



Universiteit
Leiden
The Netherlands

Virtual patient populations for quantitative pharmacology with its application in antibiotic treatment optimization

Guo, Y.

Citation

Guo, Y. (2026, September 11). *Virtual patient populations for quantitative pharmacology with its application in antibiotic treatment optimization*.

Retrieved from <https://hdl.handle.net/1887/4309451>

Version: Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/4309451>

Note: To cite this publication please use the final published version (if applicable).



Chapter 7

General discussion and summary

Modelling and simulation (M&S) in quantitative pharmacology play an important role in informing drug development and treatment optimization. In this context, accounting for inter-individual variability in pharmacokinetics (PK) and pharmacodynamics (PD) is an essential component of such M&S workflows for obtaining reliable and clinically meaningful simulation outcomes. Inter-individual variability can be associated with both patient-associated covariates predictive of variability which are identified in M&S workflows, as well as accounting for additional sources of variation in PK and PD, which can be understood through mechanism-based modelling approaches. In this thesis, we developed and applied systematic approaches for generating realistic virtual patient populations (**Section I**) and further explored how mechanism-based M&S approaches can be used to understand physiological or microbiological sources of variability in PK and PD to support the optimization of antibiotic treatment strategies (**Section II**).

7.1 Virtual populations in quantitative pharmacology

In the context of this thesis, virtual populations can be viewed as a set of individual patient-associated covariates, i.e., which aim to realistically reflect the distribution and inter-dependency of covariates in a virtual cohort using multivariate statistical models such as copulas (**Chapter 2-4**). These virtual populations capture inter-individual variability and the correlation between covariates or parameters, and are used in (population) pharmacokinetic/pharmacodynamic (PK/PD) and physiologically based PK (PBPK) models to account for differences in drug exposure and treatment response [1, 2]. In addition, mechanism-based models can be used to study how variability in microbiological or physiological patient differences or in the context of M&S model parameters, may translate into inter-individual variability in PK/PD outcomes, thereby also reflecting a virtual patient population (**Chapter 5,6**).

Copula modelling and its challenges

Virtual populations capturing distributional characteristics of patient covariates are increasingly used in quantitative pharmacology. Copula models recently emerged as a suitable method for generating realistic virtual populations [1]. As such, a need arose to develop a standardized virtual population evaluation framework, to streamline workflows, application, evaluation, and to critically assess potential limitations. To this end, in **Section I**, we introduced the theory and implementation of copula models for virtual population generation and evaluated their performance in generating adult covariate data and drug metabolizing enzyme and transporter expression profiles.

In **Chapter 2**, we describe a practical guidance to copula modelling with a newly developed model-independent validation framework to assess whether virtual populations truly reflect the real-world population, aiming to guide the application of copula-based simulations in the field. To this end, we first covered the theoretical basis of copula models and developed a standard workflow for the development, evaluation and application of these models for covariate simulations. In addition, we also discuss and address the lack of validation standards in the field. We developed a simulation-

based strategy to evaluate virtual populations by comparing the marginal distribution and dependency structures between the observed and simulated population. The developed workflow includes both visual and quantitative evaluation approaches based on density curves and contours, marginal and dependency metrics, and the visual predictive checks, including the donutVPC. The donutVPC was newly developed by extending the concept of the VPC already commonly used in pharmacometrics, towards density contour evaluation. These tools were made available to the community in the R package *pmxcopula*. As the package supports the comparison of population distributions, it can be used to a wide range of virtual population generation methods, including but not limited to copulas. Finally, we discussed and addressed some important challenges around the application of copulas, which involved the application to mixed data type or small-scale datasets.

Building on the developed workflows, in **Chapter 3**, we developed a copula-based virtual population using NHANES adult characteristics and evaluated its similarity to the observed data. The developed adult population copula model was made available through a newly developed web application (<https://cocosim.lacdr.leidenuniv.nl/>). As such, this chapter also demonstrates a concrete practical application of copulas for virtual population generation, from copula model development and evaluation to virtual population generation and data sharing. The NHANES dataset includes both continuous covariates (e.g., age and body weight), binary covariates (i.e., sex) and nominal covariates (unordered categorical; i.e., race-ethnicity). This mixed dataset presents a typical challenge in copula development, as continuous and discrete variables differ fundamentally in how probabilities and dependence are defined. Although modelling binary and ordinal variables with copulas is supported by existing approaches [3, 4, 5, 6, 7], studies that explicitly characterize the incorporation of nominal variables are relatively limited. In practice, nominal variables are typically incorporated into copulas either by converting them into ordinal variables (i.e., ordinal encoding) or by representing each category with a set of binary indicators (i.e., dummy coding) [8, 9]. In Chapter 3, we applied ordinal encoding for race-ethnicity using an exhaustive search approach to identify an ordering within race-ethnicity (i.e. Hispanic, white, African American, Asian and other race) that minimized the object function value. This approach was computationally expensive, suggesting that further methodological research and evaluation of alternative more efficient algorithms for nominal variables would be instrumental.

