reproduced from publicly available data. We tested (1) whether a monitor-derived PI/PVI numeric channel exists in the release; (2) whether PI can be derived from the raw plethysmographic waveform and

whether its scale is comparable across cases; (3) whether a within-case relative surrogate is measurable and whether it recovers the reported association; and (4) whether a comparator dynamic index (pulse pressure variation) behaves similarly. Throughout, we used explicit hypotension (MAP < 65 mmHg) as a positive control to confirm that the analytic pipeline could detect a signal known to be present.

Methods

Data sources

VitalDB (public release). 6,388 surgical cases from Seoul National University Hospital, with intraoperative waveforms at 500 Hz (including plethysmography and arterial pressure), 1-Hz numeric trends, and laboratory and clinical files [8] (Table 1). The waveform core set was defined as cases with both an invasive arterial waveform and a plethysmographic waveform (n = 3,518).

MOVER. Used only for structural comparison of waveform and outcome availability in an external, independent health system [11]. No association analysis was performed in MOVER; it was used solely to compare what an internal, device-integrated repository provides relative to a public waveform release.

Step 1: availability of a monitor-derived PI

All unique track names in the VitalDB track index were enumerated and searched for any monitor-computed perfusion-related numeric channel (PI, PI%, PVI, or similar). The original study recorded PI from a Philips IntelliVue MX800 monitor via the hospital's internal clinical data warehouse [7].

Step 2: waveform-derived PI and scale comparability

PI was derived from the VitalDB plethysmographic waveform track (SNU Analog-to-Digital Converter, `SNUADC/PLETH`) as the ratio of pulsatile to non-pulsatile amplitude (wd-PI): in 10-s windows, the signal was band-pass filtered (0.7–6 Hz), individual pulse amplitudes were measured (peak-to-trough), and wd-PI = 100 × median(amplitude)/mean DC. The across-case distribution of wd-PI was compared with the physiological range (approximately 0.02–25%) [5].

Step 3: within-case relative PI and covert hypoperfusion

Because absolute wd-PI was not comparable across cases, a case-internal relative index was constructed: rel-PI = (window wd-PI)/(case baseline wd-PI), where the baseline was the median wd-PI over all valid windows of that case (minimum 36 valid windows, i.e. 6 min). We used the overall-case median rather than a pre-induction baseline because pre-induction plethysmographic segments in the public release are frequently short, noisy, or absent, so a robust whole-case median was the most uniformly available reference; the sensitivity of the relative index to the baseline definition is addressed in Limitations. Each 10-s window was then classified into one of four mutually exclusive states:

- **covert hypoperfusion (CHP)**: MAP ≥ 65 mmHg with rel-PI < 0.5, i.e. the pulsatile perfusion signal had fallen by at least 50% from the patient's own baseline while systemic pressure remained normal;

- **overt hypotension (OHP)**: MAP < 65 mmHg with rel-PI < 0.5;
- **decoupled hypotension (DH)**: MAP < 65 mmHg with rel-PI ≥ 0.5, i.e. pressure was low without a proportional fall in the perfusion signal;
- **reference**: MAP ≥ 65 mmHg with rel-PI ≥ 0.5.

Windows with an invalid waveform or missing MAP were classified as invalid and interrupted state continuity; invalid windows were excluded from case-level summaries rather than imputed. States were aggregated into episodes requiring persistence for ≥ 2 min (12 consecutive 10-s windows). The primary exposure was the occurrence of at least one persistent CHP episode; cumulative minutes in each state were retained for threshold and dose–response analyses.

Signal-character audit

To distinguish a genuine relative-perfusion signal from a noise or auto-gain artefact, a signal-character audit was performed in a random sample of 300 cases drawn with a fixed, arbitrary seed (the seed value is recorded in Supplementary Methods S7) from the waveform core set (n = 3,518), of which 299 had valid window series (385,983 windows). The sampling frame was the full waveform core set rather than the outcome-analysable subset, because the audit concerns the properties of the derived signal itself and is independent of outcome availability. Three checks were used: split-half reliability of the windowed estimate (odd- versus even-indexed pulses within each window), within-case DC–amplitude coupling (the signature of automatic gain ranging), and the lag-1 autocorrelation of the rel-PI series (temporal smoothness). Full methods are given in Supplementary Methods S7.

Step 4: pulse pressure variation (PPV)

PPV was computed from `SNUADC/ART` (100 Hz, 30-s windows) using a robust estimator (median filtering, systolic-peak segmentation, rhythm-regularity QC to exclude arrhythmia, and PPV = 100 × (PP95 − PP5)/PP50) with physiological pulse-pressure QC. Thresholds and validity conditions were informed by perioperative haemodynamic consensus [9]. Because the aim of this step was to test whether a dynamic comparator could be computed reliably from the public release, rather than to test PPV itself, it was conducted as a pilot: 300 cases were drawn at random with a fixed, arbitrary seed (the seed value is recorded in Supplementary Methods S4) from the 3,459 core-set cases that also had an analyzable arterial waveform. This sampling pool differs from the signal-character audit sampling frame (n = 3,518) because PPV requires an analyzable arterial-pressure waveform in addition to plethysmography, whereas the signal-character audit was built on the plethysmographic waveform alone. The pilot was not sized by a formal power calculation, is underpowered for the AKI outcome, and is reported for direction only.

Outcome: KDIGO AKI

AKI was defined by KDIGO serum creatinine criteria from the VitalDB laboratory file [12]. Baseline creatinine was taken from the preoperative value or the earliest in-hospital value, with a validity filter

(baseline ≥ 0.4 mg/dL); the threshold prevents a small absolute creatinine rise from being amplified into a spurious large relative rise when the baseline is implausibly low.

Statistical analysis

Associations between exposures and AKI were assessed with threshold sensitivity analyses, continuous (Spearman) tests, and multivariable logistic regression. The main adjusted model included age, sex, BMI, ASA physical status ≥ 3, preoperative hypertension, preoperative diabetes, baseline creatinine, surgery duration, major resection, and emergency status; the PPV pilot model included age, sex, BMI, ASA physical status ≥ 3, and median MAP. Preoperative hypertension, preoperative diabetes, and major resection were included in addition to the core haemodynamic covariates because each is an established determinant of postoperative AKI and a potential confounder of the perfusion–AKI relationship, so that the adjusted estimate would not be confounded by comorbidity burden or surgical extent. Crude 2 × 2 comparisons are reported with Fisher's exact p values, and adjusted estimates with Wald p values. MAP < 65 mmHg (explicit hypotension) served as a positive control. Analyses were performed in the waveform core set and in the subgroup without explicit hypotension.

