1. Data Complexity Reduction **ClustALL** A. **Dendrogram** Deptht₁ Deptht₃ **Input Data** Deptht, **Grouped Variables** Complex **PCA PCA** disease В. **PCA** reductions Embedding₃ Embedding₁ Embedding₃ Routine clinical data Embedding_n Components 2. Stratification Process 3. Consensus-based Stratifications

For each embedding:

- **Distance metric**
- Correlation
- Gower

- **Clustering Method**
 - H-Clust
 - K-means
 - K-medoids

Evaluation of the optimal number of clusters

 $K \in 2, 3, 4, 5, 6$ (WB-ratio, Dunn and Silhouette)

Stratifications

(Embedding + Distance + Clustering)

- 1. Population-based robust stratifications by bootstrapping
- 2. Parameter-based robust stratifications

