2nd PANHELLENIC CONGRESS OF MEDICAL PHYSICS
4-6 OCTOBER 2024| EUGENIDES FOUNDATION

Development of a generative artificial intelligence algorithm for the regeneration of virtual volunteers in bioequivalence studies

Anastasios Nikolopoulos[1,2], Vangelis D. Karalis[1,2]

[1]Department of Pharmacy, National and Kapodistrian University of Athens, Athens, Greece

[2]Institute of Applied and Computational Mathematics, Foundation for Research and Technology Hellas (FORTH), Heraklion, Greece

1. Background-Aim

Sample size calculation is vital in clinical trials to ensure reliable assessments of drug products, reducing errors and ensuring ethical research. In **bioequivalence (BE) studies**, determining sample size involves key parameters like formulation differences and statistical power. Advances in AI, particularly through Generative Adversarial Networks (GANs) such as **Wasserstein GANs (WGANs)**, have enhanced data augmentation capabilities. WGANs are e**ffective in generating virtual populations from limited datasets**, aiding in robust trial design and execution.

This study investigates the use of WGANs to augment sample sizes in BE studies, evaluating their *effectiveness in generating virtual subjects that reflect a larger population and comparing their performance against traditional methods*.

2. Materials & Methods

To evaluate the performance of WGANs in BE studies, we followed a detailed methodology that replicates typical study conditions. Our approach involved several key steps: *parameter selection, subject generation through Monte Carlo simulations, tuning of WGAN hyperparameters, and sampling distributions.*