Missing data were handled as follows. The outcome required at least one postoperative serum creatinine within 7 days and a valid baseline creatinine; cases without either were excluded from the association analysis (n = 2,910 of 3,518). Cases with fewer than 36 valid waveform windows were not eligible for exposure classification, and invalid windows were excluded from case-level exposure summaries rather than imputed. Covariates entered the models as observed; complete-case analysis excluded no additional cases in the primary models (n = 2,910). No imputation was performed. All analyses used Python (numpy, scipy, statsmodels).

The use of AI-assisted tools in preparing this manuscript is disclosed in the Declarations section.

Results

The four reproduction attempts and the positive control are summarized in Table 2; analysis-set sizes are shown in Fig. 1.

No monitor-derived PI channel in the public release

Enumeration of all unique VitalDB channel names (196 unique track names across 6,388 cases) identified no monitor-computed PI, PVI, or perfusion numeric channel (Supplementary Table S1). The exposure used in the original study (a monitor-derived PI) therefore cannot be obtained from the public release and requires the institution's internal complete database. This alone precludes direct reproduction.

Waveform-derived PI is not comparable across cases

Across the full waveform core set with a valid plethysmographic median (n = 3,507 of 3,518; Fig. 2 and Table 2), the case-level median wd-PI spanned 6.4–115.9% (median 71.4%, IQR 37.5–87.1%), far beyond the physiological range (≈0.02–25%) [5], and 86% of cases lay above the physiological upper

bound. The waveform is an auto-ranged display signal, so each case has a fixed but arbitrary vertical scale. Absolute wd-PI is therefore not comparable across cases, and the published absolute thresholds (0.5 and 3.13), which fall entirely below the observed range, cannot be applied.

The within-case relative surrogate is reproducible but shows no association

Signal character (n = 299 of 300 randomly sampled cases; 385,983 windows). The audit sample was drawn at random with a fixed, arbitrary seed from the full waveform core set (n = 3,518), not from the AKI-analysable subset, because the audit addresses the properties of the derived signal and does not require outcome availability (Methods). We explicitly tested whether rel-PI was noise or an auto-gain artefact and rejected both:

- **Split-half reliability** of the windowed estimate was 0.967 (385,983 windows). Within each 10-s window, the estimate was recomputed separately from the odd-indexed and the even-indexed pulse (interleaved beats), and the two half-estimates were correlated across all windows; a random signal would give a correlation near 0 (Table 4).
- **Within-case DC–amplitude coupling** (auto-gain check) had a median case-level Spearman of 0.044 (P90 0.171); no case exceeded 0.5 and none exceeded 0.8, and the within-case DC coefficient of variation had a median of 3.3%. The gain is fixed within a case, so rel-PI variation is not an auto-range artefact.
- **Lag-1 temporal autocorrelation** of rel-PI had a median of 0.814 (P10 0.491, P90 0.959); 96.7% of cases were above 0.3 and none below 0. The signal evolves smoothly, behaving like a slow physiological process rather than white noise.

These three checks together indicate that rel-PI is measurable, reproducible, and temporally coherent, rather than noise. Its failure to reproduce the published association is therefore attributable to a construct difference (a within-case relative change versus an absolute monitor-derived PI level) and to cross-case scale incomparability, not to estimation error.

Association with AKI (n = 2,910 analysable; AKI 374, 12.85%; Table 3). The primary exposure, at least one persistent CHP episode (≥ 2 min), occurred in 1,433 cases (49.2%) and was not associated with AKI (crude OR 0.94, 95% CI 0.76–1.17, p = 0.58; adjusted OR 0.98, 95% CI 0.74–1.29, p = 0.86; Table 3 and Fig. 3). Alternative definitions of the same construct, examined as threshold sensitivity analyses, were likewise null except for the duration-based thresholds: a cumulative CHP exposure of 30 min or more showed a nominally significant crude association (OR 1.40, 95% CI 1.05–1.86, p = 0.026) that was attenuated and no longer significant after multivariable adjustment (adjusted OR 1.31, 95% CI 0.91–1.89, p = 0.148), and the same pattern appeared at 60 min (crude OR 1.78, 95% CI 1.11–2.85, p = 0.021; adjusted OR 1.49, 95% CI 0.83–2.69, p = 0.18; Table 3). The crude duration-based associations did not survive adjustment for covariates that were themselves strongly associated with AKI in these models (surgery duration, baseline creatinine, ASA physical status ≥ 3, male sex, and emergency status; Supplementary Table S3), a pattern consistent with confounding rather than an independent exposure effect. These duration thresholds define small groups: the ≥ 30-min exposure applied to only 14.3% of cases (416 of 2,910) and the ≥ 60-min exposure to 3.9% (113 of 2,910), so the nominally significant crude

estimates are themselves unstable and further argue against a real duration-dependent effect. No definition showed a consistent monotonic dose–response across thresholds (Supplementary Table S3).

These estimates, together with the threshold and dose–response definitions, are tabulated in full in Table 3. Crude p values are from Fisher's exact test; adjusted p values are Wald tests. The full model output, including all covariate coefficients and the dose–response bins, is provided in Supplementary Table S3.

In the subgroup of patients who never developed MAP < 65 mmHg (n = 1,894; AKI 197, 10.40%), the primary exposure was not associated with AKI (crude OR 0.84, 95% CI 0.62–1.14, p = 0.28; adjusted OR 0.85, 95% CI 0.57–1.27, p = 0.43). This subgroup isolates the covert construct by removing any contribution of explicit hypotension: even among patients whose mean arterial pressure never fell below 65 mmHg, the relative perfusion signal carried no detectable association with postoperative AKI.

PPV (dynamic preload index) also shows no association

In the 300-case pilot (241 analysable; AKI 30, 12.45%), PPV-based exposure was not associated with AKI (PPV > 13% at MAP ≥ 65 for ≥ 30 min: crude OR 1.17, 95% CI 0.52–2.63, Fisher p = 0.84; adjusted OR 1.36, 95% CI 0.51–3.61, p = 0.54; continuous $|\rho| \leq 0.09$; Fig. 4 and Table 3). The duration-based thresholds showed no consistent direction (Fig. 4). We emphasize that this pilot was a feasibility analysis of whether a dynamic comparator could be computed from the public release, not a test of PPV itself: with 30 AKI events it is underpowered and can exclude only very large effects, and its result is reported for direction only. No inference about the clinical value of PPV is drawn from it.

Positive control: explicit hypotension remains robust

MAP < 65 mmHg was robustly associated with AKI (crude OR 1.82, 95% CI 1.46–2.26, p < 0.001; adjusted OR 1.62, 95% CI 1.21–2.16, p = 0.0011; Table 3), and median MAP < 70 mmHg showed an OR of 3.09 (p < 0.001). All multivariable coefficients for conventional AKI risk factors were in the expected direction and magnitude, confirming that the analysis pipeline was functioning.

