# Supplementary material
# History distribution matching method for predicting effectiveness of HIV combination therapies

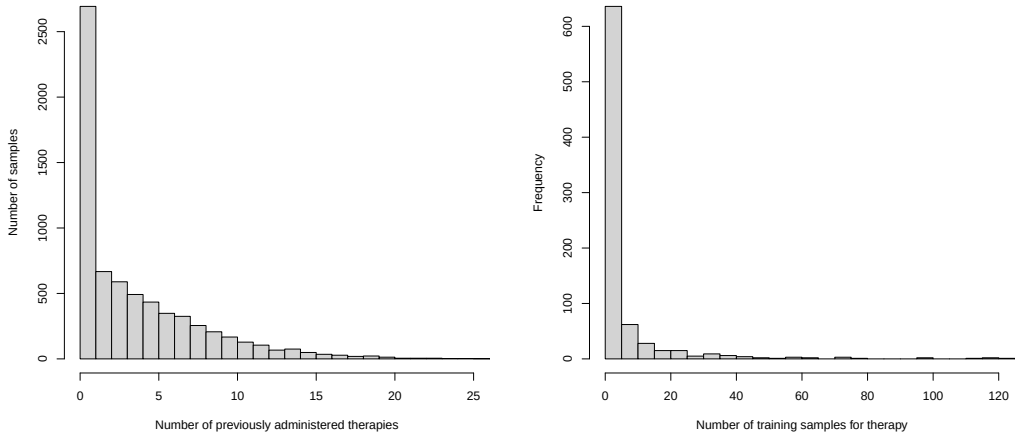
## 1 Tables and figures

Table 1: Details on the bins grouping the test samples based on their corresponding number of previous therapies.

| **Bin** | $0-2$ | $3-5$ | $>5$ |
|---|---|---|---|
| **Sample count** | 807 | 225 | 275 |
| **Success rate** | 89% | 82% | 68% |

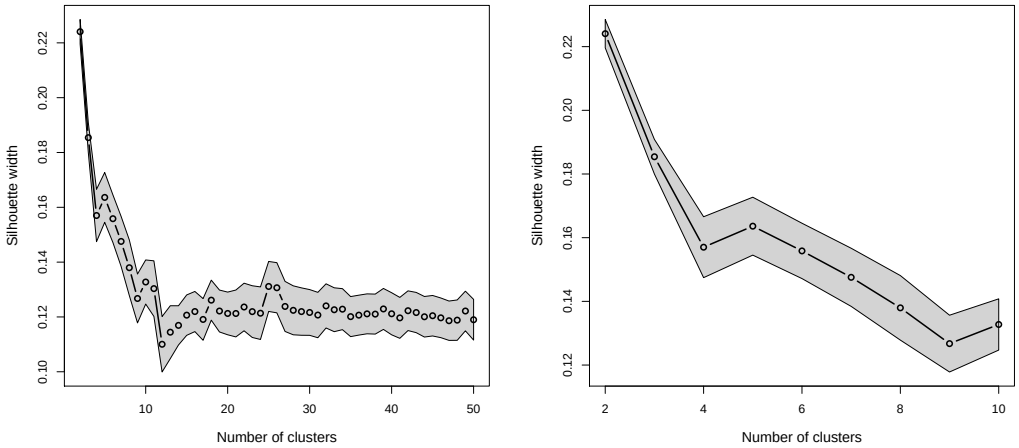
Table 2: Details on the bins grouping the test samples based on the number of training examples for their corresponding therapy combinations.

| **Bin** | $0-7$ | $8-30$ | $>20$ |
|---|---|---|---|
| **Sample count** | 217 | 242 | 848 |
| **Success rate** | 77% | 82% | 85% |




(a) (b)

Figure 1: Uneven representation in clinical data sets: a) histogram that groups all labeled samples in our clinical data set based on their corresponding number of known previous therapies. The histogram illustrates the uneven representation with respect to length of available treatment history in the data; and b) histogram that groups all labeled samples in our clinical data set based on the number of available training samples.