In **Chapter 4**, we focused on the modelling of distributional characteristics of drug metabolizing enzymes and transporters (DMETs) across individuals for virtual patient generation. We compared the performance of the multivariable normal distribution (MVN), the most widely used joint distribution model, with two alternative methodologies: an MVN model with transformation of the marginal distribution (transMVN) and copulas. The transMVN methodology was proposed to balance the computational simplicity of MVN with the modelling flexibility of copulas. We found that the hepatic DMET gene expression profiles from the GTEx database significantly deviated from those captured by a MVN model. In contrast, transMVN and copula models more accurately represented DMET expression profiles. The subsequent impact for

misspecification of the joint distribution model was examined through a PBPK case study of selegiline, which is mainly metabolized by two CYP enzymes. The median clearance and AUC predicted with transMVN and copula model simulated CYP data were found to be closer to observed data predictions than those of the MVN model, underscoring the limitation in the current routine use of the MVN model, and highlighting the importance of flexible alternatives such as transMVN and copula models to enhance patient population-level predictions with PBPK or QSP models.

A common consideration in M&S is balancing model flexibility and estimation feasibility given limited sample sizes. Improved model flexibility comes at the cost of an increased number of parameters, which inflates estimation burden and the risk of overfitting. The fundamental role of MVN in data analysis is driven by its simplicity and data-efficient estimation. This efficiency is a direct result of its strong assumptions and lean parameterization, defined solely by the mean and covariance matrix. Compared with MVN, transMVN and copula offer greater flexibility in capturing the variability and dependence structure, at the expense of a heavier parameter estimation burden. Statistical MVN tests can justify whether the null hypothesis of MVN distributed data can be rejected. When data are limited, the reduced power to reject MVN favours simpler joint distributions, which helps prevent overfitting in data-limited settings. Based on these considerations, and to help guide the selection of appropriate distributional methods, Chapter 4 proposed a decision tree to recommend the optimal joint distribution model among the MVN, transMVN and copula models, streamlining model selection while accounting for dependency structure and sample sizes. For the DMET datasets, the models recommended by decision tree aligned with those identified through a systematic performance comparison.

Time-varying copulas

Future advancements in copula-based virtual populations should focus on incorporating dependencies that vary along additional dimensions, such as time. The virtual populations considered in this thesis were constructed under the assumption that for each individual, each covariate is represented by a single measurement. This implies that within-patient heterogeneity may not be captured, such as variability due to changes over time or differences across anatomical locations of covariates of interest. For scenarios in which a single value per patient is an adequate representation, e.g., when covariates remain stable over the treatment window and only one anatomical site is relevant, this assumption is reasonable. However, treatment response may depend on time-varying covariates, such as postnatal age and renal function in neonates, or on spatially heterogeneous physiological parameters. In such cases, extending the virtual population framework to incorporate temporal and spatial information could be beneficial for model-based treatment optimization [10, 11]. Methodologically, copula-based approaches can also be extended to settings with temporal or spatial structure, for example, through copula autoregressive models [12], spatio-temporal vine copulas [13, 14], or build copulas based on parameters inferred from separate mixed-effects models of the covariates [1]. These copula-based approaches enable the modelling of more complex, dynamic or spatially heteroge-

neous dependencies. In related work not included in this thesis, we have explored the use of copula-based approaches to capture time-varying dependencies between the common covariates in neonates, including gestational age, birth weight, postnatal age, and serum creatinine. Preliminary results indicated that copula models built on a mechanism-based mixed-effects model generated more realistic paediatric virtual populations than approaches based on pooled observations or spline-based mixed-effects copula models.