Discussion

In this reproducibility assessment, a monitor-derived perfusion-index exposure could not be reconstructed from publicly available perioperative waveform data. The monitor-computed PI numeric channel used in the original study was absent from the public release; PI re-derived from the raw plethysmographic waveform was on a non-comparable, auto-ranged scale across cases; and a within-case relative surrogate, although measurable and demonstrably reproducible, did not reproduce the reported association with AKI. The MAP-based positive control remained robust, confirming that the exposure–outcome pipeline could detect a genuine signal. The failure of reproduction is therefore a property of the exposure rather than of the analysis.

The most informative finding concerns why the surrogate fails, because it distinguishes this study from a simple negative result. A natural first interpretation would be that the waveform-derived relative index is too noisy to be useful. We tested this directly and rejected it: across 385,983 windows in 299 cases, the

windowed estimate had a split-half reliability of 0.967, showed essentially no coupling between within-case DC level and pulse amplitude (median Spearman 0.044; no case above 0.5), the signature expected if the signal were being auto-ranged, and was strongly autocorrelated in time (median lag-1 autocorrelation 0.814). rel-PI thus behaves like a smooth, repeatable physiological signal. The absence of an association with AKI therefore cannot be dismissed as measurement error; it reflects a construct difference, in that rel-PI describes a change relative to a patient's own baseline, whereas the published exposure is an absolute PI level. These are not interchangeable, and a relative surrogate constructed to circumvent the scale problem is, by construction, not the exposure that was originally reported.

The reproducibility failure can be described as a three-layer problem. At the **channel** layer, the exposure is unobtainable: the public release does not include the monitor-computed PI/PVI numeric channel, so the exact variable studied cannot be recovered from open data at all. At the **scale** layer, the raw plethysmographic waveform that is released is not on an absolutely interpretable scale: across-case values spanned 6.4–115.9% (median 71.4%; 86% of cases above the physiological upper bound of ≈25%), far outside the physiological range, so published absolute thresholds, whether a Youden-optimal value such as 3.13 or a clinical threshold such as 0.5, cannot be transported to waveform-derived estimates [5,7]. At the **construct** layer, even a carefully built relative index, which is internally valid and reproducible, measures a different quantity than the absolute exposure and need not carry the same association. The three layers are cumulative: solving the scale problem by relative normalization does not and cannot recover the original construct. The first two constraints are moreover independent: even if a future release included the monitor-derived PI numeric channel, an auto-ranged waveform would still not substitute for the monitor-native value in secondary analysis, and conversely a comparable waveform scale would not create the missing channel. Both conditions must hold before a published monitor-derived exposure can be reproduced at all.

Our results should not be read as evidence against the clinical value of PI, nor against the original study. Low intraoperative PI has been associated with adverse outcomes in independent cohorts [6], and the biological rationale for PI as a perfusion marker is well established [4]. Rather, our findings clarify the conditions under which a monitor-based exposure can be reproduced from open data. The practical implication is that waveform-derived haemodynamic indices, including PI, the pleth variability index, PPV, and similar measures, cannot be used as drop-in substitutes for monitor-computed values without first verifying (a) that the corresponding numeric channel is available and (b) that any waveform from which the index is derived is on a comparable, interpretable scale. When neither holds, effect estimates are not comparable across studies, however carefully the surrogate is computed.

The pleth variability index (PVI) deserves a specific note. It is computed by the monitor from the same plethysmographic waveform as PI and therefore inherits the same two constraints: its numeric channel is absent from the public release, and the released waveform is auto-ranged. The same applies to other monitor-computed indices that would have to be re-derived from released waveforms.

Several recommendations follow. Data contributors should, where permissible, release monitor-computed numeric channels alongside raw waveforms, since these are the variables that clinical studies actually

analyse. Where investigators can access only released waveforms, manuscripts should explicitly state that waveform-derived indices are not equivalent to monitor-native variables; authors should report how scale and gain were handled, validate the surrogate against a reference measurement where feasible, and clearly describe the measured construct together with the resulting limits on cross-study comparison. Journal editors and reviewers assessing reproducibility claims should ask which specific channel the exposure came from and whether that channel is present in the dataset claimed to reproduce it. Negative and feasibility results of this kind should be reported, both to reduce publication bias and to prevent repeated attempts to reconstruct exposures that the data cannot support [10].

Our comparison with an independent dataset is instructive. MOVER, an external and structurally different perioperative repository, provides dense intraoperative numeric data and structured postoperative outcomes and can therefore support low-blood-pressure phenotypes and competing-risk analyses [11]. The same fundamental constraint, however, applies to its waveform-signal indices: cross-database comparability of plethysmographic-derived measures depends on device- and scale-specific handling. The reproducibility challenge we describe is thus not unique to VitalDB but is a general feature of reproducing monitor-computed exposures from open waveform archives.

This study has limitations. First, it is confined to a single public database, and the specific threshold values and channel inventories are properties of that release; we also cannot verify whether the non-public internal version of the plethysmographic waveform is scaled identically to the released version, although the monitor-computed PI channel is absent from the public release in either case. Second, the within-case relative surrogate depends on how the case baseline is defined. We used the median wd-PI over all valid windows of the case; a different definition, for example a stable pre-induction period, would produce a different relative index, although the construct of a within-case relative change would remain distinct from an absolute monitor level. Third, AKI was defined by the serum-creatinine criterion of KDIGO only; the urine-output criterion could not be applied with the available data, so some cases of AKI, particularly those presenting primarily with oliguria, may have been missed. Such misclassification would be expected to bias association estimates toward the null, which strengthens rather than weakens the conclusion that the surrogate does not reproduce the reported association. Fourth, the PPV pilot was deliberately small and underpowered and is reported for direction only; it does not constitute a definitive test of PPV, and its null result should not be generalized to PPV itself. Finally, the within-case relative surrogate is not the gold-standard monitor PI and, as the audit shows, measures a different construct; the audit supports its measurability and reproducibility but not its equivalence to the published absolute exposure. We could not apply the original study's absolute threshold directly, because the required channel and scale are unavailable; this is the substantive finding rather than an analytic shortcoming.

Conclusions

A monitor-derived perfusion-index exposure cannot be reproduced from publicly available perioperative waveform data: the monitor channel is absent and the released waveform scale is not comparable across cases. A within-case relative surrogate is measurable and reproducible, but it is a distinct construct and

does not recover the reported association. These constraints arise from how open waveform archives are constructed rather than from VitalDB alone, and they extend to other monitor-computed plethysmographic indices such as the pleth variability index. Investigators seeking to reproduce monitor-computed haemodynamic exposures from open perioperative data should verify channel availability and scale comparability before drawing conclusions from effect estimates.

