Longitudinal data modeling for Alzheimer's disease: Examples of Predictive and Generative models

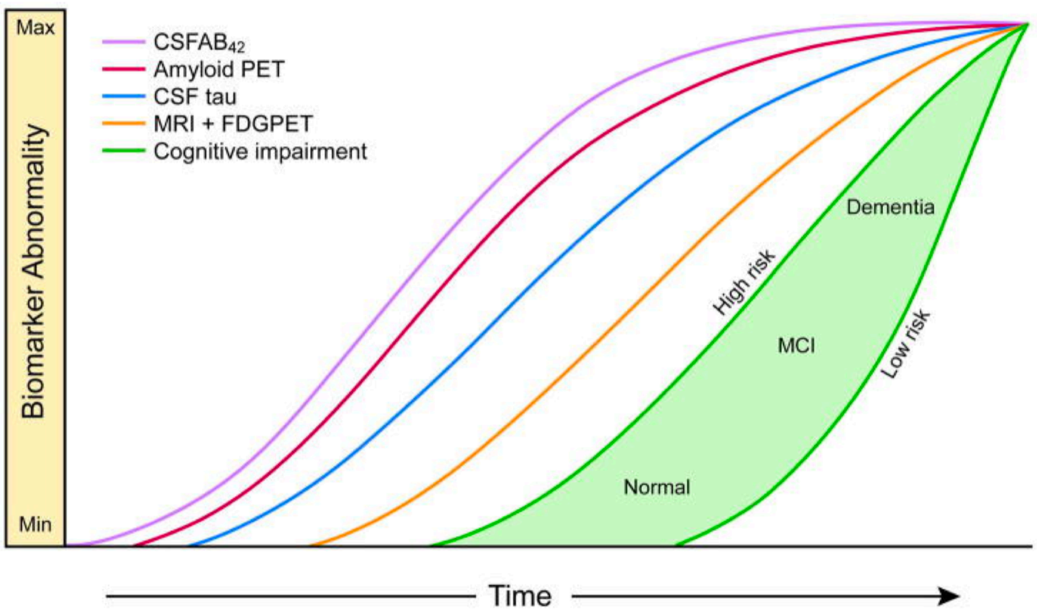
Supervisor:
Prof. Dr. Martin Hofmann-Apitius
Co-supervisor:
Prof. Dr. Holger Fröhlich

Presented by:
Meemansa Sood

Fraunhofer SCAI

Alzheimer’s Disease

- Understanding the pathophysiological process of the disease
- Stage specific disease progression
- Understanding the role of biomarkers in the predictive analyses of progression of disease

