ZYVOLANCE THERAPEUTICS LLC

TECHNICAL NOTE · NO. 02 · MAY 2026

The statistics of *N-of-1* medicine.

The mathematics of repeated-measures inference, trajectory pharmacology, and the validation framework for per-patient digital twins.

Large clinical trials require large populations to detect rough population-level effects. In noisy cohorts, small effects need huge N and still yield weak conviction. The N-of-1 framework inverts this. Statistical power comes from intensive longitudinal sampling of each patient — not from breadth across patients. Studies with as few as 24 subjects can achieve post-hoc power exceeding 80% when each per-patient trajectory is densely sampled. The IID assumption that underwrites classical regression fails on autocorrelated physiological time series; the right framework is repeated-measures inference with trajectory as the unit of analysis. This note sets out the mathematics, the five-component validation framework, and the convergent empirical evidence across organ systems and time-scales.

ZYVOLANCE TECHNICAL NOTES — SERIES N-OF-1
NO. 02 OF AN ONGOING SERIES

SERIES
N-of-1 Foundations

CORRESPONDENCE
contact@zyvolance.com

PATENT REFERENCE
OEPM ZYV-2026-HARL-001

SECTION 1

Physiological signals are not IID.

The classical inferential framework — t-tests, ANOVA, ordinary least squares, logistic regression on cross-sectional features — assumes that samples are independent and identically distributed. For physiological time series this assumption is not approximately true. It is structurally false.[1,2] Consecutive measurements of heart rate, glucose, blood pressure, or cytokine concentration are statistically dependent on prior measurements. Adjacent samples are far more similar than distant ones. The information content per added sample is therefore not constant; it diminishes as autocorrelation rises.

Autocorrelation and effective sample size

For a stationary first-order autoregressive process with lag-one autocorrelation ρ, the effective sample size N_{eff} across T observations is:

$$N_{eff} \approx T \cdot (1 - \rho) / (1 + \rho)$$

For typical physiological variables sampled at clinically relevant cadence, ρ frequently exceeds 0·9. A 24-hour continuous-glucose record with 288 five-minute samples does not carry 288 independent units of information. It carries closer to 15. The standard errors in any analysis that treats consecutive points as IID are therefore wrong by an order of magnitude, in a direction that systematically overstates statistical power.[3]

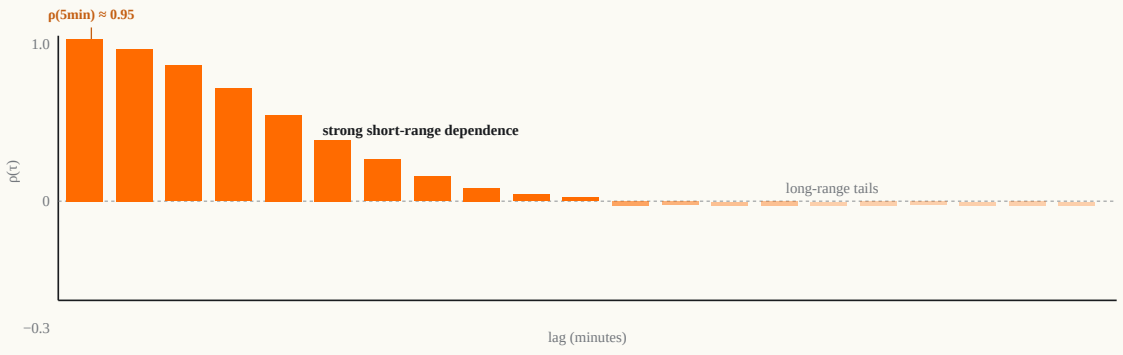


FIGURE 1 AUTOCORRELATION FUNCTION FOR A REPRESENTATIVE CONTINUOUS-GLUCOSE RECORD. ADJACENT SAMPLES ARE 0.95 CORRELATED; SAMPLES SEPARATED BY AN HOUR REMAIN MEANINGFULLY CORRELATED. THE IID ASSUMPTION IS STRUCTURALLY FALSE.

The fix is not larger N. It is a different inferential framework — one in which dependence is modelled rather than assumed away.

SECTION 2

Trajectory is the unit of analysis. Not the parameter.