List of abbreviations

AC alternating current; **AKI** acute kidney injury; **ART** invasive arterial pressure; **ASA** American Society of Anesthesiologists; **AUROC** area under the receiver operating characteristic curve; **BMI** body mass index; **CHP** covert hypoperfusion; **CI** confidence interval; **DC** direct current; **DH** decoupled hypotension; **IQR** interquartile range; **KDIGO** Kidney Disease: Improving Global Outcomes; **MAP** mean arterial pressure; **MOVER** Medical Informatics Operating Room Vitals and Events Repository; **OHP** overt hypotension; **OR** odds ratio; **PI** perfusion index; **PLETH** plethysmographic waveform; **PPV** pulse pressure variation; **PVI** pleth variability index; **QC** quality control; **rel-PI** within-case relative perfusion index; **SNUH** Seoul National University Hospital; **SNUADC** SNU Analog-to-Digital Converter (VitalDB waveform track); **wd-PI** waveform-derived perfusion index.

Declarations

Ethics approval and consent to participate

This study analysed only publicly available, fully de-identified data from the VitalDB database. The acquisition and open release of the VitalDB data were approved by the Institutional Review Board of Seoul National University Hospital (IRB No. H-1408-101-605), and written informed consent was waived because the released data are anonymised [8]; the original data collection was also registered at ClinicalTrials.gov (NCT02914444). Data collection was performed in accordance with the relevant guidelines and regulations of that institutional ethics committee. No identifiable human data were accessed by the present authors. Because this was a secondary analysis of a publicly available, fully de-identified dataset, no additional ethics approval or informed consent was required.

Consent for publication

Not applicable. This manuscript does not contain data from any individual person.

Availability of data and materials

The datasets analysed in this study are publicly available: VitalDB (<https://vitaldb.net/dataset>) and MOVER (<https://doi.org/10.1093/jamiaopen/ooad084>). VitalDB data were used under the data-use agreement of the data contributors and are not redistributed by the present authors. All analysis code and the derived result files that reproduce every reported number are archived at Zenodo

(<https://doi.org/10.5281/zenodo.23397596>; version 1.0.0) [13], and the analysis scripts are also provided as Supplementary Material.

Competing interests

The authors declare that they have no competing interests.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Authors' contributions

Wenyue Xiang contributed to the study design, acquired and verified the source datasets, contributed to clinical interpretation of the haemodynamic variables, and critically revised the manuscript. Zhiming Huang contributed to data curation and quality control, independently verified the reported numerical results against the analysis output files, checked the tables and figures for consistency with those outputs, and critically revised the manuscript. Weixian Xie conceived and designed the study, built the analysis pipeline, performed the statistical analyses, drafted the manuscript, and supervised the work. All authors read and approved the final manuscript and agree to be accountable for all aspects of the work.

Acknowledgements

The authors thank the VitalDB team at Seoul National University Hospital for releasing the VitalDB dataset, and the investigators of the MOVER repository for making their dataset publicly available.

Declaration of AI-assisted tools in the writing process

During the preparation of this work the authors used a generative AI assistant to support drafting and language editing of the manuscript and to support the development of the analysis and plotting code. After using this tool, the authors reviewed and edited all content, and verified every reported number against the source output files. The authors take full responsibility for the content of this publication.

Authors' information (optional)

Wenyue Xiang: Affiliated Qingyuan Hospital, Guangzhou Medical University, Qingyuan People's Hospital, Qingyuan 511518, Guangdong, China. E-mail: <xwy8759@163.com>. ORCID: 0009-0004-6423-4815.

Zhiming Huang: Affiliated Qingyuan Hospital, Guangzhou Medical University, Qingyuan People's Hospital, Qingyuan 511518, Guangdong, China. E-mail: <hzm7501@126.com>. ORCID: 0009-0002-6496-4450.

Weixian Xie (corresponding author): Affiliated Qingyuan Hospital, Guangzhou Medical University, Qingyuan People's Hospital, No. 35 Yinquan North Road, Qingcheng District, Qingyuan 511518,

Guangdong, China. E-mail: <2026691021@gzhmu.edu.cn>. Telephone: +86-18826610702. ORCID: 0009-0005-5285-8716.

References

1. O'Connor ME, Hewson RW, Kirwan CJ, Ackland GL, Pearse RM, Prowle JR. Acute kidney injury and mortality 1 year after major non-cardiac surgery. Br J Surg. 2017;104(7):868–876. doi:10.1002/bjs.10498.

2. Sun LY, Wijeysundera DN, Tait GA, Beattie WS. Association of intraoperative hypotension with acute kidney injury after elective noncardiac surgery. Anesthesiology. 2015;123(3):515–523. doi:10.1097/ALN.0000000000000765.

3. Salmasi V, Maheshwari K, Yang D, Mascha EJ, Singh A, Sessler DI, Kurz A. Relationship between intraoperative hypotension, defined by either reduction from baseline or absolute thresholds, and acute kidney and myocardial injury after noncardiac surgery: a retrospective cohort analysis. Anesthesiology. 2017;126(1):47–65. doi:10.1097/ALN.0000000000001432.

4. Coutrot M, Dudoignon E, Joachim J, Gayat E, Vallée F, Dépret F. Perfusion index: physical principles, physiological meanings and clinical implications in anaesthesia and critical care. Anaesth Crit Care Pain Med. 2021;40(6):100964. doi:10.1016/j.accpm.2021.100964.

5. Lima AP, Beelen P, Bakker J. Use of a peripheral perfusion index derived from the pulse oximetry signal as a noninvasive indicator of perfusion. Crit Care Med. 2002;30(6):1210–1213. doi:10.1097/00003246-200206000-00006.

6. Agerskov M, Thusholdt ANW, Holm-Sørensen H, Wiberg S, Meyhoff CS, Højlund J, Secher NH, Foss NB. Association of the intraoperative peripheral perfusion index with postoperative morbidity and mortality in acute surgical patients: a retrospective observational multicentre cohort study. Br J Anaesth. 2021;127(3):396–404. doi:10.1016/j.bja.2021.06.004.

7. Kang P, Park JB, Yoon HK, Ji SH, Jang YE, Kim EH, Lee JH, Lee HC, Kim JT, Kim HS. Association of the perfusion index with postoperative acute kidney injury: a retrospective study. Korean J Anesthesiol. 2023;76(4):348–356. doi:10.4097/kja.22620.