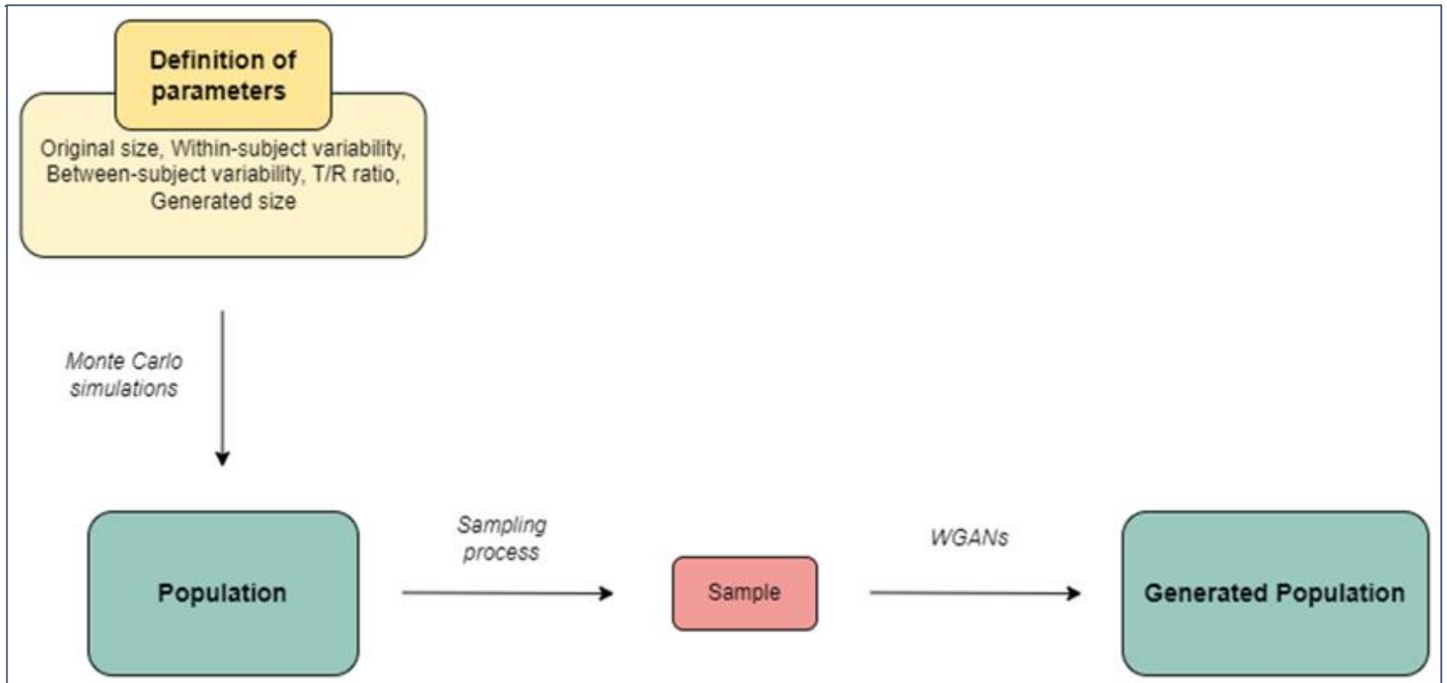

***Figure 1**: Flowchart of the research methodology.*

The process began with generating initial samples based on defined study parameters. These samples were then used as input for the WGANs to create new distributions *(Figure 1)*. *The effectiveness of the generated distributions was assessed by comparing them to the original distributions and evaluating their BE acceptance according to regulatory standards.* **To ensure robustness**, each procedure was repeated hundreds of times with varied parameters, such as variability, T/R ratios, sample percentages, and original population sizes. This repetition aimed to **produce reliable results and validate the performance of the WGANs**.

2. Materials & Methods

The methodology utilized WGANs, which consist of **two main components: a generator and a critic**. The generator produces new data intended to resemble the original data distribution, while the critic evaluates these data for authenticity. **The feedback from the critic helps refine the generator**, enhancing its ability to produce realistic outputs *(Figure 2)*.

