Generation of realistic virtual adult populations using the NHANES database: a Copula approach

Yuchen Guo [1], Laura B. Zwep [1], Catherijne A.J. Knibbe [1], Tingjie Guo [1], J. G. Coen van Hasselt [1]

[1] Leiden Academic Center for Drug Research, Leiden University, Leiden, The Netherlands

INTRODUCTION

- Virtual populations (VPs) are commonly used in pharmacometric simulations of clinical trials and optimization of dosing strategies.
- VPs associated covariate variables to be simulated should be representative of the real-world target population.
- We previously demonstrated favorable performance of Copula in comparison to alternative simulation approaches[1].
- A general workflow and a user-friendly platform are needed to guide and support the generation of VPs using Copulas.

Aim

1. To demonstrate a model development workflow for the development of Copula-based virtual patient simulation models
2. To develop a Copula-based model to support the simulation of adult population based on the real-world population in the NHANES[2] database

WORKFLOW
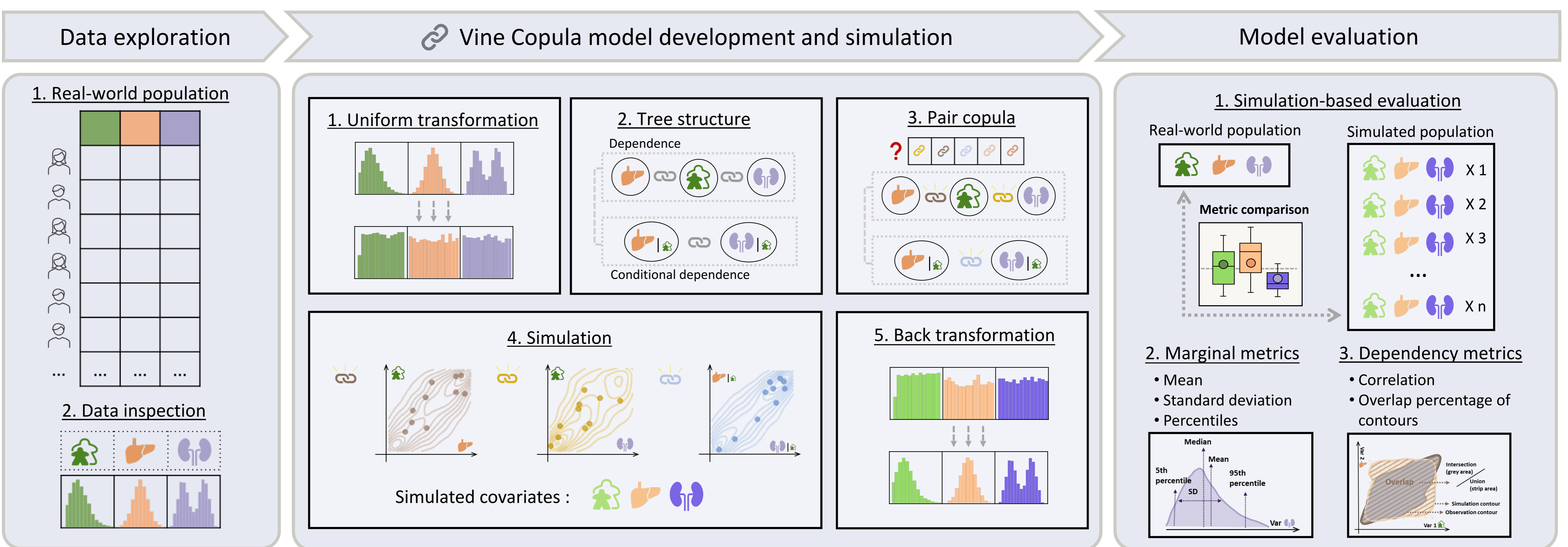
NHANES CASE ANALYSIS

Methods

Data source	NHANES database (2009~2018 releases)
Population	Adults (18~80 years old)
Number of individuals	27008
Covariates	Sex, race, age, height, weight, fat, ALT, AST, ALP, albumin, bilirubin (BR), serum creatinine (SCR)
Candidate models	Vine copulas with parametric pair copula models
Model selection criterion	AIC
Software	R 4.3.0 *rvinecopulib* package *shiny* package