8. Lee HC, Park Y, Yoon SB, Yang SM, Park D, Jung CW. VitalDB, a high-fidelity multi-parameter vital signs database in surgical patients. Sci Data. 2022;9(1):279. doi:10.1038/s41597-022-01411-5.

9. Fellahi JL, Futier E, Vaisse C, Collange O, Huet O, Loriau J, Gayat E, Tavernier B, Biais M, Asehnoune K, Cholley B, Longrois D. Perioperative hemodynamic optimization: from guidelines to implementation—an experts' opinion paper. Ann Intensive Care. 2021;11(1):58. doi:10.1186/s13613-021-00845-1.

10. Lim L, Lee HC. Open datasets in perioperative medicine: a narrative review. Anesth Pain Med (Seoul). 2023;18(3):213–219. doi:10.17085/apm.23076.

11. Samad M, Angel M, Rinehart J, Kanomata Y, Baldi P, Cannesson M. Medical Informatics Operating Room Vitals and Events Repository (MOVER): a public-access operating room database. JAMIA Open. 2023;6(4):ooad084. doi:10.1093/jamiaopen/ooad084.

12. Kidney Disease: Improving Global Outcomes (KDIGO) Acute Kidney Injury Work Group. KDIGO clinical practice guideline for acute kidney injury. Kidney Int Suppl. 2012;2(1):1–138. doi:10.1038/kisup.2012.1.

13. Xiang W, Huang Z, Xie W. Reproduction code and derived results: perfusion indices cannot be reproduced from publicly available perioperative waveform data (version 1.0.0). Zenodo. 2026. doi:10.5281/zenodo.23397596.

Figure legends

Fig. 1. Study flow. The monitor-derived perfusion-index exposure reported by Kang et al. [7] was subjected to four sequential reproduction attempts on the public VitalDB release: (1) channel availability; (2) waveform-derived PI and scale comparability; (3) a within-case relative surrogate (rel-PI) tested against KDIGO AKI; and (4) a dynamic comparator (PPV). Explicit hypotension (MAP < 65 mmHg) served as a positive control. The analysis sets derived from the waveform core set (n = 3,518) are annotated in the flow: 3,507 cases for the wd-PI scale comparison, 2,910 cases (AKI 374) for the association analysis, 299 of 300 randomly sampled cases (385,983 windows) for the signal-character audit, and 241 analysable of 300 randomly drawn cases for the PPV pilot. The signal-character audit sample was drawn at random with a fixed, arbitrary seed from the full core set (n = 3,518), independently of outcome availability.

Fig. 2. Waveform-derived perfusion index is not comparable across cases. Distribution of the case-level median wd-PI across the waveform core set (n = 3,507; log scale). The shaded band marks the physiological PI range (≈0.02–25%) [5]; dotted lines mark the published absolute thresholds (0.5 and 3.13) [7], which lie entirely below the observed range. 86% of cases exceed the physiological upper bound.

Fig. 3. Within-case relative perfusion index and postoperative AKI (n = 2,910; AKI 374). Left: odds ratios for covert-hypoperfusion (CHP) exposures, with the primary exposure (a persistent CHP episode of at least 2 min, i.e. MAP ≥ 65 mmHg with rel-PI < 0.5 for ≥ 2 min) shown crude and adjusted, together with threshold and no-hypotension-subgroup variants, and for the positive control (explicit hypotension). Right, signal-character audit of the within-case relative index (299 of 300 randomly sampled cases; 385,983 windows), which addresses three alternative explanations for the null association: (a) estimate reproducibility: split-half reliability of the windowed estimate, computed from interleaved (odd- versus even-indexed) pulses within each window and plotted across rel-PI deciles (overall r = 0.967), ruling out a random estimate; (b) auto-gain check: within-case coupling between the DC level and the pulse amplitude (median Spearman 0.044; no case > 0.5), ruling out an auto-ranging artefact; (c) temporal smoothness: lag-1 autocorrelation of rel-PI (median 0.814; 96.7% of cases > 0.3), ruling out white noise.

The audit shows that rel-PI is measurable and reproducible, rather than noise or an artefact, so its null association reflects a construct difference rather than estimation error.

Fig. 4. Pulse pressure variation (PPV) and AKI, 300-case pilot (241 analysable; AKI 30). Odds ratios for PPV-based exposures; continuous PPV showed $|\rho| \leq 0.09$. The pilot was a feasibility analysis and was not powered to test PPV; results are reported for direction only and should not be read as evidence that PPV is unrelated to AKI.

Tables

All four tables are cited in the text (Table 1 in Data sources; Tables 2 and 3 in Results; Table 4 in the signal-character audit) and are supplied as separate files at submission.

- **Table 1.** Data-source inventory.
- **Table 2.** Reproduction-attempt matrix.
- **Table 3.** Associations between covert hypoperfusion (CHP) exposures and postoperative KDIGO AKI. The primary exposure is a persistent CHP episode of at least 2 min; the remaining blocks are

threshold and subgroup sensitivity analyses, the PPV pilot, and an explicit-hypotension positive control.

- **Table 4.** Signal-character audit of the within-case relative index.

Table 1. Data-source inventory

Data source	Institution / origin	Cases	Waveforms	Numeric resolution	Key outcomes available	Role in this study
VitalDB (public)	Seoul National University Hospital, Korea	6,388	500 Hz (incl. PLETH, ART)	1–7 s	Creatinine (lab), clinical	Primary dataset — 4 reproduction steps
— waveform core set	subset with invasive ART ∩ PLETH	**3,518**	500 Hz both channels	—	—	Main exposure/outcome analysis
MOVER (external)	UC Irvine Medical Center, USA	65,729 (EPIC) + 23,928 (SIS)	high-fidelity (incl. PLETH)	1 min (EPIC)	Creatinine, structured post-op complications	Structural comparison only (external system)

Reference ranges: VitalDB [8]; MOVER [11].