Machine learning

The use of machine learning (ML) approaches and copulas have certain overlap in e.g., the goal of modelling multivariate data including dependence characterization [15]. In the realm of ML, some unsupervised ML algorithms, such as generative models, can be used to model multivariate distributions and generate virtual populations [16, 17]. MVN assumptions are also involved in several ML algorithms for efficient algorithmic implementations, such as linear regression and its regularized variants (e.g., Ridge and Lasso), Gaussian mixture models, and Gaussian-based Bayesian inference [18, 19].

Copulas capture the joint distribution of observed variables by characterizing dependence structure separately from the marginal distributions. Marginal distributions, or margins, are informative only about how the data were collected or recorded. As a result, any monotonic transformation of the data only affects the margins but not the estimated copula. This margin-invariant characterization of multivariate dependence has implications for ML approaches, including MVN-based ML algorithms. Therefore, the use of copulas in the improvement of ML methods is an active research area [20]. For example, copulas can be used as plug-in modules in ML frameworks to increase the flexibility, as in copula-based Bayesian networks [21] and copula-based neural networks [22]; to better characterize the dependence, as in copula based classifier [23]; and to improve computational efficiency, for instance by combining vine copula with autoencoders in generative models [24]. Conversely, ML methods are also used to advance copulas. For example, neural networks can support efficient estimation of marginal distributions and high dimensional copula structures [25, 26].

Copulas and ML methods are grounded in two distinct application domains: a statistical domain aimed for inference and an ML domain focused on predictions. The different objective leads to different model development approaches. Copulas use all data to estimate the model parameters, aiming to accurately capture the joint distribution of variables. In contrast, conventional supervised ML models rely on data partitioning (e.g., training and test sets) to assess the predictive performance and generalization ability. However, if copulas are used inside an ML workflow, for example, to augment the available dataset for developing an ML model, the training and test datasets are defined for the downstream task [27].

Mechanistic approaches to generate virtual populations

Real-world patient characteristics may often not fully address ultimately observed variability in PK or PD responses. The use of mechanistic models (e.g., PBPK or QSP models) may address some of those gaps to understand what physiological, or in this

thesis microbial, factors may contribute to observed inter-individual variability, i.e., through backward calibration of observed treatment outcome data using mechanistic model [19, 20, 21].

A standard procedure involves first defining a parameter space with realistic ranges and then simulating all parameter combinations. Parameter sets that yield unrealistic outcomes are discarded, while those producing plausible treatment responses are retained as virtual patients. This two-stage framework forms the basis for model-calibrated virtual population generation, and recent advances have been focused on improving computational efficiency using ML surrogate models [28, 29, 30, 31]. In the context of such approaches, biological parameters in mechanistic models can be interdependent, yet, this interdependence is often neglected in virtual population generation using such models because the full prior correlation structure is often unknown [32]. Even without complete knowledge of the correlation structure, partial correlation information can still be inferred empirically from published clinical trials or external databases. Methods like MVN, transMVN and copulas, such as developed in Chapter 4, can describe this information to exclude unrealistic parameter combinations, facilitating faster virtual population generation and improved QSP predictions. Among these methods, copulas offer an advantage in capturing tail dependencies, where extreme values occur simultaneously. Correct characterization of tail dependencies can enhance the prediction for rare but critical scenarios, for example, the extremely high cytokine levels during a cytokine storm. Ignoring tail dependence can result in serious consequences. Mechanistic virtual population generation complemented with copulas are thus of interest for future research.

7.2 Model-informed antibiotic treatment optimization

M&S approaches can be used in the context of drug development as well as further treatment optimization. Antibiotic therapy is increasingly compromised by rapid antimicrobial resistance, which, in combination with the low discovery rate of novel antibiotic classes [33], has forced researchers to focus on optimizing treatment strategies of existing antibiotics. In this context, Section II focused on the application of M&S-based strategies for antibiotic treatment optimization in bacterial lung infections and characterization of factors driving potential variability in PK or PD.