Results

- Marginal metrics of virtual population aligned with observed population (*Fig.1*).
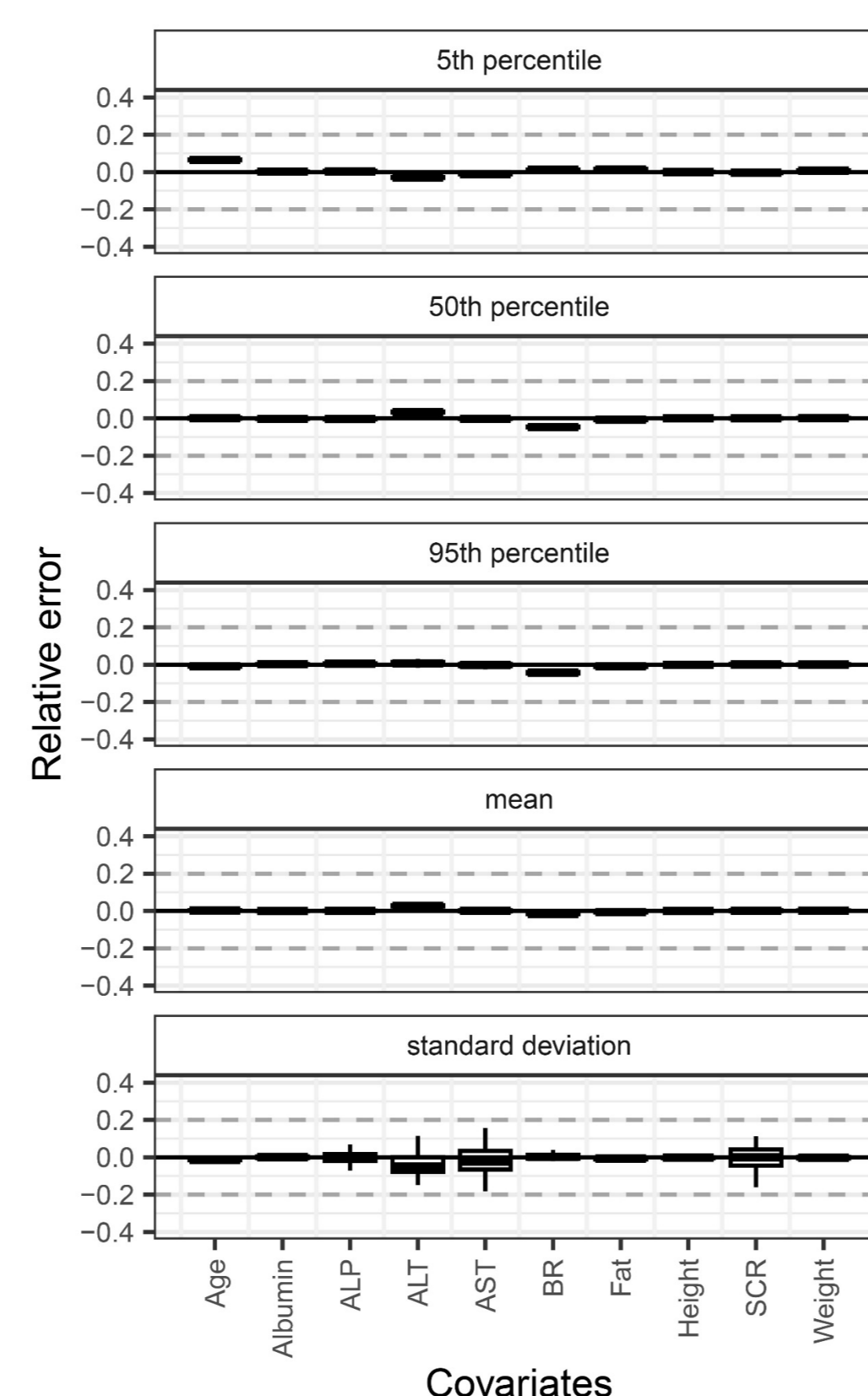

Fig.1. Relative error of marginal metrics over 100 simulations compared to the statistics of the observed population.