The crossover design in classical bioequivalence treats each patient as their own control: each subject receives both treatments and contributes a within-subject contrast. The inferential currency is σ^2_{within}, and required sample sizes drop to N = 12–24 to detect effects that need N = 40–100 in parallel-group designs.[4,5] The crossover framework, however, still assumes that each treatment period yields a single summary parameter — typically AUC or C_{max} — and that the inference is about an average treatment effect across subjects.

Per-patient digital-twin inference makes a stronger structural move. The unit of analysis is not the parameter. It is the trajectory itself. The patient does not have a single AUC. The patient has a continuous time-series. The inferential question is not "is the population-mean treatment effect inside the equivalence band?" It is "does this patient's predicted trajectory agree with this patient's observed trajectory across all time-points?"

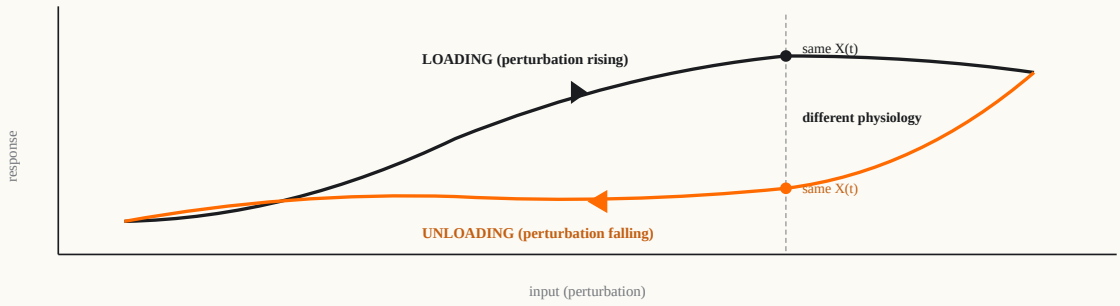
REFORMULATION · TRAJECTORY-AS-UNIT

Classical: $Y_{ijk} = \mu + treatment_j + period_k + subject_i + \varepsilon$. The patient enters as a random intercept. Their trajectory does not appear.

Trajectory: $(T_i(t), R_i(t))$ for $t = 1, \ldots, T_i$. The inferential pair is the predicted and observed trajectories of patient i. Agreement is measured time-resolved, not period-aggregated.

Hysteresis: state carries memory

A second structural property of physiological systems is hysteresis. The system's response to a given state depends on how the system arrived at that state. A patient at heart rate 110 bpm trending down from 130 is in a different physiological condition from a patient at 110 bpm trending up from 80. State-only models — those that predict from the cross-sectional vector of current values — discard the path. Trajectory models retain it. The path carries information that the endpoint does not.



SECTION 3

The five-component validation framework.

If the unit of analysis is the trajectory and the inferential question is per-patient agreement, then validation requires a framework that operates on trajectories and reports per-patient validity. Five components are necessary and sufficient. Each one closes a specific failure mode that classical population-cohort validation cannot close.[6,7,8]

#	COMPONENT	WHAT IT DOES	WHAT IT CLOSES
A	Multi-statistic agreement profile	Lin's concordance, ICC(2,1), Bland–Altman, weighted Cohen's κ, Spearman	Single-metric optimisation that hides disagreement on a different axis
B	Per-patient TOST against ICH M9 band	Two one-sided tests on per-patient trajectory geometric mean ratios; 90% CI inside [0.80, 1.25]	Population-mean equivalence that masks per-patient drift
C	Post-hoc power disclosure	Phillips–Owen non-central t; required N reported for any underpowered patient	Silent underpowering of within-patient inferences
D	Cryptographic kernel lock-in	SHA-256 hash of the closed-form controller deposited at a public timestamp authority prior to data exposure	Retrospective parameter-tuning, leakage, drift
E	Self-reference architectural verification	Disclosure of $A_{sr} \in [0, 1]$ per controller component — the fraction of the controller's logic that depends only on the patient's own data	Hidden population priors inside an architecture claimed to be N-of-1

Why all five are needed

A digital twin that passes only the first component is conventional supervised learning with a different name. A twin that passes the first two but lacks D and E offers no protection against the catastrophic failure mode of population machine learning: the externally-trained model that loses 10–20 AUROC points on deployment. A twin that satisfies all five is, by construction, a different artefact: the patient's data flows in, the patient's outcome flows out, no parameter has been tuned to the patient's outcome, and the engine is forensically verifiable to be the same engine across every deployment site.