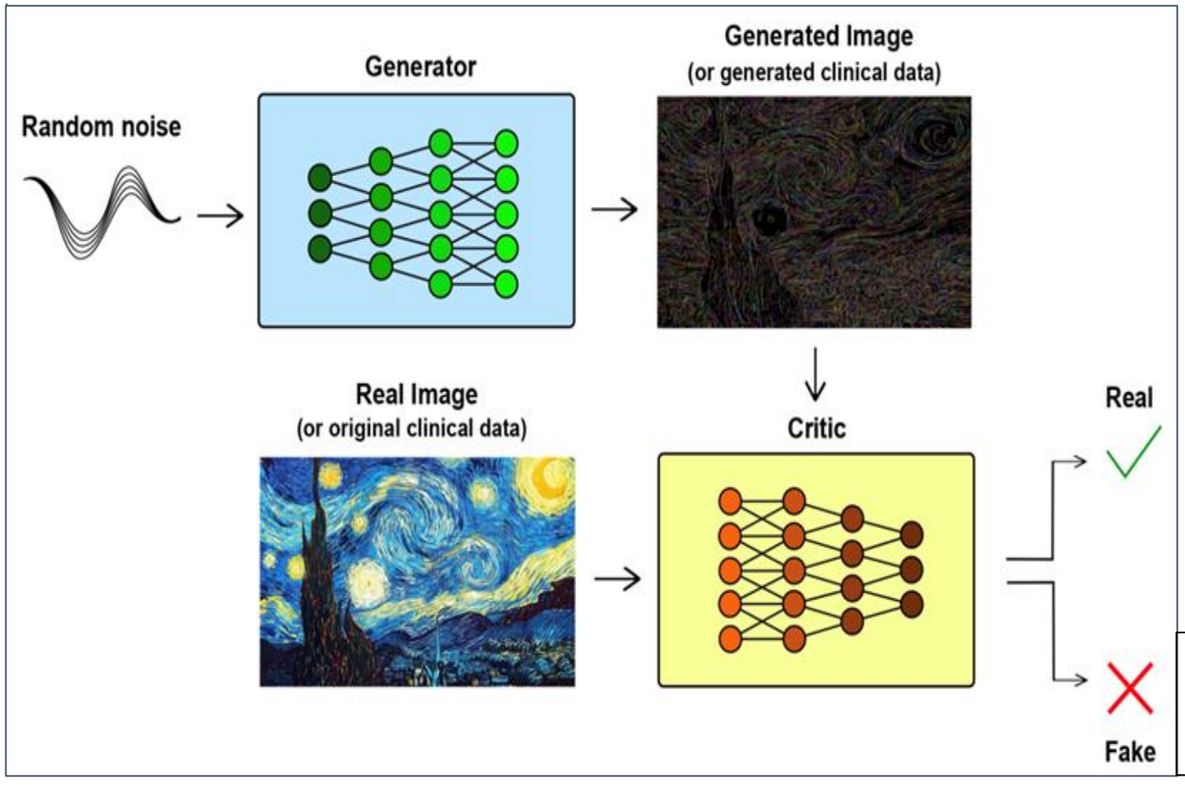

***Figure 2**: Operation of WGANs.*

Hyperparameter tuning was crucial to optimize the WGAN model. We **evaluated various settings**, including epochs, latent dimensions, batch size, activation functions, hidden layers, and learning rates. Monte Carlo simulations were conducted using a 2x2 crossover design. Subjects were randomly assigned to two groups, R and T, across two periods. We generated distributions with specific means and standard deviations, introducing variability to reflect real-world conditions. Samples from these distributions were used as input for the WGANs, and the generated populations were compared to the original and sample distributions. *Statistical analysis, including ANOVA and confidence interval construction*, was performed to **assess BE acceptance**.

3. Results

We assessed the performance of WGANs in BE studies by *comparing the acceptance and similarity percentages of generated populations to original populations*.

Figures 3 reveal that for scenarios with **N = 24, the generated populations effectively mirrored the original in terms of BE acceptance percentages, even with minimal input data** (e.g., as few as six patients). The WGANs generated distributions that consistently met or exceeded the BE acceptance criteria of the original populations. In contrast, traditional sampling methods struggled to replicate these characteristics, especially in high variability and large T/R ratio scenarios.

