



## MDClone Synthetic file comparison report

This report compares the quality of MDClone's synthetic data you have produced with the original data. If you are not familiar with the report please read the [guide explaining how to read and understand the report](#).

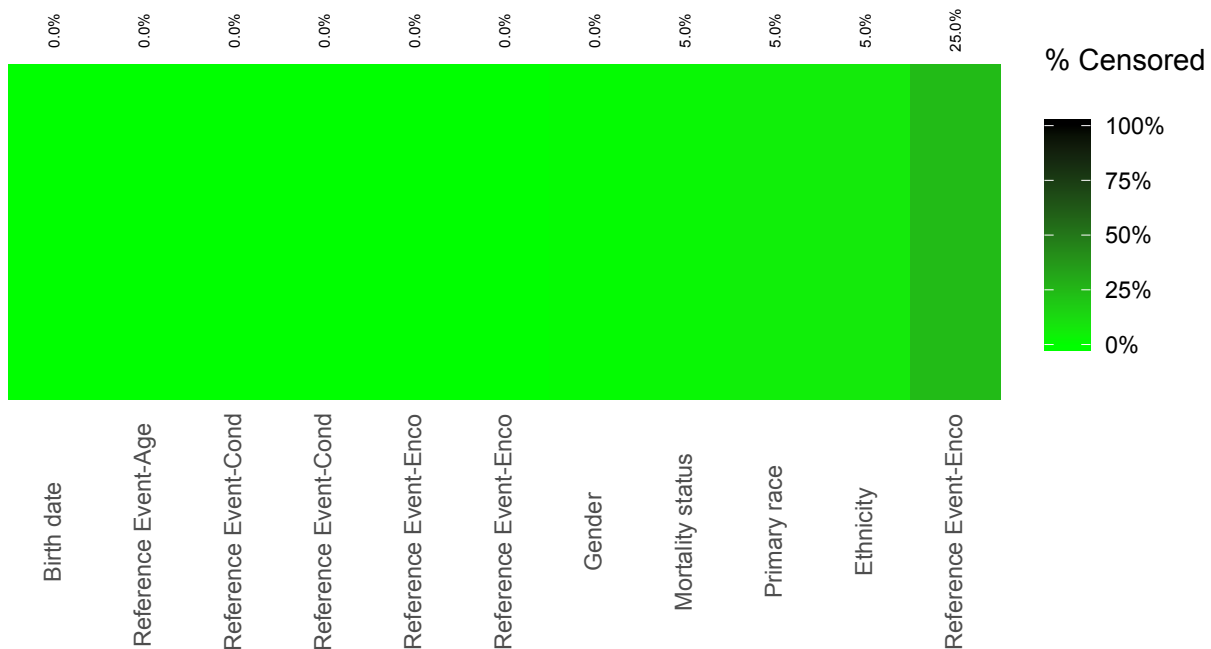
### General

Comparison was created at Friday, July 02, 2021 19:49

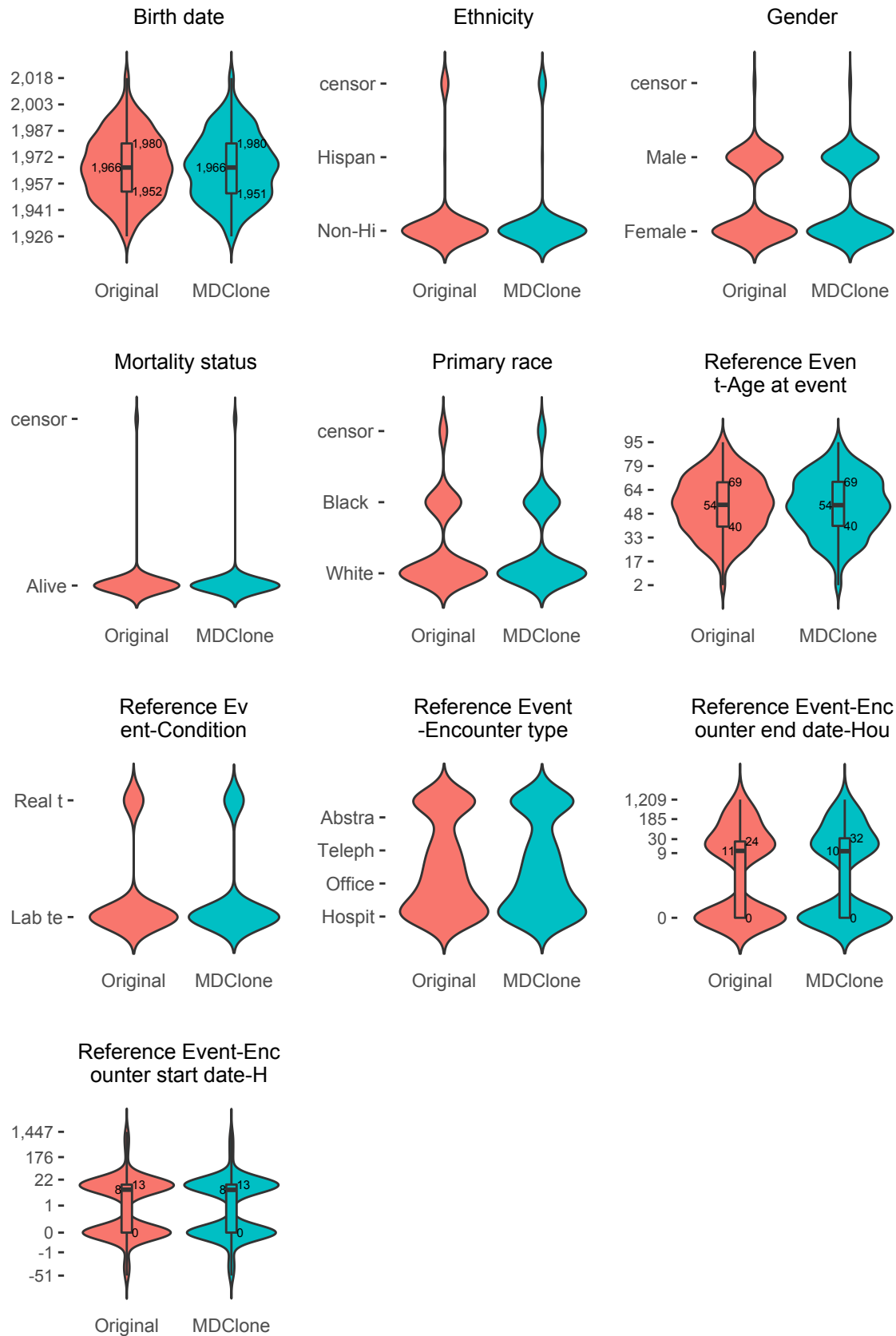
This file includes three parts: this page includes general information about the data and the comparison, followed by figures comparing the distribution of each variable in both original and synthetics data sets. Last there is a comparison of the correlations between each pair of variables.

Synthetic Data includes 425 rows of data.

Approximately ~24.0% of the rows have some censoring. The following figure shows how much censoring occurs for each variable. A light green means no censoring (using the full original data) and black means heavy censoring.



## Variables' distribution



## Correlations

Finally, here is a graph of Spearman's correlation between all pairs of variables, comparing the correlation in the original data set with the correlation of each pair in MDCClone's synthetic data.

When the correlations in the synthetic data are identical to the correlations in the original data they are on the diagonal.