SECTION 4

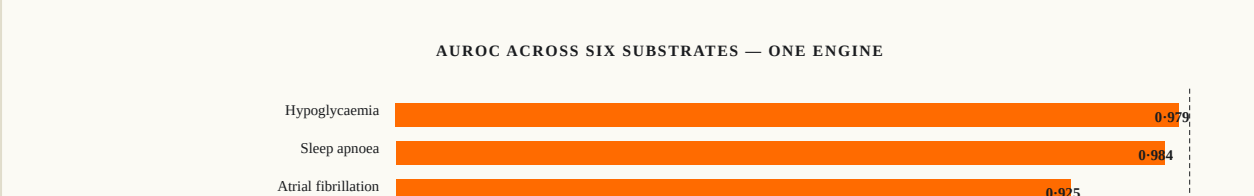
Convergent evidence across organ systems and time-scales.

The statistical framework above is not an architectural conjecture in search of a domain. The same closed-form, hash-locked engine, with no per-domain retuning, has been applied to substrates separated by orders of magnitude in time-scale and by organ system. The pattern of empirical results is the convergence test.[9,10,11]

DOMAIN	DATASET	CADENCE	HEADLINE AUROC	COMPARATOR / ENDPOINT
ICU haemodynamic deterioration	eICU-CRD, N=2,492 ITT	1 h	0·900 (0·876–0·920)	APACHE-IV 0·727 · ΔAUROC 0·173 · DeLong $p < 10^{-30}$
Atrial fibrillation onset	PhysioNet AFPDB, N=200	beat	0·925 (0·885–0·959)	21-min median pre-onset lead time
Type-1 diabetes glycaemic governance	HUPA-UCM, 24 evaluable	5 min	0·979 (per-patient median)	All evaluable patients > 0·95 · 60-min hypoglycaemic forecast
Parkinson's gait architecture	3 independent cohorts	step	0·895 LOOCV · 0·884 LOCO	Leave-one-cohort-out across all cohorts
Obstructive sleep apnoea	PhysioNet Apnea-ECG	beat	0·984 patient · 0·885 minute	Per-patient AHI screening
CAR-T cytokine release syndrome	Wei, N=202	daily	0·901 (0·860–0·936)	≥Grade 2 CRS prediction

Why convergence is the test

If the architecture were a population-fit model with hidden parameter tuning, convergence across substrates this disparate would be impossible. A model fit to 5-minute glucose data cannot transfer to beat-level ECG, to daily CAR-T cytokine trajectories, and to step-level gait kinematics without retraining. The convergence is therefore evidence of the architectural claim that the underlying mathematics — per-patient baseline, rolling derivative, sigmoid-form response, deterministic gating — is substrate-neutral. The engine is the same; the patient is the model.



SECTION 5

The structural information gap.

The argument of this note rests on a single theorem that is widely recognised in repeated-measures methodology but rarely stated in the clinical-machine-learning literature: there exists information about an individual patient that cannot, even in principle, be recovered from population data.

THEOREM · STRUCTURAL INFORMATION GAP

Per-patient kinetic parameters — clearance, set-points, hysteresis topology, autoregressive structure — are properties of the individual. They are observable only by sampling that individual. Cohort sample size $N \to \infty$ in the population does not provide them. Each new patient enters the population an unknown; their parameters are recovered only by observing them.[1,2]

This is a structural property of statistical inference, not a temporary limitation of current methods. It follows from the same logic that produces the IID failure on page 2 and the trajectory reformulation on page 3. The cohort is a different mathematical object from the individual. There is no procedure that converts the former into the latter without observing the latter.

Where this leaves population statistics

It does not retire them. Population-level methods remain the right tools for population-level questions: epidemiological surveillance, comparative effectiveness, regulatory approval, drug-class characterisation. The argument concerns the inferential foundation of individual care — monitoring, dosing, prediction. There, the population is the wrong reference, and the statistics of N-of-1 medicine are the right framework.