Antimicrobial pharmacodynamics

Antibiotic resistance is closely linked to bacterial lifestyle. Compared with planktonic bacteria, biofilm-associated bacteria typically exhibit greater resistance to antibiotics, yet have been studied much less. In **Chapter 5**, we characterized PK/PD relationships for both colistin and imipenem in biofilm-associated *Pseudomonas aeruginosa* infections using time-kill assay data. The developed models captured reduced antibiotic diffusion, slower growth rates, and decreased sensitivity of resistant subpopulations in biofilm bacteria. We also observed differences in the treatment outcome between the two bacterial lifestyles under humanized drug concentration profiles. Although these simulations omit factors such as host immune responses, they high-

light the impact of bacterial lifestyles under clinically relevant conditions and provide insight into how microbial properties and associated model parameters may affect observed PD response.

Chapter 5 utilizes *in vitro* time-kill assay data that are affected by known experimental factors, such as extensive binding of colistin to polystyrene microplates and degradation of imipenem under the assay conditions [34, 35]. Conventional PK/PD models based on time-kill assay treat antibiotic concentration as static. In our case, this assumption can distort PK/PD relationships. Bacterial regrowth may result from drug loss rather than resistance development. Although *in vitro* PK was not measured during the time-kill assay, drug binding and degradation were addressed by incorporating the *in vitro* PK information from previous studies. This approach allows accurate characterization of PK/PD relationships and should be applied when drug concentrations change substantially over time. This extends beyond microbial assays to any *in vitro* experiments that measure drug effects over time. In addition to colistin and imipenem, other drugs are also subject to binding- and degradation-related losses. For example, mecillinam has been reported to exhibit a 2-hour half-life in MOPS medium at 37°C and pH 7.4 [36]; doxycycline and minocycline showed significant binding to polypropylene test tubes [37].

In Chapter 5, we show bacterial lifestyle strongly influences antibiotic responses, highlighting the need for bacteria lifestyle-guided treatment strategies. However, in routine clinical practice, antibiotic regimens are mainly optimized for planktonic bacteria. In clinical microbiology, antibiotic susceptibility testing (AST) is routinely performed under planktonic conditions, and its key output, the minimal inhibitory concentration (MIC), is inherently designed for this lifestyle. MIC-based PK/PD indices, such as the ratio of maximal free concentration to MIC (fC_{max}/MIC), are widely used to guide antibiotic dosing by defining efficacy targets that can be achieved through dose escalation, prolonged or continuous infusions and loading doses [38]. In contrast, the management of biofilm-associated infections presents challenges. While the minimal biofilm inhibitory concentration (MBIC) has been proposed [27], clinical adoption remains low, and there is a lack of evidence to support the use of biofilm AST to guide treatment of *Pseudomonas aeruginosa* pulmonary infections in CF patients [39, 40, 41]. This reflects a limited correlation between *in vitro* biofilm susceptibility and clinical outcomes. Although in Chapter 5 we derived PK/PD targets for both planktonic and biofilm-associated infections, the application of biofilm-specific PK/PD targets remain confined to translational research [42, 43, 44]. Current clinical guidelines for biofilm associated infections (e.g., chronic lung infection and device related infections) emphasize longer duration, combination therapy, local/high concentration delivery, and source control, but does not specify PK/PD targets tailored to biofilm lifestyle for routine dosing decisions [45].

Spatial drug distribution

Ensuring attainment of sufficient antibiotic exposure at the site of infection is another prerequisite for bacterial eradication and prevention of resistance. This may be of particular relevance in patients with cystic fibrosis (CF), who typically suffer from chronic

lung infections due to the presence of a thick mucus layer in the respiratory tract, where treatment including antibiotics may often be associated with reduced efficacy. To determine whether the mucus layer in CF patients may act as a diffusion barrier for drug delivery, in **Chapter 6**, we developed a spatial PK/PD modelling framework to predict the temporal and spatial kinetics of drug concentration and its effect. We found that under intravenous PK profiles, molecule size plays an important role in drug movement within mucus. Small molecules (radius < 1nm) could freely penetrate mucus. In contrast, large molecules/particles showed impaired diffusion. In addition, drug-mucin interactions and muco-ciliary clearance showed a minor impact, in contrast to prior reports. To demonstrate how such spatial PK models may be coupled to PD models, a spatial PK/PD model was developed based on the imipenem PK/PD model developed in Chapter 5, which revealed that baseline bacterial organization influenced eradication outcomes. Overall, this study adds quantitative and mechanistic understanding of drug diffusion in airway mucus, and highlights how consideration of spatial heterogeneity may be critical in understanding PK/PD responses in some conditions. Moreover, methodologically, the current study represents one of the few M&S examples that have considered spatial PK models.