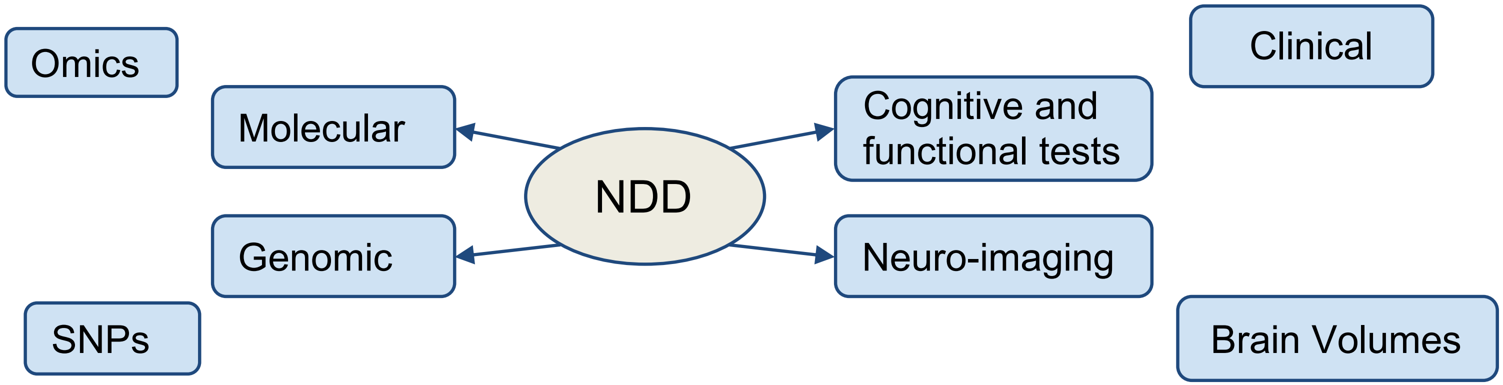


The figure is adapted from Jack, et. al. 2013

Fraunhofer SCAI

Scientific Challenge

- Understand and model diseases in a longitudinal manner
- Need to agglomerate and understand multi-scale data

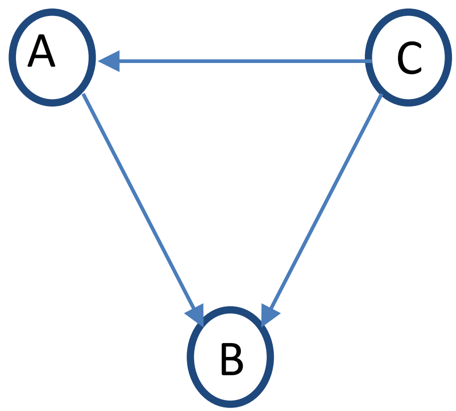


- Issues:
  - Failure of several clinical trials around established hypotheses
  - Intermittent patient drop outs
  - Lack of enough neurological data

Machine Learning Strategies

Bayesian Networks (BNs)

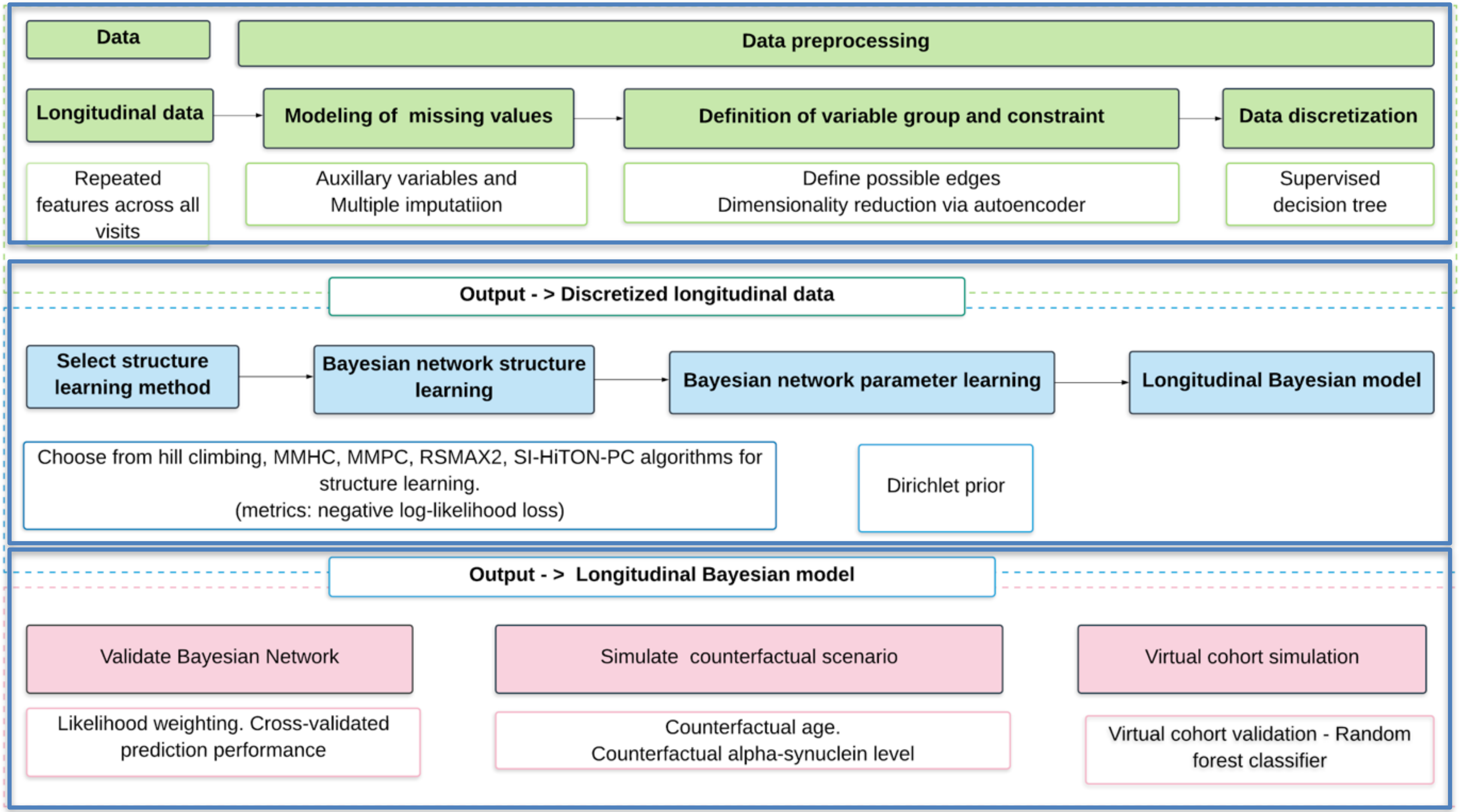
- BNs gives probabilistic relationships among a set of variables
- Generative as well as predictive model