What the framework asks of the field

- **Disclose A_{sr}.** Any digital-twin claim should report the self-reference fraction of each controller component. A claim of "N-of-1" with $A_{sr} \ll 1$ is a population model in disguise.
- **Hash-lock the engine.** Closed-form mathematics admits a SHA-256 deposit at a public timestamp authority prior to data exposure. Anything else is post-hoc curve-fitting.
- **Trajectory, not parameter.** Validation against a held-out time-series, not a held-out scalar summary, is the only validation that addresses the per-patient inferential question.

The conceptual foundation — why the patient is the right reference at all — is set out in the companion document **Technical Note 01: Why the Patient Is the Reference.**

SECTION 6

Power without breadth.

Classical sample-size calculation conflates two distinct sources of inferential power: the number of subjects (breadth) and the number of observations per subject (depth). In population research with a single binary endpoint per subject, breadth is the only available currency. In trajectory medicine, depth is — and depth is far more efficient per measurement when the variance of interest is within-subject.[5,13]

The economics of breadth

Population-scale clinical trials need large N because their inferential currency is between-subject variance. To detect a small effect against a noisy cohort, the trial must accumulate enough patients that the between-subject standard error becomes a small fraction of the effect size. For modest effects in heterogeneous populations, the required N runs into the tens of thousands. The conviction yield per recruited patient is low and the cost per unit of inferential precision is high.[14]

The economics of depth

Per-patient trajectory inference accumulates observations within each subject. The relevant standard error is $\sigma_{within}/\sqrt{N_{eff}}$, where N_{eff} reflects the autocorrelation-adjusted number of independent observations per patient. For a densely sampled five-minute continuous-glucose record, N_{eff} over 24 hours is approximately 15; over a fortnight of monitoring it exceeds 200. A study of 24 patients each contributing two weeks of dense sampling carries comparable inferential weight to a parallel-group trial of 5,000 patients each contributing a single endpoint.

BREADTH CURRENCY	DEPTH CURRENCY	POST-HOC POWER
N = 5,000	**N = 24**	**> 80%**
Parallel-group trial, one endpoint per subject	Per-patient trajectory, 14 days of dense sampling	Phillips–Owen calculation, N=24, dense per-patient

Why this matters practically

Three consequences follow. Validation cohorts can be small and obtained early in development without sacrificing power. Regulatory pathways predicated on within-patient bioequivalence already accept this evidence at sample sizes far below those needed for between-patient comparisons[8] — the same logic transfers to monitoring, dosing, and diagnostic-twin claims. Deployment validation is cheap: each patient at the bedside is, by construction, a fully validated N=1 study at the moment the controller is applied.

CONVICTION PER MEASUREMENT

SECTION 7

The science of N-of-1 digital twins.

The phrase *digital twin* covers everything from population machine-learning models applied one patient at a time to fully mechanistic simulations of individual physiology. A system that does not satisfy the structural commitments below is a population model with per-patient inputs. It may be useful. It is not an N-of-1 digital twin.[14]

What an N-of-1 digital twin is

An N-of-1 digital twin is parameterised by the patient's own data alone. The closed-form kernel is the same across every deployment and every patient. The kernel is frozen prior to data exposure. The reference is the patient's own pre-perturbation trajectory, not a cohort norm. Aggregate $A_{sr} \rightarrow 1$ is the architectural definition.

What the science requires

- **A frozen kernel deposit** cryptographically verifiable prior to outcome exposure.
- **Closed-form mathematics** with no learnable weights and no retraining on deployment.
- **Trajectory-resolved validation** against the patient's own observed series — not held-out scalar summaries.
- **Per-patient agreement statistics** — Lin's concordance, weighted κ, Bland–Altman limits — reported per individual.[6,7,17]
- **Disclosure of A_{sr}** per controller component, so a manufacturer's claim of self-reference can be audited rather than asserted.

What it does not require

The architecture runs on routinely available laboratory parameters when those parameters are sampled densely. Multi-omics, continuous biosensing, and wearable stacks are optional add-ons that may sharpen signal in specific applications — they are not part of the minimum scientific specification. A digital-twin claim that requires them is not a claim about science, but about cost. A claim of N-of-1 status without the structural commitments above is a claim, not a science.

