Computational validation of a population-pharmacokinetic exposure engine: a pre-registered reproduction study for three antimicrobials

Youssef Ghaly
GATMEDI
Correspondence: yg@gatmedi.org
Preprint — version 1 July 2026.

Abstract

Model-informed dose individualisation depends on a computational engine that computes drug exposure from a pharmacokinetic model and a patient's covariates; an error in that engine propagates to every downstream recommendation, so its correctness is a precondition for any individualisation claim. We report a pre-registered computational validation of the deterministic exposure engine underlying a substrate-based dose-support system. The protocol, four checks, numeric bars, and case set were fixed and cryptographically committed before any result was generated, and the engine was content-hashed and produced deterministic output. Across seven quantitative cases spanning three antimicrobials, the engine reproduced published exposure values within the pre-registered 2% tolerance, with a maximum relative error of 1.77%; three further reference-uncertain cases met their published pharmacodynamic targets. Numerical integration of the compartmental model agreed with the closed-form analytic solution to within 0.1% (maximum 0.077%), and reference-subject parameter identity was exact. Every case presenting missing or unusable model inputs was refused rather than answered (ten of ten), with no numeric result emitted. Repeated execution produced an identical content hash. The study establishes model-implementation fidelity — that the engine computes exposure correctly given published parameters — and does not address prospective accuracy against new patient data or any clinical-outcome measure; beta-lactam target attainment additionally rests on an assumed unbound fraction, and coverage is three antimicrobials. Correct, refusal-safe exposure computation is thereby demonstrated as the foundation layer on which individualisation is built.

1. Introduction

The pharmacokinetics of narrow-therapeutic-index antimicrobials vary widely between critically ill patients, and dosing to a pharmacodynamic target therefore requires computing a patient's predicted exposure — the area under the concentration-time curve, the elimination half-life, or the fraction of the dosing interval with free concentration above the minimum inhibitory concentration — from a population pharmacokinetic model and the patient's covariates. Any individualisation or decision-support method built on such computation inherits its arithmetic: an error in the exposure engine propagates unseen into every dose it recommends. The correctness of that engine, given a model and its parameters, is therefore a precondition that must be established before individualisation performance can be assessed.

The methods the engine implements are standard and public: compartmental pharmacokinetic models fitted by nonlinear mixed-effects estimation, numerical integration of the resulting ordinary-differential-equation system, and the pharmacokinetic/pharmacodynamic indices used in clinical practice (an AUC

over MIC target for vancomycin, a percentage of the interval above the MIC for beta-lactams). This study does not claim novelty in any of these. The object under test is a particular deterministic exposure engine, and the claim assessed is model-implementation fidelity: that, given a published model and its parameters, the engine reproduces the published exposures and refuses when its inputs are insufficient. The engine and the compiled substrate from which it reads parameters are proprietary; the present validation concerns only observable input-output behaviour, checked against the public literature.

The scope is deliberately narrow. The engine was validated for three antimicrobials at an early stage of compilation, well below the intended scale of the underlying substrate; the claims here are bounded accordingly and make no statement about drugs or populations outside the tested set. What follows sets out the pre-registration discipline, the case set and its fixed denominators, the four checks and their bars, and the results, including the cases the engine declined and the comparators that fell outside the gating tolerance.

2. Methods

2.1 Scope of disclosure

The population pharmacokinetic models used as inputs are published and are cited so that the validation is independently checkable. The exposure engine, and the compiled substrate from which it reads parameters, are proprietary and out of scope; the descriptions below are confined to standard, publicly documented pharmacometric computation and do not describe how the substrate or its parameter sets are assembled.

2.2 Pre-registration and reproducibility

The four checks, every numeric bar, and the full case set were written down and cryptographically committed before any result was generated; the committed threshold file is identified by its SHA-256 digest in the Appendix. The engine was content-hashed prior to the run, and its outputs are deterministic: repeated execution reproduced an identical content hash. No bar was moved and no case was added or removed after results were seen. Provenance identifiers — engine, threshold-file, parameter-file, and result content hashes, and software versions — are listed in the Appendix.