Table 2. Reproduction-attempt matrix

Step		Target (from source study [7])	Method applied	Channel available in public release?	Result	Verdict
	1	Monitor-derived PI numeric channel	Enumerate all 196 unique track names; search PI / PVI / perfusion	**No**	No monitor-computed perfusion channel exists	❌ **Exposure unobtainable**
	2	Absolute PI level (thresholds 3.13 / 0.5)	Derive wd-PI from `SNUADC/PLETH` (0.7–6 Hz, per-beat amplitude/DC)	Waveform present, scale auto-ranged	Across-case wd-PI (n = 3,507): median **71.4%**, IQR 37.5–87.1%, range **6.4–115.9%**; 86% of cases > 25% (physiological ≈0.02–25% [5])	❌ **Scale not comparable**
	3	PI–AKI association	Within-case relative surrogate (rel-PI, baseline-normalized) + KDIGO AKI	Surrogate measurable; reproducible (split-half r = 0.967)	No association (persistent CHP ≥ 2 min: crude OR 0.94, p = 0.58; adjusted OR 0.98, p = 0.86; no-hypotension subgroup OR 0.84)	❌ **Not reproduced (construct differs)**
	4	Dynamic preload comparator	PPV from `SNUADC/ART` (300 randomly drawn cases)	Waveform present	No association (crude OR 1.17, p = 0.84; underpowered pilot, direction only)	❌ Not reproduced
Control		Explicit hypotension (known)	MAP < 65 mmHg	Numeric MAP available	**Robust** association (adjusted OR 1.62, p = 0.0011)	✅ **Pipeline validated**

Table 3. Associations between covert hypoperfusion (CHP) exposures and postoperative KDIGO AKI. Primary exposure = persistent CHP episode ≥ 2 min (first row); the remaining rows are threshold and subgroup sensitivity analyses, the PPV pilot, and the explicit-hypotension positive control

Exposure	Exposed (n)	AKI %	Crude OR (95% CI)	p	Adjusted OR (95% CI)	p
rel-PI / covert hypoperfusion (CHP)						
Persistent CHP episode (≥ 2 min)	1,433	12.5	0.94 (0.76–1.17)	0.58	0.98 (0.74–1.29)	0.86
Any CHP window (> 0 min)	2,522 (86.7%)	13.3	1.37 (0.97–1.95)	0.087	1.47 (0.91–2.39)	0.118
CHP time ≥ 10% of valid monitoring	776	11.9	0.88 (0.69–1.14)	0.35	1.20 (0.88–1.63)	0.26
CHP ≥ 30 min	416	16.4	1.40 (1.05–1.86)	0.026	1.31 (0.91–1.89)	0.148
CHP ≥ 60 min	113	20.4	1.78 (1.11–2.85)	0.021	1.49 (0.83–2.69)	0.18
Subgroup with no explicit hypotension (n = 1,894): persistent CHP episode (≥ 2 min)	773	9.4	0.84 (0.62–1.14)	0.28	0.85 (0.57–1.27)	0.43
PPV (300-case pilot; 241 analysable; AKI 30, 12.45%)						
PPV > 13% at MAP ≥ 65 for ≥ 30 min	153	13.07	1.17 (0.52–2.63)	0.840 (Fisher)	1.36 (0.51–3.61)	0.54
P2CHP time ≥ 20%	162	9.26	0.44 (0.20–0.94)	0.038 (protective direction)	—	—
Continuous PPV (Spearman ρ)	—	—	abs(ρ) ≤ 0.09	≥ 0.17	—	—
Positive control						
MAP < 65 mmHg (explicit hypotension)	—	—	1.82 (1.46–2.26)	< 0.001	**1.62 (1.21–2.16)**	**0.0011**
Median MAP < 70 mmHg	—	—	3.09	< 0.001	—	—

Analysis set: waveform core set with valid baseline creatinine and ≥ 1 postoperative measurement (n = 2,910; AKI 374, 12.85%). Outcome: KDIGO AKI [12]. Crude p values are from Fisher's exact test; adjusted p values are Wald tests. Adjusted models: age, sex, BMI, ASA physical status ≥ 3, preoperative hypertension, preoperative diabetes, baseline creatinine, surgery duration, major resection, emergency status. PPV adjusted model (pilot): age, sex, BMI, ASA ≥ 3, median MAP. CHP = covert hypoperfusion: MAP ≥ 65 mmHg with rel-PI < 0.5 (a ≥ 50% fall from the case baseline); a persistent CHP episode required ≥ 2 min of continuous CHP; "Any CHP window" counts cases with at least one CHP window without the persistence requirement (sensitivity definition). "—" = not applicable / not estimated.

Table 4. Signal-character audit of the within-case relative index (n = 299 of 300 randomly sampled cases; 385,983 windows)

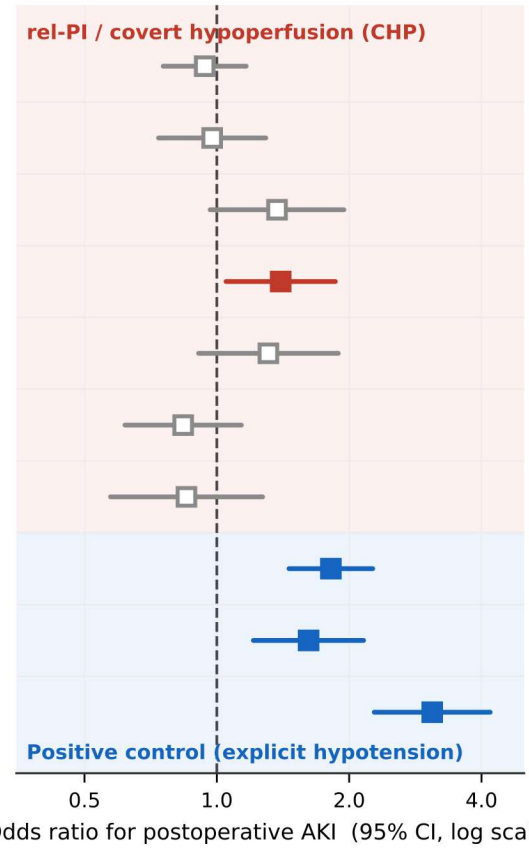
Test	Quantity	Result	Interpretation
Split-half reliability	r	**0.967**	Reproducible, not random
DC–amplitude coupling (auto-gain check)	median case-level Spearman	**0.044** (0% > 0.5)	Gain fixed within case → not auto-range artefact
Within-case DC variability	coefficient of variation (median)	**3.3%**	Stable baseline
lag-1 temporal autocorrelation	median	**0.814** (96.7% > 0.3)	Smooth, physiological-like

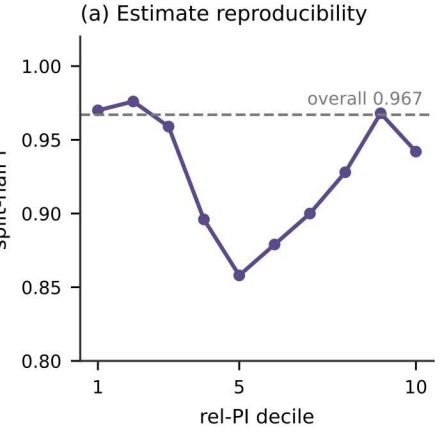
The audit sample was drawn at random with a fixed, arbitrary seed (seed = 2026; Supplementary Methods S7) from the full waveform core set (n = 3,518), not from the AKI-analysable subset, and is therefore independent of outcome availability. Corresponds to Fig. 3 (right panel) and Supplementary Table S4.