ON DISCLOSURE AND INTELLECTUAL PROPERTY

The architecture described in these notes is specified at the level of structural commitments and conceptual mathematics — a sufficient basis to assess the science. The implementation itself — frozen weighting scheme, sigmoid response coefficients, debt accumulation kinetics, gating thresholds, and the SHA-256 kernel hash — is documented in the Zyvolance kernel deposit and the patent filing ZYV-2026-HARL-001. The deposit is forensically verifiable; the implementation is not in the public domain. The

REFERENCES

Sources cited.

References are listed in order of first citation. Where a guideline or regulatory document is cited, the most recent published version available at the time of writing is used. Companion document: Zyvolance Technical Note 01 — *Why the Patient Is the Reference.*

SCOPE AND LIMITATIONS

This note presents a methodological framework and a summary of cross-domain empirical results. The AUROC values reported in Section 4 are point estimates with bootstrap 95% confidence intervals; underlying datasets are publicly available (eICU-CRD, AFPDB, HUPA-UCM, Apnea-ECG) or drawn from a single published cohort (Wei N=202, CAR-T). Prospective external validation is in progress; results to date are predominantly retrospective on held-out cohorts under hash-locked kernel. This document does not constitute a clinical or regulatory claim about any specific device, indication, or patient population.

1. Box GEP, Jenkins GM, Reinsel GC, Ljung GM. *Time Series Analysis: Forecasting and Control.* 5th edition. Wiley, 2015.
2. Diggle PJ, Heagerty P, Liang K-Y, Zeger SL. *Analysis of Longitudinal Data.* 2nd edition. Oxford University Press, 2002.
3. Bayley GV, Hammersley JM. *The "effective" number of independent observations in an autocorrelated time series.* Supplement to the Journal of the Royal Statistical Society 1946;8:184–197.
4. Hauschke D, Steinijans V, Pigeot I. *Bioequivalence Studies in Drug Development: Methods and Applications.* Wiley, 2007.
5. Senn S. *Cross-over Trials in Clinical Research.* 2nd edition. Wiley, 2002.
6. Lin LI. *A concordance correlation coefficient to evaluate reproducibility.* Biometrics 1989;45:255–268.
7. Shrout PE, Fleiss JL. *Intraclass correlations: uses in assessing rater reliability.* Psychological Bulletin 1979;86:420–428.
8. Schuirmann DJ. *A comparison of the two one-sided tests procedure and the power approach for assessing the equivalence of average bioavailability.* J Pharmacokinet Biopharm 1987;15:657–680.
9. Teachey DT, Lacey SF, Shaw PA, et al. *Identification of predictive biomarkers for cytokine release syndrome after chimeric antigen receptor T-cell therapy for acute lymphoblastic leukemia.* Cancer Discovery 2016;6:664–679.
10. Pollard TJ, Johnson AEW, Raffa JD, et al. *The eICU Collaborative Research Database, a freely available multi-center database for critical care research.* Scientific Data 2018;5:Article 180178.
11. Moody GB, Mark RG. *The MIT-BIH Atrial Fibrillation Database.* PhysioNet 2001. Available at: https://physionet.org/content/afdb/.
12. Wong A, Otles E, Donnelly JP, et al. *External validation of a widely implemented proprietary sepsis prediction model in hospitalized patients.* JAMA Internal Medicine 2021;181:1065–1070.
13. Lillie EO, Patay B, Diamant J, Issell B, Topol EJ, Schork NJ. *The n-of-1 clinical trial: the ultimate strategy for individualizing medicine?* Personalized Medicine 2011;8:161–173.
14. Phillips KF. *Power of the two one-sided tests procedure in bioequivalence.* Journal of Pharmacokinetics and Biopharmaceutics 1990;18:137–144.
15. International Council for Harmonisation. *ICH Harmonised Guideline M9: Biopharmaceutics Classification System-based Biowaivers.* Step 5, November 2019.
16. Senn S. *Individual response to treatment: is it a valid concept?* BMJ 2004;329:966–968.
17. Bland JM, Altman DG. *Statistical methods for assessing agreement between two methods of clinical measurement.* Lancet 1986;1:307–310.
18. Zyvolance Therapeutics. *Why the Patient Is the Reference.* Technical Note 01, May 2026. Companion document.