2.3 Case set and models

The case set comprised twelve evaluated cases and ten refusal cases. Quantitative reproduction was assessed against a published worked example for vancomycin (Wei 2022, a neurosurgical population) and against published model-derived half-lives for piperacillin (Klastrup 2020, continuous infusion); qualitative target attainment was assessed for piperacillin (Hahn 2021, extracorporeal membrane oxygenation) and meropenem (Peng 2025, continuous renal replacement therapy), which are reported as reference-uncertain because no single published exposure number was locally pinnable for them. Two reference subjects were used to check parameter identity. The ten refusal cases were published records deliberately flagged as unusable for simulation — for reasons including parameters given only in supplementary form, garbled or unextractable units, physiologically implausible volumes, or a missing required covariate. The pre-registered denominators were seven quantitative cases, three attainment cases, two identity cases, and ten refusal cases.

2.4 Metrics and pre-registered bars

Four checks were specified. Reproduction accuracy required the engine's computed exposure to fall within 2% of the published value for every quantitative case, with the reference-uncertain cases instead required to meet their published pharmacodynamic target. Numerical integrity required the numerically integrated exposure to agree with the closed-form analytic solution to within 0.1%, and required reference-subject parameter identity to be exact. Refusal safety required every unusable-input case to raise an explicit error with no numeric result emitted (100%). Reproducibility required repeated execution to produce an identical content hash.

3. Results

3.1 Reproduction accuracy

All seven quantitative cases fell within the pre-registered 2% tolerance, with a maximum relative error of 1.77% (Table 1, Figure 1). The four vancomycin cases reproduced the published AUC over 24 h to within 0.63 to 1.77% across the mannitol and no-mannitol covariate settings and the AUC 400 and 600 targets, and the three piperacillin cases reproduced the model-derived half-life to within 0 to 0.84% across creatinine-clearance values of 30, 80, and 130 mL/min.

Case	Drug	Metric	Computed	Published	Rel. error %	Bar %
Wei AUC_{24} 400 (mannitol)	Vancomycin	AUC_{24} (mg·h/L)	397.5	400.0	0.63	2.0
Wei AUC_{24} 600 (mannitol)	Vancomycin	AUC_{24} (mg·h/L)	589.4	600.0	1.77	2.0
Wei AUC_{24} 400	Vancomycin	AUC_{24} (mg·h/L)	405.9	400.0	1.48	2.0
Wei AUC_{24} 600	Vancomycin	AUC_{24} (mg·h/L)	593.2	600.0	1.13	2.0
Klastrup t½ (CrCL 30)	Piperacillin	t½ (h)	4.264	4.3	0.84	2.0
Klastrup t½ (CrCL 80)	Piperacillin	t½ (h)	2.108	2.1	0.38	2.0
Klastrup t½ (CrCL 130)	Piperacillin	t½ (h)	1.4	1.4	0.00	2.0

Table 1. Quantitative reproduction of published exposures (pre-registered bar: relative error ≤ 2%). All seven cases pass; the maximum relative error is 1.77%.