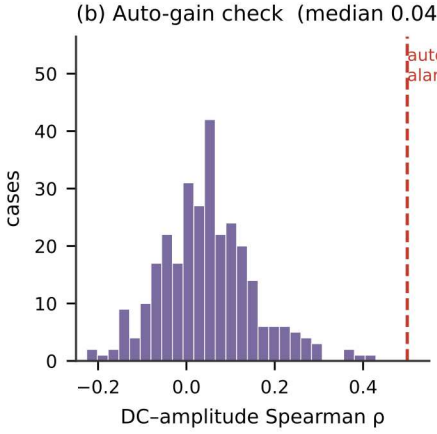
Fig 3. Within-case relative perfusion index and postoperative AKI

Signal character of the within-case relative index (n = 299 cases; 385,983 windows): rel-PI is measurable and reproducible, and is not a noise

Persistent CHP episode ≥ 2 min (crude)	0.94 (0.76–1.17)
Persistent CHP episode ≥ 2 min (adjusted)	0.98 (0.74–1.29)
Any CHP window > 0 min (crude)	1.37 (0.97–1.95)
CHP ≥ 30 min (crude)	**1.40 (1.05–1.86)**
CHP ≥ 30 min (adjusted)	1.31 (0.91–1.89)
No-hypotension subgroup (crude)	0.84 (0.62–1.14)
No-hypotension subgroup (adjusted)	0.85 (0.57–1.27)
MAP < 65 mmHg (crude)	**1.82 (1.46–2.26)**
MAP < 65 mmHg (adjusted)	**1.62 (1.21–2.16)**
Median MAP < 70 mmHg (crude)	**3.09 (2.28–4.19)**







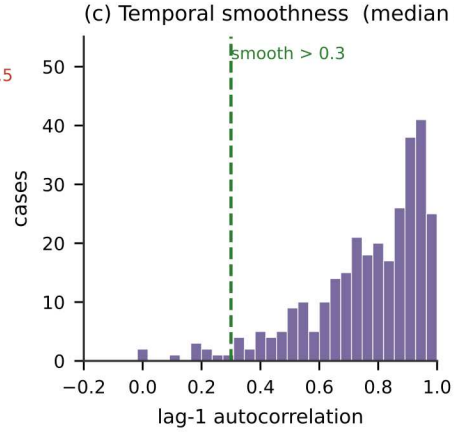


Fig 2. Waveform-derived perfusion index is not comparable across cases

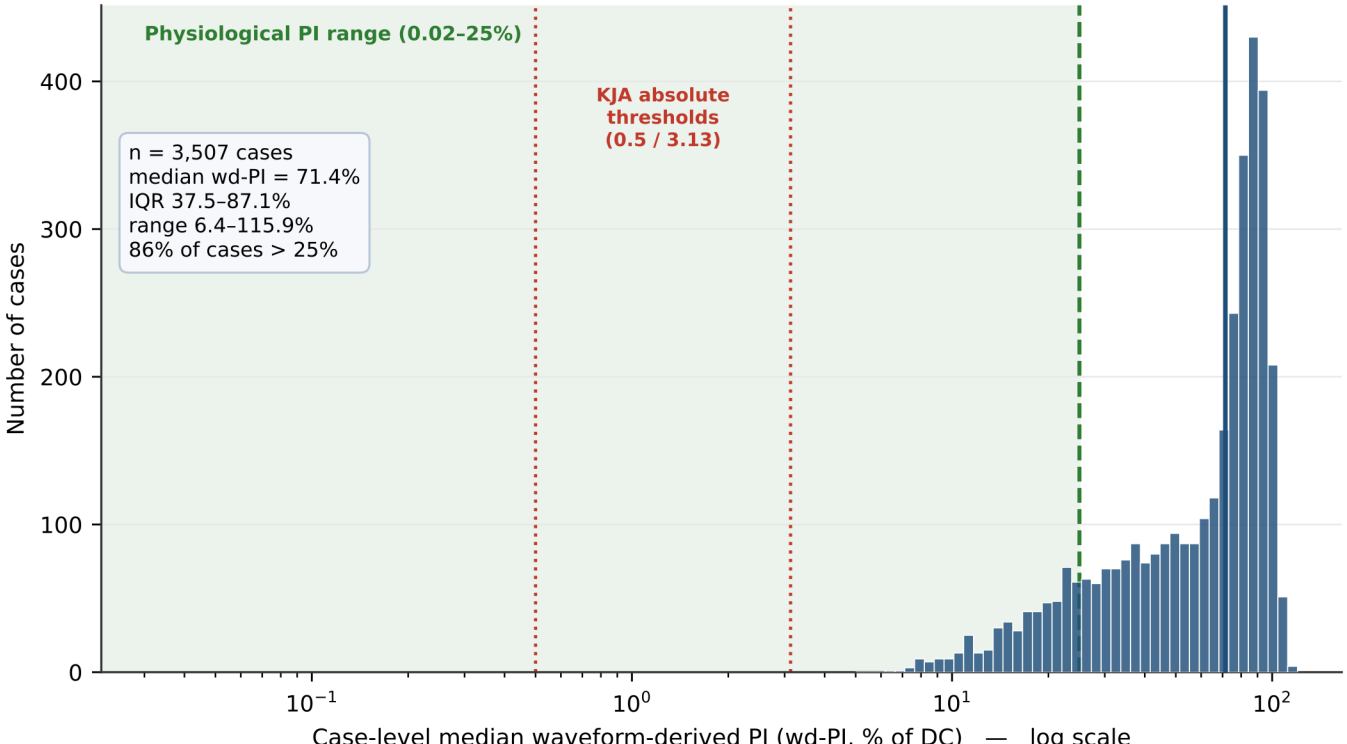
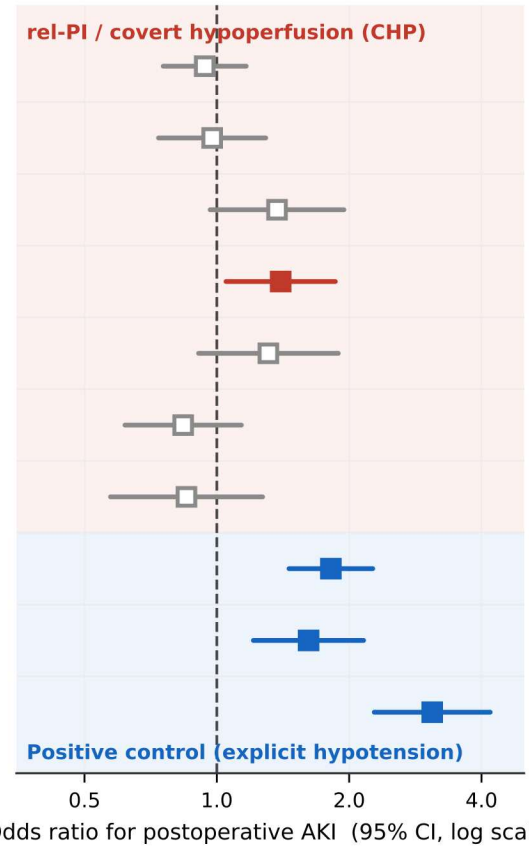
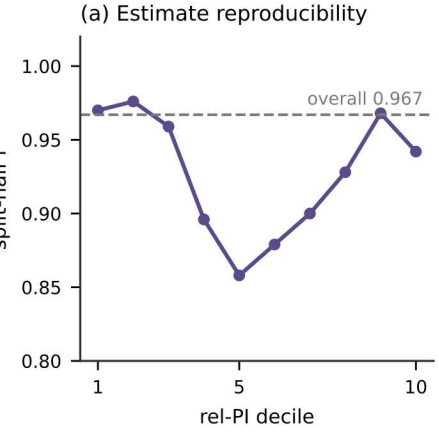