- Simulated correlations of each each pair of covariates closely tracked that of observed population (*Fig.2 upper panel*).
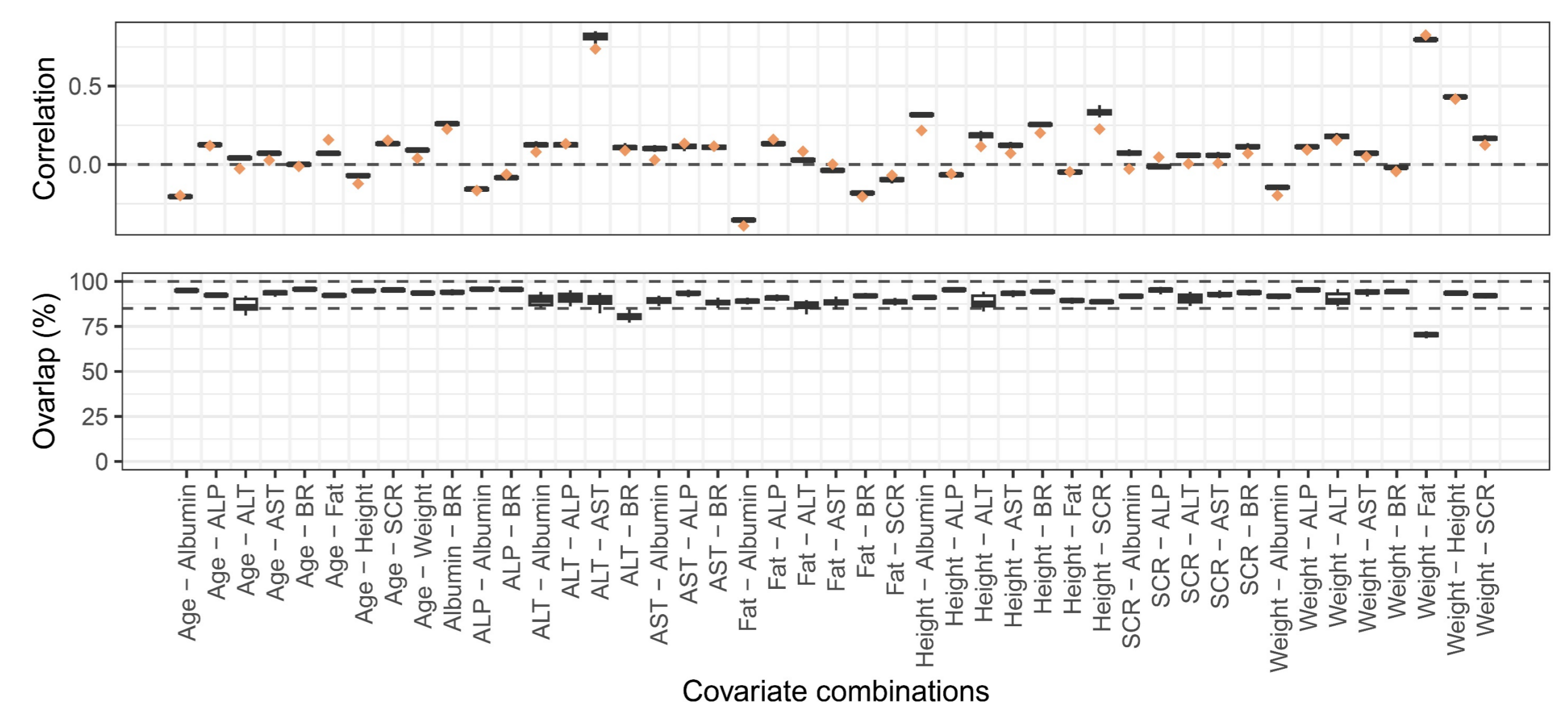- Most covariate pairs achieved over 85% overlap of 95th contour (*Fig.2 lower panel*).

Fig.2. Boxplots of dependency measures of each combination of ten covariates. Upper panel: correlation of simulated(black) and observed population(orange). Lower panel: overlap of 95th contours of each pair covariates.

Generate your virtual population via our shiny app based on NHANES copula!

https://cocosim.lacdr.leidenuniv.nl/
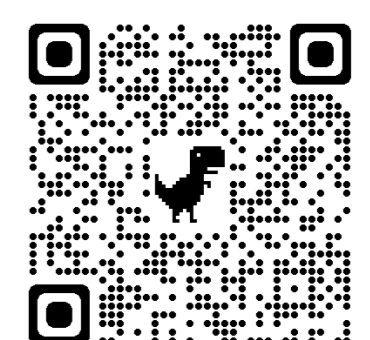
CONCLUSIONS

- We developed and evaluated a copula model for simulation of commonly used covariates in adult population based on the NHANES database, which can be further used to support clinical trial simulation and dose optimization.
- A web application was developed to facilitate the utility of the developed Copula model.

1. Zwep, Laura B., et al. PAGE 30 (2022) Abstr 10099 [www.page-meeting.org/?abstract=10099]
2. Centers for Disease Control and Prevention (CDC). National Center for Health Statistics (NCHS). National Health and Nutrition Examination Survey Data. Hyattsville, MD: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention

y.guo@lacdr.leidenuniv.nl