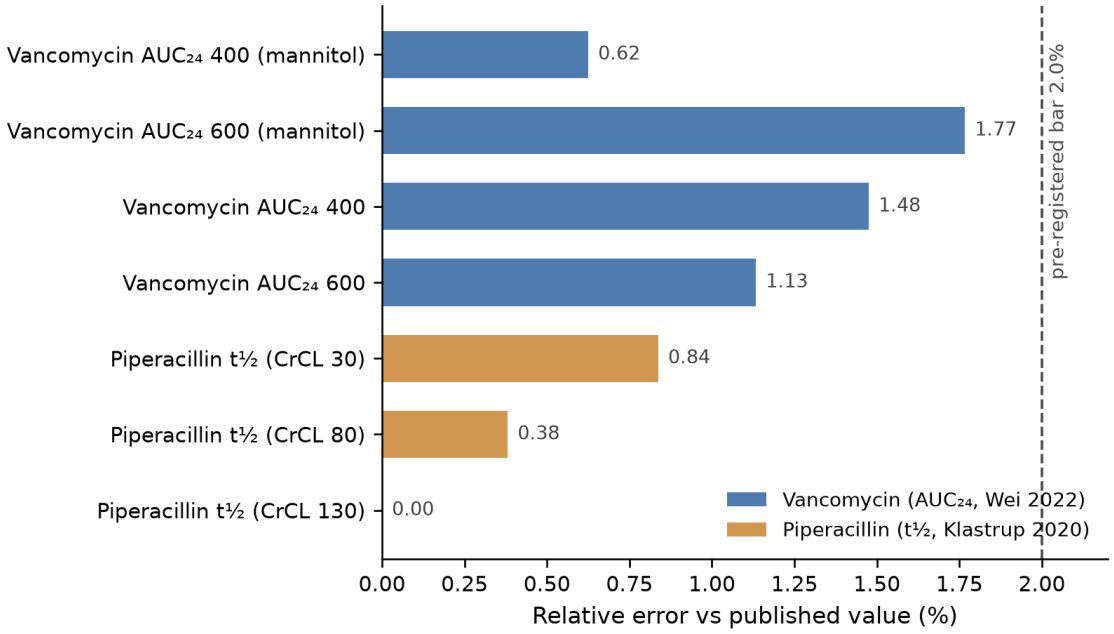

Figure 1. Reproduction accuracy. Relative error of the engine's computed exposure against the published value for each quantitative case, with the pre-registered 2% bar. All seven cases fall below the bar.

3.2 Target attainment (reference-uncertain cases)

The three reference-uncertain cases each met their published pharmacodynamic target, reaching 100% of the dosing interval with free concentration above the MIC (Table 2). For the meropenem continuous-renal-replacement cases the clearance covariate coefficient could not be extracted from the source (it is illegible in the source table), so these cases were computed at the renal-replacement-population typical clearance rather than individualised within the creatinine-clearance range; this is recorded as a per-case flag and in the Limitations.

Case	Drug	Computed %fT>MIC	Target	Meets target
Hahn continuous 16 g	Piperacillin	100.0	100%	yes
Peng 1 g q8h (3 h infusion)	Meropenem	100.0	100%	yes
Peng 500 mg q6h (3 h infusion)	Meropenem	100.0	100%	yes

Table 2. Reference-uncertain attainment cases. Each meets its published %fT>MIC target. These cases are excluded from the ≤2% quantitative set because no single published exposure value was locally pinnable.

3.3 Numerical integrity

The numerically integrated exposure agreed with the closed-form analytic solution to within the pre-registered 0.1% across all evaluated cases, with a maximum divergence of 0.077% for the meropenem 500 mg q6h case (Figure 2); the continuous-infusion cases agreed exactly. Parameter identity for the two reference subjects was exact, with a maximum absolute deviation of zero.

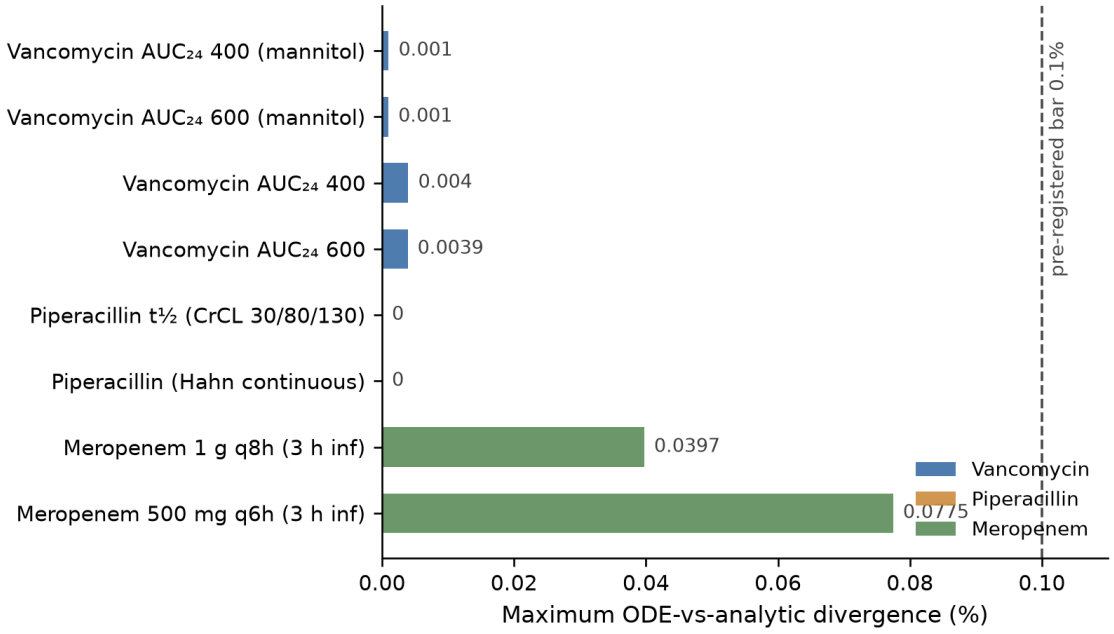

Figure 2. Numerical integrity. Maximum divergence between the numerically integrated and the closed-form analytic exposure for each case, against the pre-registered 0.1% bar. The largest divergence is 0.077%; reference-subject parameter identity (not shown) was exact.