(a) (b)

Figure 2: Silhouette widths with their corresponding standard deviations for different number of clusters: (a) two to fifty; and (b) two to ten. The standard deviations are estimated with the bootstrap method ($B = 100$).

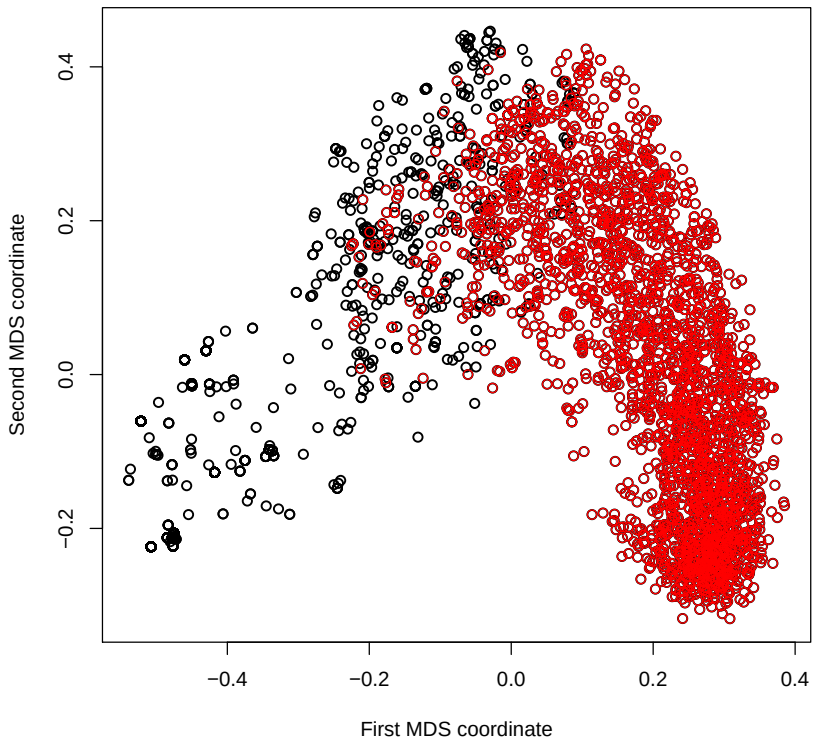



Figure 3: Clustering results for the training data set.

## 2 Model parameters

All model parameters are tuned by using the tuning set chosen in the time-oriented evaluation scenario described in the paper. More specifically, the tested values for the prior and regularization parameters of the logistic regression models are chosen as powers of 10 from the interval $[10^{-4}, 10^{3}]$, and the values for the smoothing parameter $\gamma$ of the sample-specific weights are chosen equidistantly from $[0, 1]$.

## 3 Assigning a label to HIV therapy samples

Each therapy sample in our HIV clinical data set is labeled as success or failure based on the virus load measured in the course of its corresponding therapy. If the virus load drops bellow 400 *cp/ml* in the period from 21 days after the start of the therapy to the end of the therapy we label it with *success* (1); otherwise with a *failure* (−1). Figure 4 shows a schema of the labeling procedure. In this way we create a labeled data set that includes 6336 therapy samples with 638 distinct therapy combinations.

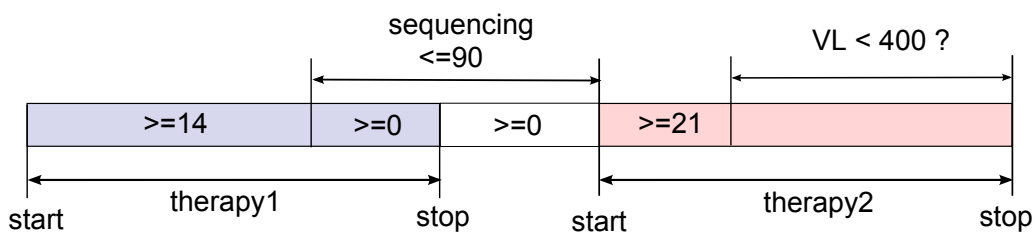



Figure 4: Assigning a label and a viral sequence to *therapy2*, where *therapy1* and *therapy2* are two consecutive therapies administered to a patient.