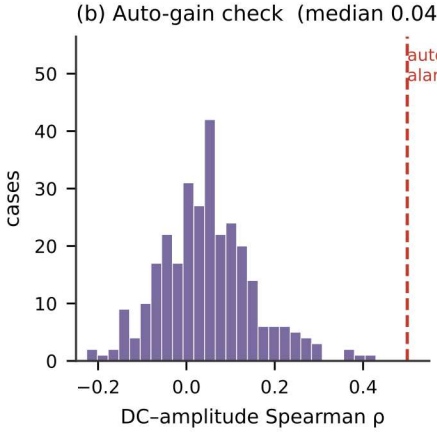
Fig 3. Within-case relative perfusion index and postoperative AKI

Signal character of the within-case relative index (n = 299 cases; 385,983 windows): rel-PI is measurable and reproducible, and is not a noise

Persistent CHP episode ≥ 2 min (crude)	0.94 (0.76–1.17)
Persistent CHP episode ≥ 2 min (adjusted)	0.98 (0.74–1.29)
Any CHP window > 0 min (crude)	1.37 (0.97–1.95)
CHP ≥ 30 min (crude)	**1.40 (1.05–1.86)**
CHP ≥ 30 min (adjusted)	1.31 (0.91–1.89)
No-hypotension subgroup (crude)	0.84 (0.62–1.14)
No-hypotension subgroup (adjusted)	0.85 (0.57–1.27)
MAP < 65 mmHg (crude)	**1.82 (1.46–2.26)**
MAP < 65 mmHg (adjusted)	**1.62 (1.21–2.16)**
Median MAP < 70 mmHg (crude)	**3.09 (2.28–4.19)**







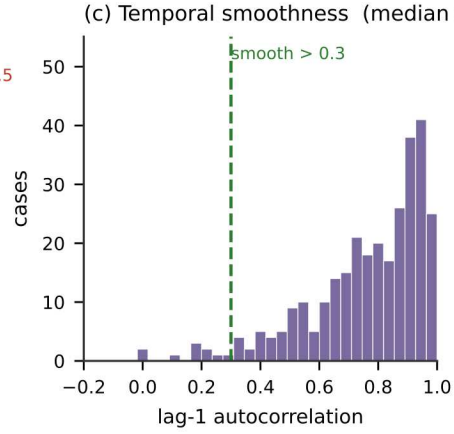
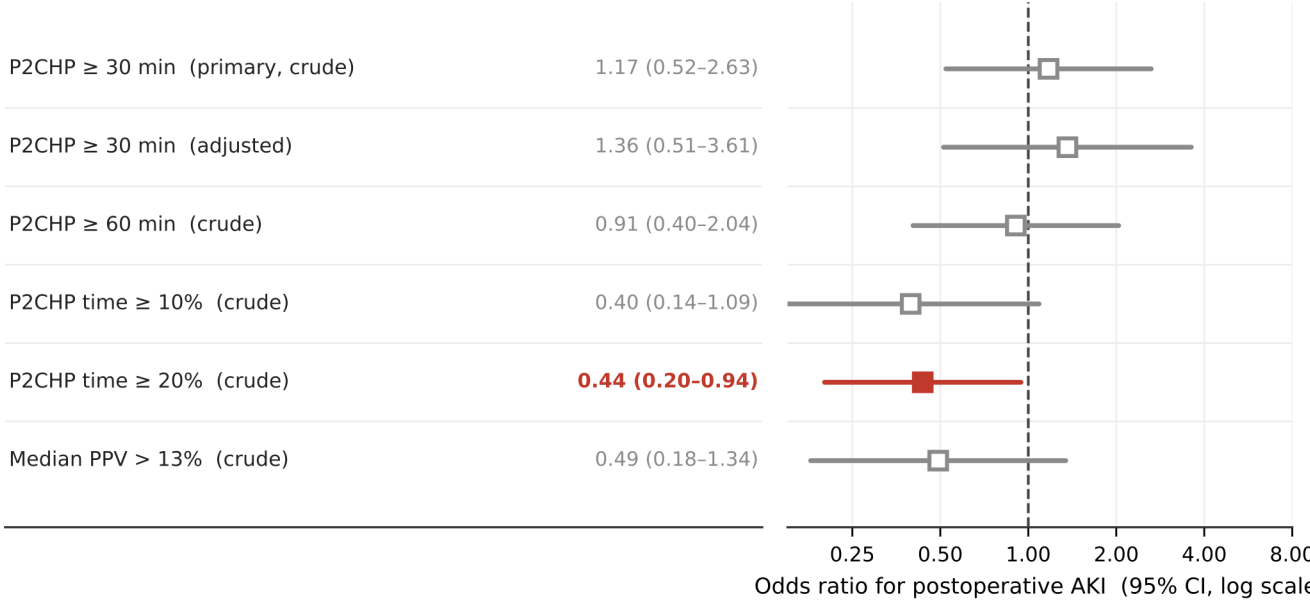


Fig 4. Pulse pressure variation (PPV) and AKI — 300-case pilot

Analysis set n = 241 (AKI 30, 12.4%). Underpowered pilot — direction only; continuous PPV |ρ| ≤ 0.09.



Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryTableS1channelenumeration.csv
- SupplementaryMaterial.docx