Covariates and key parameters selection enables population variability

Virtual populations enable the prediction of treatment exposure and outcome variability by incorporating variability in covariates or key parameters. The methodology for identifying which covariates or parameters drive this variability depends on the underlying models. In top-down models, such as the population PK model used in Chapter 5, covariates like body weight and GFR are typically identified based on the statistical significance in improving overall model fit together with biological plausibility. These identified covariates form the basis for virtual population that supported the prediction of population-level bacteria response. In contrast, bottom-up models often involve a large number of parameters and face challenges of parameter identifiability. As a result, sensitivity analysis is commonly applied to prioritize parameters that strongly influence model outputs. Then, variability is incorporated into virtual population for these influential parameters. The screening process was also used in the physiologically based spatial PK model developed in Chapter 6, where sensitivity analysis revealed that variation in the physiological parameters (i.e., mucin concentration and mucociliary clearance) had minimal impact on the drug diffusion, indicating limited relevance for predicting population-level PK and PD variability.

7.3 Open science

Data sharing

The future development of the field of M&S in quantitative pharmacology depends on not only the advancement of quantitative approaches but also equitable accessibility of high-quality patient data. While the number of published models has grown

exponentially, the underlying patient data are rarely made public due to stringent privacy policies. When individual level data can be shared, anonymization and de-identification techniques, such as age banding and k-anonymity, can help protect patient privacy by masking or removing key identifiers [46]. When such data cannot be shared, summary metrics or models derived from the raw data are often shared and used to generate synthetic datasets. In most quantitative pharmacology M&S studies, patient characteristics are typically reported as mean, median and range. This approach ignores the correlation between patient characteristics and therefore may not accurately represent the observed data. Copula models offer greater flexibility in capturing both variability and dependence structures, enabling to generate synthetic datasets that better reflect the observed data without distorting the underlying information [47]. In addition, ML approaches, such as regression models, Bayesian networks and variational autoencoders, can also support data sharing [48, 49]. These methods not only offer a viable starting point for researchers without direct data access, but also help circumvent the delay caused by lengthy data-sharing approvals, as synthetic data can be used as a robust proxy for preliminary modelling and help streamline the study workflow.

The robustness and applicability of copulas or other models depend on the quality and representativeness of the input data. Single-center datasets may lack the heterogeneity in real-world populations, so reliable copulas generally benefit from large and diverse datasets that capture real-world patient populations. This demands a joint effort from healthcare institutions and industry partners to pool resources to better reflect the populations encountered in clinical practice. By transforming clinical data into the shared resource, such initiatives support model-informed drug development and precision dosing, and ultimately enable improved real-world care.

Knowledge transfer

Quantitative pharmacological M&S approaches operate at the intersection of mathematics, statistics, biology and pharmacology. Because of this unique positioning, the field inevitably advances through knowledge transfer, often by adapting methodologies from other domains to solve complex pharmacological problems. This thesis illustrates such cross-disciplinary evolution through the application of copulas as a methodology for virtual population generation. While copulas, as a statistical method, are widely accepted and applied in finance, insurance, hydrology and climate, their integration into quantitative pharmacology is an emerging and accelerating trend [1, 50, 51, 52]. Similarly, Chapter 6 employed the use of partial differential equations, a core mathematical framework in fluid mechanics in physics, yet relatively novel for modelling pharmacological data.

Knowledge transfer across disciplines may be hampered by differences in terminology, underlying assumptions and conventions. These challenges occur not only in collaborations between researchers from different backgrounds, but also at the individual level when, for example, a pharmacometrician seeks to adopt statistical methodologies. For this reason, Chapter 2 presents a tutorial on copula-based virtual population generation, covering both theory and practical coding guidance. The

same concepts may be expressed using different “languages”. Meaningful exchange therefore requires careful clarification of definitions and assumptions before productive discussion can begin. This challenge is particularly evident in collaborations between experimentalists and modellers, where progress depends on shared understanding of experimental design, data-generating assumptions, and the abstraction in modelling [53].

Recognizing the necessity of knowledge transfer is only the first step. Effective transfer can be facilitated by sharing well-documented and reproducible code, including clear records of software versions and even computational environments. At the community level, additional educational initiatives, such as conference keynote lectures and journal tutorials that provide step-by-step guidance on cross-disciplinary workflows, are essential. Ultimately, the most important bridge between these worlds is the interdisciplinary researchers who can translate, integrate and synthesize ideas across fields.