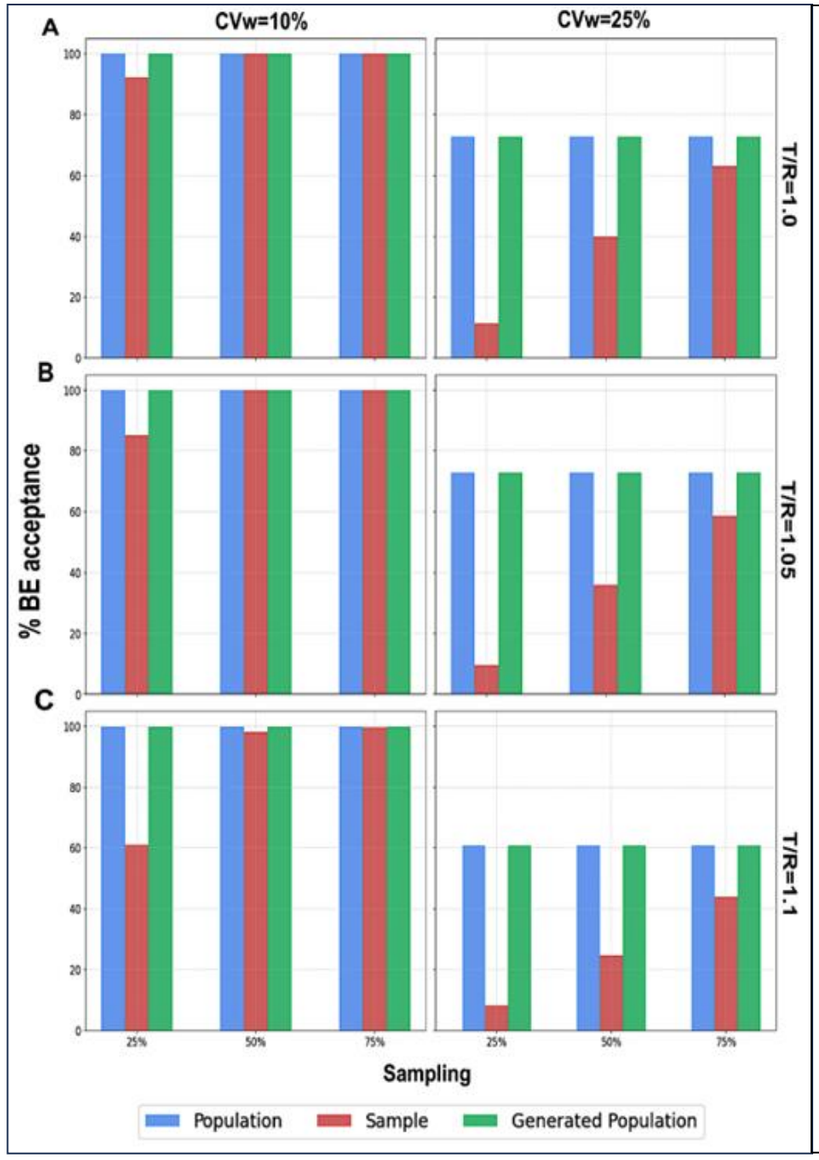

***Figure 3**: Bioequivalence acceptance percentages for the population, sample, and generated subjects at different sampling rates. The population size in all scenarios was 24, with residual variability set at 10% and 25%. Panels (A), (B), and (C) represent T/R ratios of 1.0, 1.05, and 1.1, respectively.*

3. Results

Further analysis demonstrated that WGANs achieved **100% similarity in BE acceptance percentages between generated populations and the original for all tested scenarios, including challenging conditions with small sample sizes and high variability**. *Figure 4* show that the WGANs outperformed traditional samples, maintaining high similarity percentages.