3.4 Refusal safety

Every one of the ten unusable-input cases was refused, nine raising a model-blocked error and one a parameter error, and no numeric result was emitted in any case (Table 3). The refusals correspond to published records the engine treats as unusable — because their parameters are supplementary-only, their units unextractable, their volumes physiologically implausible, or a required covariate is absent — together with one case in which a required covariate was withheld from an otherwise usable model.

Record presented	Expected	Outcome	Number emitted
crrt-dosing-review (Li 2020)	ModelBlocked	ModelBlocked	no
meropenem-peng-2022-crrt-meta	ModelBlocked	ModelBlocked	no
meropenem-rancic-2024	ModelBlocked	ModelBlocked	no
meropenem-smarrt-roberts-2025	ModelBlocked	ModelBlocked	no
meropenem-tao-2025	ModelBlocked	ModelBlocked	no
meropenem-zhao-2022	ModelBlocked	ModelBlocked	no
piptazo-kong-2025	ModelBlocked	ModelBlocked	no
piptazo-smarrt-roberts-2025	ModelBlocked	ModelBlocked	no
vancomycin-tesfamariam-2024	ModelBlocked	ModelBlocked	no
vancomycin-wei-2022 (missing covariate)	ParameterError	ParameterError	no

Table 3. Refusal-safety matrix. Ten records with missing or unusable inputs; all are refused with the expected error and no numeric result (pre-registered bar: 100%).

3.5 Non-gating comparators

Two additional beta-lactam comparators from continuous-venovenous-haemofiltration non-compartmental analyses were computed but were not part of the gating tolerance, because small-sample non-compartmental medians are not internally self-consistent to 2%. For meropenem, the computed steady-state concentration differed from the measured value by 3.66% and the corresponding exposure by 3.53%; for piperacillin, the steady-state concentration check was internally circular and a weight-free exposure-consistency comparison differed by 10.8%, consistent with sampling before steady state. These comparators are reported to document the additional data rather than as validation of model fidelity.

3.6 Reproducibility

Repeated execution of the locked engine and harness produced a byte-identical content hash, and the committed locked reference values were intact.

4. Discussion

The engine reproduced published antimicrobial exposures within a pre-registered 2% tolerance, integrated the compartmental model in agreement with the analytic solution to within 0.1%, and refused every case whose inputs it could not use. Together these establish model-implementation fidelity: the arithmetic on which any downstream individualisation depends is correct for the tested models, and the engine does not emit a number when it lacks the inputs to compute one. The refusal behaviour is a property of the engine observable from the outside — it declines unusable published records rather than interpolating or guessing — and it distinguishes such an engine from tools that return a value regardless of input adequacy.

The result is a foundation rather than an endpoint. It is the precondition for the individualisation work reported separately, in which the same engine is used to update a population prior with a patient's own measurements; it does not by itself demonstrate accuracy against prospectively measured patient

concentrations, and it is bounded to the three antimicrobials and the models tested here. Reproduction of a published model's own outputs confirms faithful implementation, not the external validity of the underlying model.

5. Limitations

1. The engine optimises a pharmacokinetic/pharmacodynamic surrogate (AUC over MIC, or percentage of the interval above the MIC), not mortality or any clinical outcome; it is decision-support, not a prescription.
2. Reproduction demonstrates model-implementation fidelity — that the engine correctly implements published models — and not prospective accuracy against newly measured patient concentrations.
3. Beta-lactam target attainment rests on an assumed unbound fraction, which is a modelling assumption surfaced per case rather than a per-patient measurement.
4. Coverage is three antimicrobials (vancomycin, piperacillin, meropenem); no claim is made for other drugs, which were not tested.
5. For the meropenem continuous-renal-replacement cases, the clearance covariate coefficient was not extractable from the source, so those cases were computed at the population-typical clearance without within-range individualisation.
6. The two additional non-compartmental comparators were not internally self-consistent to 2% and are reported as non-gating; they neither pass nor contribute to the validation verdict.
7. The validation uses population priors and a deterministic forward computation; individual Bayesian forecasting is a separate capability assessed elsewhere.