Tools

Beyond methodological development itself, future impact of M&S approaches also depends on the accessibility of the resulting tools. Openly available tools, supported by intuitive interfaces, transparent workflows and well-chosen examples could lower the barrier to adoption and help convey intuition effectively than abstract descriptions alone. Furthermore, stable maintenance and user feedback-driven issue resolution are essential for sustained application of the tool. At the same time, user friendliness as a key to ensure uptake by stakeholders and thus enhance the impact of open science. This thesis provides tools tailored for both technical modeller and non-technical users. Chapter 2 introduced the *pmxcopula* R package, an open-source tool for assessing whether virtual population is representative of the real-world population. Chapter 3 presented a user-friendly web application that allows non-modellers to generate realistic virtual population using the NHANES copula. These efforts align with broader open-source initiatives within quantitative pharmacology, such as the open systems pharmacology community, which aims to provide robust and accessible M&S tools for life-sciences [54].

7.4 Conclusions

In conclusion, this thesis advanced and demonstrated the value of novel approaches to generate realistic virtual populations, as well as the use of quantitative mechanism-based M&S strategies to further characterize sources of variability in the context of antibiotic treatment optimization. As such, the approaches developed in this thesis may serve as a framework that can be expanded towards other patient populations, therapeutic indications and modalities in drug development and patient care.

References

1. Zwep LB, Guo T, Nagler T, Knibbe CAJ, Meulman JJ, and Hasselt JGC van. Virtual Patient Simulation Using Copula Modeling. *Clinical Pharmacology and Therapeutics*. 2024 Apr; 115:795–804
2. Allen R, Rieger T, and Musante C. Efficient Generation and Selection of Virtual Populations in Quantitative Systems Pharmacology Models. *CPT: Pharmacometrics & Systems Pharmacology*. 2016; 5:140–6
3. Nikoloulopoulos AK. Copula-Based Models for Multivariate Discrete Response Data. *Copulae in Mathematical and Quantitative Finance*. Ed. by Jaworski P, Durante F, and Härdle WK. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013 :231–49
4. Mitskopoulos L, Amvrosiadis T, and Onken A. Mixed vine copula flows for flexible modeling of neural dependencies. *Frontiers in Neuroscience*. 2022 Sep; 16:910122
5. Lin L, Wang R, Zhang R, and Zhao C. The Checkerboard Copula and Dependence Concepts. *SIAM Journal on Financial Mathematics*. 2025; 16:426–47
6. Molendijk T. Copula-based regression for discrete data. phd. TU Delft, 2025
7. Leon AR de and Wu B. Copula-based regression models for a bivariate mixed discrete and continuous outcome. *Statistics in Medicine*. 2011; 30:175–85
8. Feldman J and Kowal DR. Nonparametric Copula Models for Multivariate, Mixed, and Missing Data. *Journal of Machine Learning Research*. 2024; 25:1–50
9. Pan S. Conditional inferences and predictions based on copula models. phd. University of British Columbia, 2023
10. Abduljalil K, Jamei M, Rostami-Hodjegan A, and Johnson TN. Changes in Individual Drug-Independent System Parameters during Virtual Paediatric Pharmacokinetic Trials: Introducing Time-Varying Physiology into a Paediatric PBPK Model. *The AAPS Journal*. 2014 May; 16:568–76
11. Zhang S et al. Integrating single cell sequencing with a spatial quantitative systems pharmacology model spQSP for personalized prediction of triple-negative breast cancer immunotherapy response. *Immunoinformatics (Amsterdam, Netherlands)*. 2021 Oct; 1-2:100002
12. Brechmann EC and Czado C. COPAR—multivariate time series modeling using the copula autoregressive model. *Applied Stochastic Models in Business and Industry*. 2015 Jul; 31:495–514
13. Gräler B. spcopula: Modelling spatial and spatio-temporal dependence with copulas in R. 2013
14. Ausin MC and Lopes HF. Time-varying joint distribution through copulas. *Computational Statistics & Data Analysis. The Fifth Special Issue on Computational Econometrics* 2010 Nov; 54:2383–99
15. Elidan G. Copulas in machine learning. *Copulae in Mathematical and Quantitative Finance: Proceedings of the Workshop Held in Cracow, 10-11 July 2012*. Springer, 2013 :39–60
16. Janke T, Ghanmi M, and Steinke F. Implicit Generative Copulas. *Advances in Neural Information Processing Systems*. Vol. 34. Curran Associates, Inc., 2021 :26028–39
17. Nair R, Mohan DD, Frank S, Setlur S, Govindaraju V, and Ramanathan M. Generative adversarial networks for modelling clinical biomarker profiles with race/ethnicity. *British Journal of Clinical Pharmacology*. 2023; 89:1588–600
18. Ogutu JO, Schulz-Streeck T, and Piepho HP. Genomic selection using regularized linear regression models: ridge regression, lasso, elastic net and their extensions. *BMC Proceedings*. 2012 May; 6:S10
19. Murphy KP. Conjugate Bayesian analysis of the Gaussian distribution. *DEF : Development education forum*. 2007; 1:16
20. Größer J and Okhrin O. Copulae: An overview and recent developments. *Wiley Interdisciplinary Reviews: Computational Statistics*. 2022; 14:e1557
21. Pircalabelu E, Claeskens G, and Gijbels I. Copula directed acyclic graphs. *Statistics and Computing*. 2017; 27:55–78
22. Simchoni G and Rosset S. Flexible Copula-Based Mixed Models in Deep Learning: A Scalable Approach to Arbitrary Marginals. *International Conference on Artificial Intelligence and Statistics*. PMLR, 2025 :91–9

