# Improved GSimp - a Flexible Missing Value Imputation Method to Support Generic Drug Development and Regulatory Assessment

**Jing Wang, Xiajing Gong, Liang Zhao, Meng Hu**
**Division of Quantitative Methods and Modeling, Office of Research and Standards, Office of Generic Drugs**
**Center for Drug Evaluation and Research, U.S. Food and Drug Administration, Silver Spring, Maryland, U.S.A.**

## BACKGROUND

- Missing values are commonly seen in pharmacokinetics (PK) or pharmacodynamics (PD) bioequivalence (BE) studies.
- They typically appear as a mixture of different types, such as Missing Not at Random (MNAR) and Missing Completely at Random (MCAR).
- Handling missing values poses non-trivial challenges for generic drug development and regulatory assessment.

## OBJECTIVES

- The objective of this study is to develop a data imputation method suitable for handling missing values in PK or PD BE studies to support generic drug development and regulatory assessment.

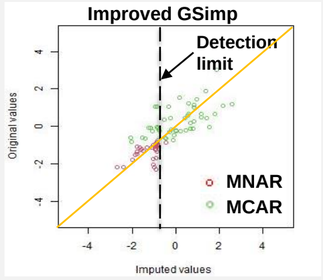
## METHODS

- An iterative Gibbs sampler-based left-censored missing value imputation approach (GSimp) showed superior performance over other mainstream data imputation methods with MNAR data [1].
- We introduced an improved GSimp method ("Improved GSimp" thereafter) by making two improvements to the original GSimp algorithm.
  - Implementing probability-based predictions for imputed values
  - Handling missing data from the mixed missing types
- Simulations with mixtures of different missing types and missing proportions were used to compare the performance of Improved GSimp with other mainstream methods. Normalized root mean square error (NRMSE) and scatter plots were used to evaluate the data imputation performance.
- A real PD BE study data were used to demonstrate the usefulness of the Improved GSimp method in facilitating BE assessment. Missing proportion of the real PD study dataset was 0.11, due to limit of detections.

## RESULTS

### ➢ Improved GSimp

Improvements extend the original GSimp to mixtures of different types of missing values with enhanced reliability and accuracy for imputed values

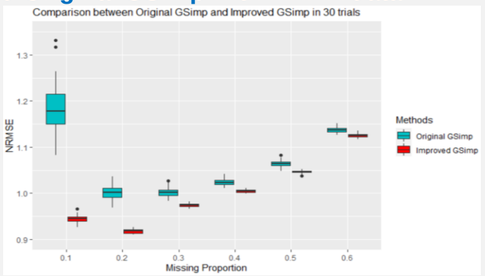



Case example: True value = -1.6752

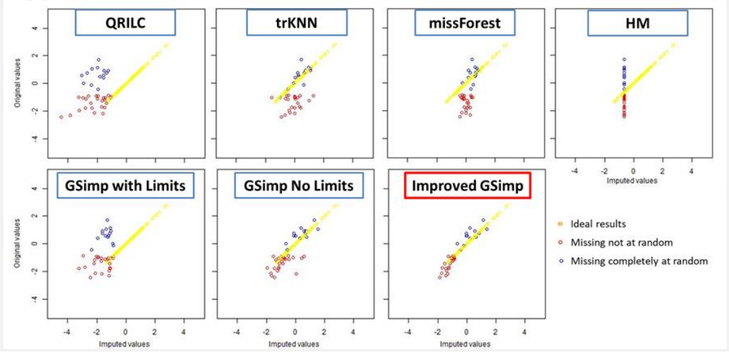
| Imputation Process # | Improved GSimp Imputed Values | Original GSimp Imputed Values |
|---|---|---|
| 1 | -1.6704 | -1.7982 |
| 2 | -1.6684 | -1.6199 |
| 3 | -1.6753 | -1.8384 |
| 4 | -1.6681 | -1.5587 |
| 5 | -1.7003 | -1.7538 |
| 6 | -1.6628 | -1.5551 |
| 7 | -1.6457 | -1.6716 |

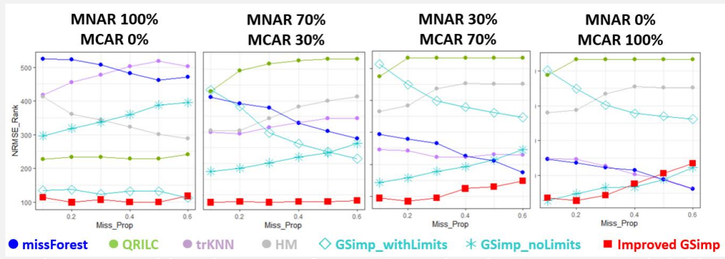
### ➢ Compare Improved and Original GSimp with MNAR data

Imputation process 30 times for simulated MNAR datasets, missing proportion from 0.1 to 0.6. Improved GSimp showed significant lower imputation error than the original GSimp in all




### ➢ Improved GSimp vs. other methods with mixtures of missing types

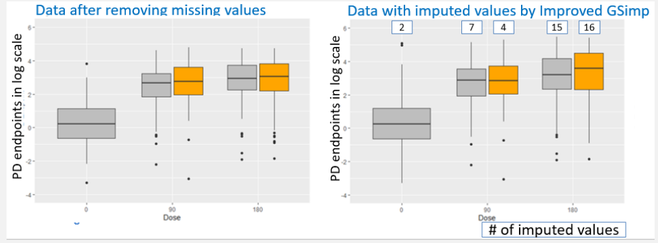



For datasets with mixtures of MNAR and MCAR, Improved GSimp outperformed the original GSimp and other methods.

### ➢ Improved GSimp applied to a real PD BE study

With the imputed values by Improved GSimp, the analysis showed that the calculated 90% confidence interval fell in the acceptance limit of BE.




## CONCLUSION(S)

In this study, we developed an improved GSimp data imputation method for handling mixed types of missing data. The superior imputation accuracy and reliable performance of the improved GSimp showed its great potential to address the missing value imputation challenges in PK or PD BE assessment for generic drug development and regulatory assessment.

## REFERENCES

[1] Lazar, Cosmin, et al. Journal of proteome research 15.4 (2016): 1116-1125.
[2] Wei, Runmin, et al. PLoS computational biology 14.1 (2018): e1005973.
[3] Lenz, Michael, et al. Plos one 15.12 (2020): e0243487.

## ACKNOWLEDGEMENTS

Dr. Jing Wang was supported by an appointment to the Research Participation Program at the Center for Drug Evaluation and Research FDA, administered by the Oak Ridge Institute for Science and Education through an interagency agreement between the U.S. Department of Energy and the U.S. FDA. This poster reflects the views of the authors and should not be construed to represent FDA's views or policies.