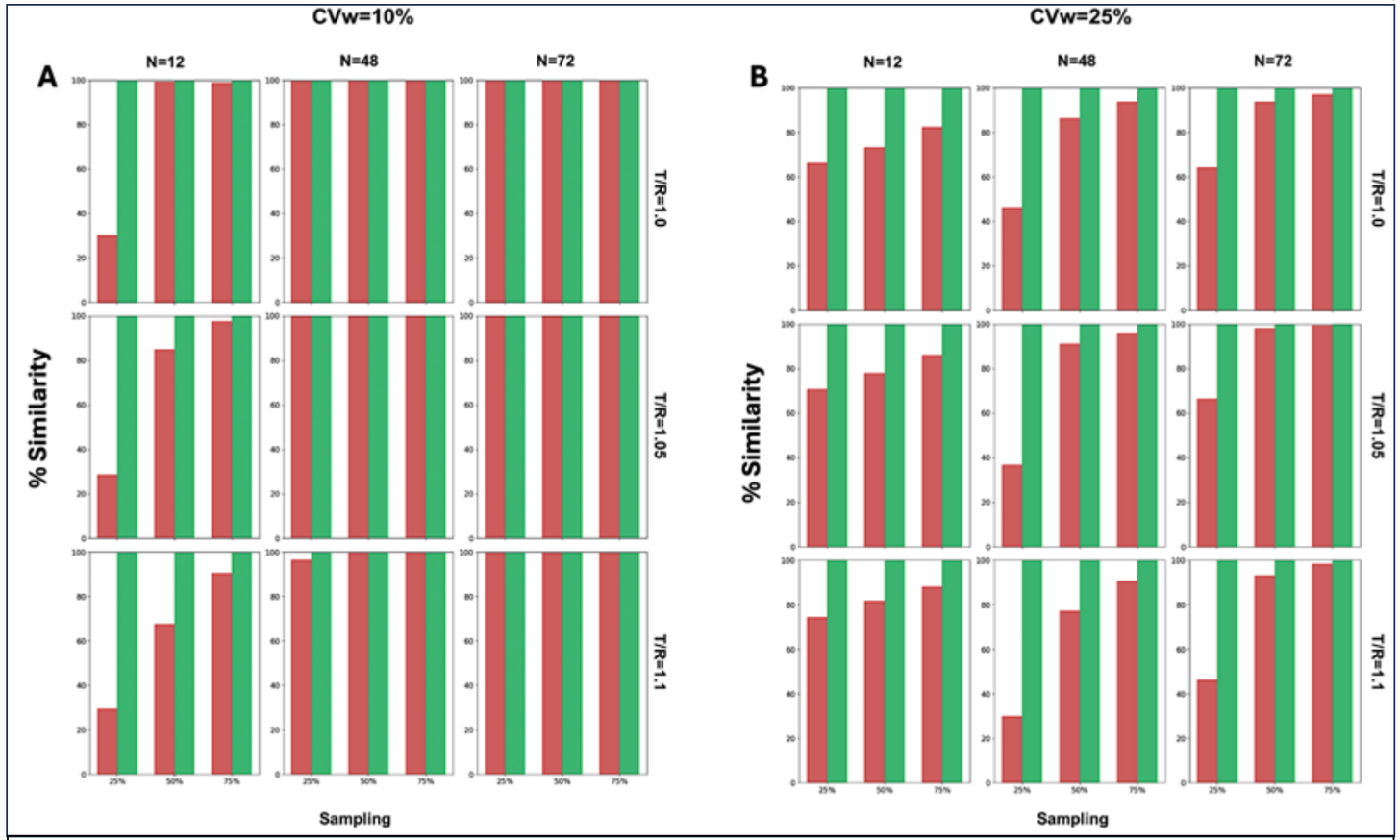

***Figure 4**: Percentage of similarity between the population and the sample (red) and between the population and generated subjects (green) for populations of N = 12, 48, and 72. The analysis covered two levels of residual variability (10% and 25%), with horizontal panels representing T/R ratios of 1.0, 1.05, and 1.1.*

3. Results

For scenarios where the generated population was **twice the size of the original**, **WGANs showed improved BE acceptance percentages, particularly in cases with larger T/R ratios and higher CVw values**. This is highlighted in *Figure 5*, where the WGANs exhibited a positive difference in BE acceptance compared to the original, especially with smaller original sample sizes. The consistent robustness of the WGANs across various clinical study settings, as indicated by constant hyperparameters, underscores their effectiveness and reliability for generating virtual populations.