23. Ozdemir O, Allen TG, Choi S, Wimalajeewa T, and Varshney PK. Copula based classifier fusion under statistical dependence. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2017; 40:2740–8
24. Tagasovska N, Ackerer D, and Vatter T. Copulas as High-Dimensional Generative Models: Vine Copula Autoencoders. *Advances in Neural Information Processing Systems*. Vol. 32. Curran Associates, Inc., 2019
25. Chen CH, Song F, Hwang FJ, and Wu L. A probability density function generator based on neural networks. *Physica A: Statistical Mechanics and its Applications*. 2020; 541:123344
26. Zeng Z and Wang T. Neural Copula: A unified framework for estimating generic high-dimensional Copula functions. *arXiv preprint arXiv:2205.15031*. 2022
27. Meyer D, Nagler T, and Hogan RJ. Copula-based synthetic data augmentation for machine-learning emulators. *Geoscientific Model Development*. 2021 Aug; 14:5205–15
28. Myers RC, Augustin F, Huard J, and Friedrich CM. Using machine learning surrogate modeling for faster QSP VP cohort generation. *CPT: Pharmacometrics & Systems Pharmacology*. 2023; 12:1047–59
29. Matteson A, Hagen D, Swat MH, Matthews R, Kierzek AM, and Apgar JF. A Scalable Cloud-Based QSP Modeling System for Virtual Patient Generation. *American Conference of Pharmacometrics*. International Society of Pharmacometrics, 2024
30. Pawłowski T, Bokota G, Lazarou G, Kierzek AM, and Sroka J. Emulation of Quantitative Systems Pharmacology models to accelerate virtual population inference in immuno-oncology. *Methods (San Diego, Calif.)* 2024 Mar; 223:118–26
31. Androulakis IP et al. The dawn of a new era: can machine learning and large language models reshape QSP modeling? *Journal of Pharmacokinetics and Pharmacodynamics*. 2025; 52:36
32. Myers RC, Augustin F, Huard J, and Friedrich CM. Using machine learning surrogate modeling for faster QSP VP cohort generation. *CPT: Pharmacometrics & Systems Pharmacology*. 2023; 12:1047–59
33. Aminov RI. A Brief History of the Antibiotic Era: Lessons Learned and Challenges for the Future. *Frontiers in Microbiology*. 2010 Dec; 1
34. Karvanen M, Malmberg C, Lagerbäck P, Friberg LE, and Cars O. Colistin is extensively lost during standard in vitro experimental conditions. *Antimicrobial Agents and Chemotherapy*. 2017
35. Swanson DJ, DeAngelis C, Smith IL, and Schentag JJ. Degradation kinetics of imipenem in normal saline and in human serum. *Antimicrobial Agents and Chemotherapy*. 1986 May; 29:936–7
36. Brouwers R, Vass H, Dawson A, Squires T, Tavaddod S, and Allen RJ. Stability of β -lactam antibiotics in bacterial growth media. *PLoS ONE*. 2020 Jul; 15:e0236198
37. Ciarlone AE, Fry BW, and Ziemer DM. Some observations on the adsorption of tetracyclines to glass and plastic labware. *Microchemical Journal*. 1990 Oct; 42:250–5
38. Tilanus A and Drusano G. Optimizing the Use of Beta-Lactam Antibiotics in Clinical Practice: A Test of Time. *Open Forum Infectious Diseases*. 2023 Jul; 10:ofad305
39. Drevinek P et al. New concepts in antimicrobial resistance in cystic fibrosis respiratory infections. *Journal of Cystic Fibrosis*. 2022 Nov; 21:937–45
40. Moskowitz SM et al. Randomized trial of biofilm testing to select antibiotics for cystic fibrosis airway infection. *Pediatric Pulmonology*. 2011 Feb; 46:184–92
41. Waters V and Ratjen F. Standard versus biofilm antimicrobial susceptibility testing to guide antibiotic therapy in cystic fibrosis. *The Cochrane Database of Systematic Reviews*. 2012 Nov; 11:CD009528
42. Hengzhuang W, Høiby N, and Ciofu O. Pharmacokinetics and pharmacodynamics of antibiotics in biofilm infections of *Pseudomonas aeruginosa* in vitro and in vivo. *Methods in Molecular Biology*. 2014
43. Hengzhuang W, Wu H, Ciofu O, Song Z, and Høiby N. Pharmacokinetics/pharmacodynamics of colistin and imipenem on mucoid and nonmucoid *Pseudomonas aeruginosa* biofilms. *Antimicrobial agents and chemotherapy*. 2011 Sep; 55:4469–74
44. Hengzhuang W, Wu H, Ciofu O, Song Z, and Høiby N. In vivo pharmacokinetics/pharmacodynamics of colistin and imipenem in *Pseudomonas aeruginosa* biofilm infection. *Antimicrobial agents and chemotherapy*. 2012 May; 56:2683–90