6. Conclusion

Under a pre-registered and reproducible protocol, a deterministic exposure engine reproduced published antimicrobial exposures within 2%, integrated the compartmental model in agreement with the analytic solution within 0.1%, and refused every unusable-input case without emitting a number. Correct, refusal-safe exposure computation is thereby demonstrated for three antimicrobials as the foundation layer on which model-informed individualisation is built; prospective validation against measured patient data remains a separate, later step.

Competing interests

The author is the founder of GATMEDI, the company developing the engine evaluated in this study.

Funding

This work was funded internally by GATMEDI. It received no external funding.

Data and code availability

The engine, its source code, and the underlying compiled substrate are proprietary and are not available. The study used the published population-pharmacokinetic models cited in the References, which are

publicly available; the engine and the evaluation harness that computed and checked the exposures are not.

Ethics

This study used only computational reproduction and checking against published models. It involved no human subjects and no patient data, and did not require ethics approval.

References

1. Wei S, Zhang D, Zhao Z, Mei S. Population pharmacokinetic model of vancomycin in postoperative neurosurgical patients. Front Pharmacol. 2022;13:1005791. doi:10.3389/fphar.2022.1005791.
2. Klastrup V, Thorsted A, Storgaard M, Christensen S, Friberg LE, Öbrink-Hansen K. Population pharmacokinetics of piperacillin following continuous infusion in critically ill patients and impact of renal function on target attainment. Antimicrob Agents Chemother. 2020;64(7):e02556-19. doi:10.1128/AAC.02556-19.
3. Hahn J, Min KL, Kang S, Yang S, Park MS, Wi J, Chang MJ. Population pharmacokinetics and dosing optimization of piperacillin-tazobactam in critically ill patients on extracorporeal membrane oxygenation and the influence of concomitant renal replacement therapy. Microbiol Spectr. 2021;9(3):e00633-21. doi:10.1128/Spectrum.00633-21.
4. Peng Y, Liu Y, Cheng Z, Zhang Q, Xie F, Zhu S, Li S. Population pharmacokinetics of prolonged infusion for meropenem: tailoring dosing recommendations for Chinese critically ill patients on continuous renal replacement therapy with consideration for renal function. Drug Des Devel Ther. 2025;19:1105-1117. doi:10.2147/DDDT.S489603.
5. Rybak MJ, Le J, Lodise TP, et al. Therapeutic monitoring of vancomycin for serious methicillin-resistant Staphylococcus aureus infections: a revised consensus guideline and review by ASHP, IDSA, PIDS, and SIDP. Am J Health-Syst Pharm. 2020;77(11):835-864. doi:10.1093/ajhp/zxaa036.
6. European Committee on Antimicrobial Susceptibility Testing (EUCAST). Breakpoint tables for interpretation of MICs and zone diameters, version 16, 2026. https://www.eucast.org/clinical_breakpoints.
7. Virtanen P, Gommers R, Oliphant TE, et al.; SciPy 1.0 Contributors. SciPy 1.0: fundamental algorithms for scientific computing in Python. Nat Methods. 2020;17(3):261-272. doi:10.1038/s41592-019-0686-2.

Appendix. Provenance and reproducibility

The run was banked from a frozen, content-hashed engine under a pre-registered, SHA-256-locked threshold file. Deterministic outputs reproduce on repeated execution.

Item	Value
Result content hash (SHA-256)	49ee1f1c7ff64e8221e9d6b99357a2568219794b3457bf3554b33543f9b79821
Locked threshold file (SHA-256)	16ed9ee39d60e69d9cda353eea8ca7b7f019c4b9d03d1c02dc2a2cefa045e6b9
Exposure-engine core (SHA-256)	f78d70dbee80127c431313e41b66b25237e7f24fa5f7a7c9f1e979d19849f915
Harness run_validation (SHA-256)	6be03a69c7c300a9d20e6543872a1be4d65fcba6f845497f604e1499947db187
Parameter file (SHA-256)	f238ae50c2a210654aa52eba051eb463d1f5db3be31973ada4a40c4fc9fab5a9
Random seed	0
Denominators	7 quantitative + 3 attainment + 2 identity + 10 refusal cases
Wall-clock runtime	19.834 s
Software	Python 3.12.4; NumPy 2.0.1; SciPy 1.17.1; PyYAML 6.0.3