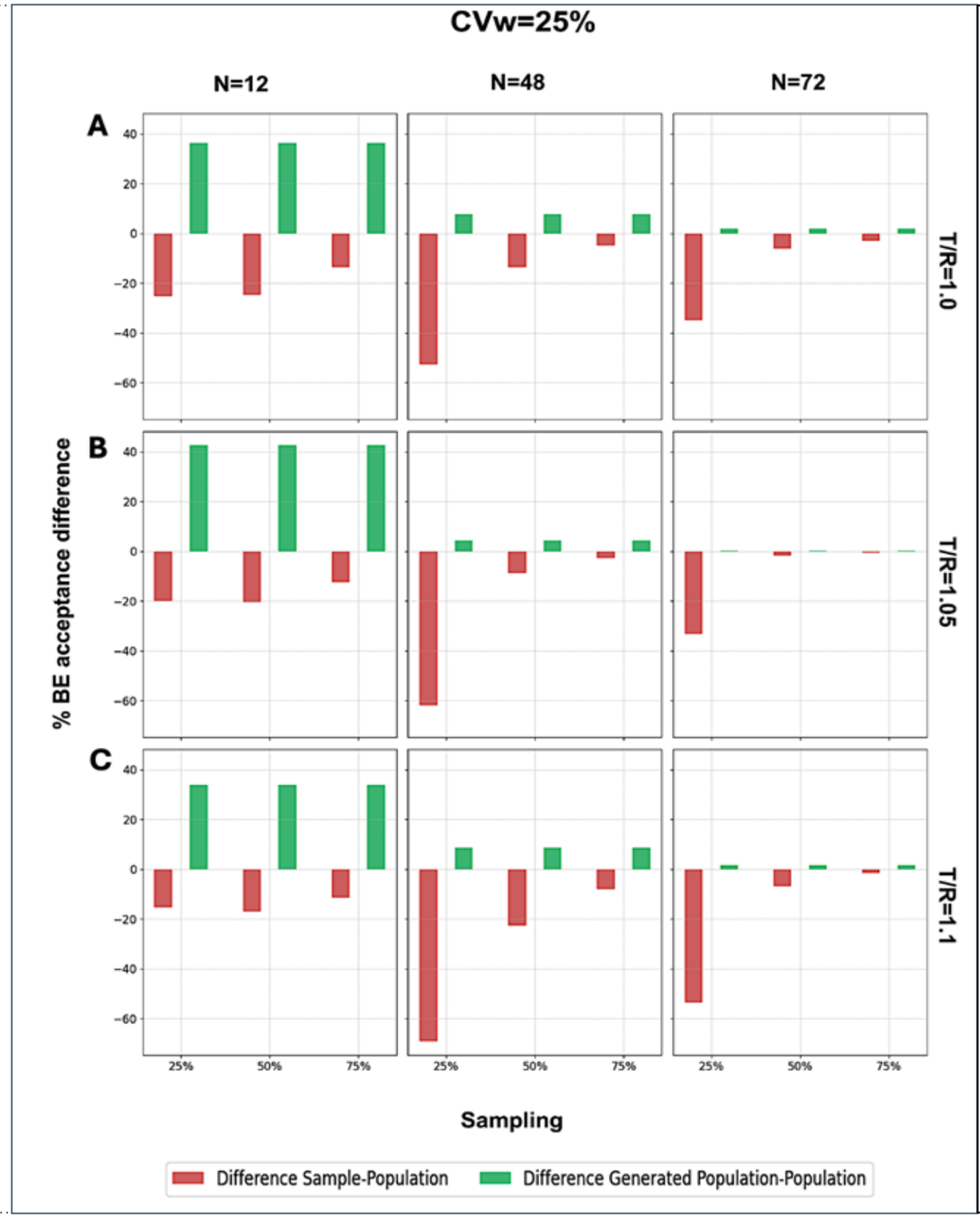

***Figure 5**: Difference in bioequivalence acceptance percentages (positive shown in green bars, negative in red bars) for populations of N = 12, 48, and 72 with CVw = 25%, across different sampling percentages. Panels (a), (b), and (c) depict T/R ratios of 1.0, 1.05, and 1.1, respectively.*

4. Conclusions

Building on our previous research into creating "virtual patients" for clinical studies using WGANs, **this study further investigates their applicability and efficiency in the BE "environment"** , by exploring various scenarios with different levels of variability, sampling percentages, and T/R ratios.

Our findings reveal **significant advantages of integrating AI generative algorithms into clinical trials**. These include *reduced human exposure, shorter study durations, simplified trial processes, decreased workloads for physicians and clinics, and substantial cost savings for sponsors and health agencies*. Importantly, the populations generated by WGANs consistently achieved **BE acceptance rates equal to or greater than those from original populations, outperforming traditional sampling methods, especially in scenarios with high variability and large T/R ratios**.

These results suggest that AI-driven **generative algorithms can deliver comparable or even superior outcomes with fewer real data points**. This supports the broader adoption of AI-based methods in clinical trials. However, successful implementation will require regulatory bodies to establish specific guidelines and criteria for using AI-generated virtual subjects to address potential issues and ensure consistency.