45. Høiby N et al. ESCMID guideline for the diagnosis and treatment of biofilm infections 2014. *Clinical Microbiology and Infection: The Official Publication of the European Society of Clinical Microbiology and Infectious Diseases*. 2015 May; 21 Suppl 1:S1–25
46. LIU X, LI XB, MOTIWALLA L, LI W, ZHENG H, and FRANKLIN PD. Preserving Patient Privacy When Sharing Same-Disease Data. *ACM journal of data and information quality*. 2016 Oct; 7:17
47. Houssou R, Augustin MC, Rappos E, Bonvin V, and Robert-Nicoud S. Generation and Simulation of Synthetic Datasets with Copulas. 2022 Mar
48. Egmond MB van et al. Privacy-preserving dataset combination and Lasso regression for healthcare predictions. *BMC Medical Informatics and Decision Making*. 2021 Sep; 21:266
49. Sella N, Guinot F, Lagrange N, Albou LP, Desponds J, and Isambert H. Preserving information while respecting privacy through an information theoretic framework for synthetic health data generation. *npj Digital Medicine*. 2025 Jan; 8:49
50. Hartung N and Khatova A. Information-theoretic evaluation of covariate distributions models. *Journal of Pharmacokinetics and Pharmacodynamics*. 2025; 52:21
51. Jutras-Dubé P, Al-Khasawneh MB, Yang Z, Bas J, Bastin F, and Cirillo C. Copula-based transferable models for synthetic population generation. *Transportation Research Part C: Emerging Technologies*. 2024 Dec; 169:104830
52. Fayette L, Brendel K, and Mentré F. Using Fisher Information Matrix to predict uncertainty in covariate effects and power to detect their relevance in Non-Linear Mixed Effect Models in pharmacometrics. *Journal of Pharmacokinetics and Pharmacodynamics*. 2025 Jul; 52:38
53. Gunawardena J. Models in biology: 'accurate descriptions of our pathetic thinking'. *BMC biology*. 2014; 12:29
54. Lippert J et al. Open Systems Pharmacology Community—An Open Access, Open Source, Open Science Approach to Modeling and Simulation in Pharmaceutical Sciences. *CPT: Pharmacometrics & Systems Pharmacology*. 2019 Dec; 8:878–82