- BN can model temporal processes
- Model can (partially) learn causal relationships from multi-modal data

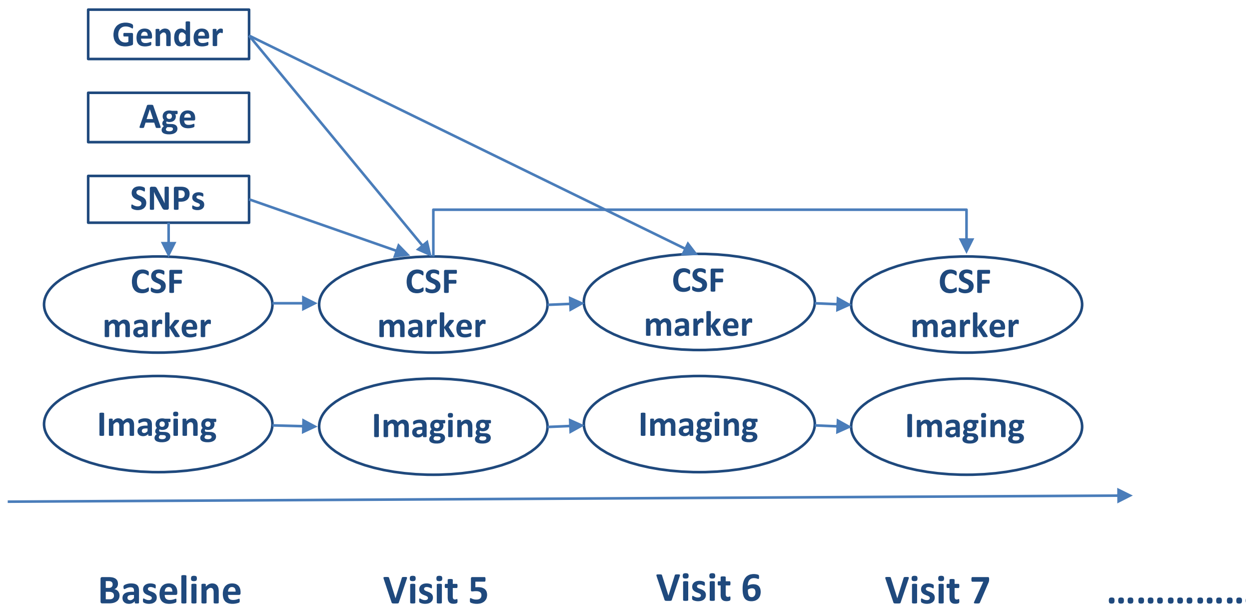
Fraunhofer SCAI

Overview of Workflow



Fraunhofer SCAI

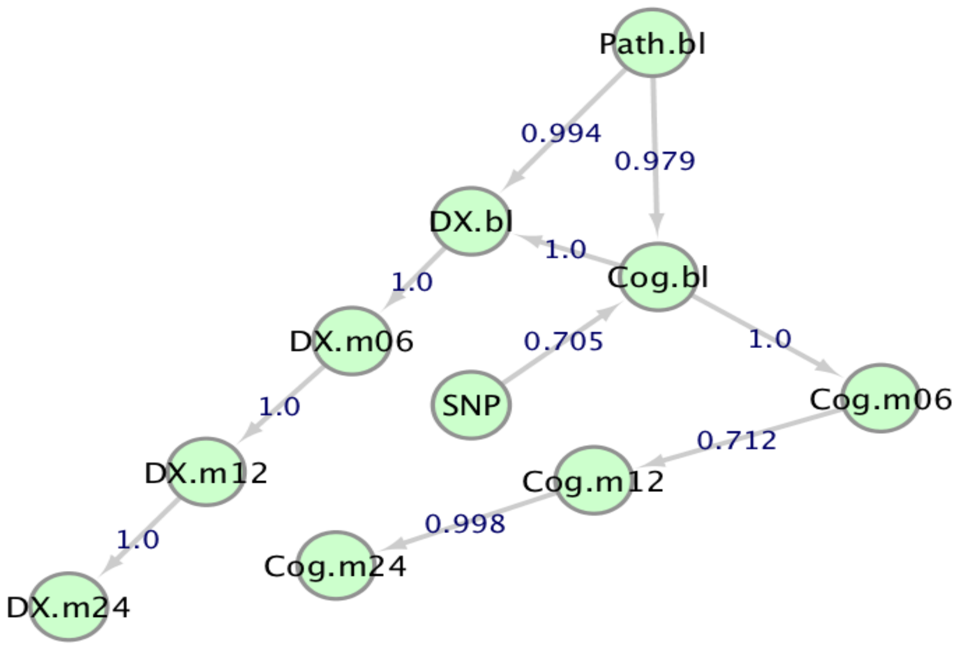
BN as generative model

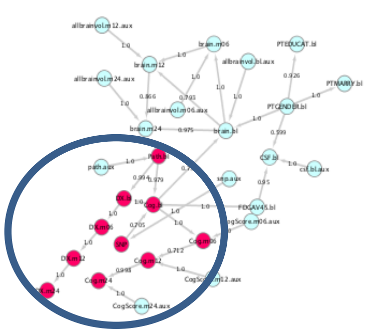


- Unfolding of time dependent variables
- Constraints for possible edges

Fraunhofer SCAI

Results of BN Structure Learning





```
Variable Importances:
                     variable relative_importance
1                    CDRSB.bl            1.000000
2                     MMSE.bl            0.972977
3                   ADAS11.bl            0.953378
4                      FAQ.bl            0.928205
5                   ADAS13.bl            0.908452
6          RAVLT.immediate.bl            0.723133
7 RAVLT.perc.forgetting.bl               0.571820
8           RAVLT.learning.bl            0.517162
9         RAVLT.forgetting.bl            0.026748
```


Fraunhofer SCAI

BN used as a predictive model

Baseline feature

- Predict each meta score at month 6

Baseline + month 6

- Predict each meta score at month 12

Baseline+ month 6 + month 12

- Predict each meta score at month 24

Baseline...to. ....month24..

- Predict each meta score until last month

- **Cross validated Predictions**
- **10 fold cross validation performance**

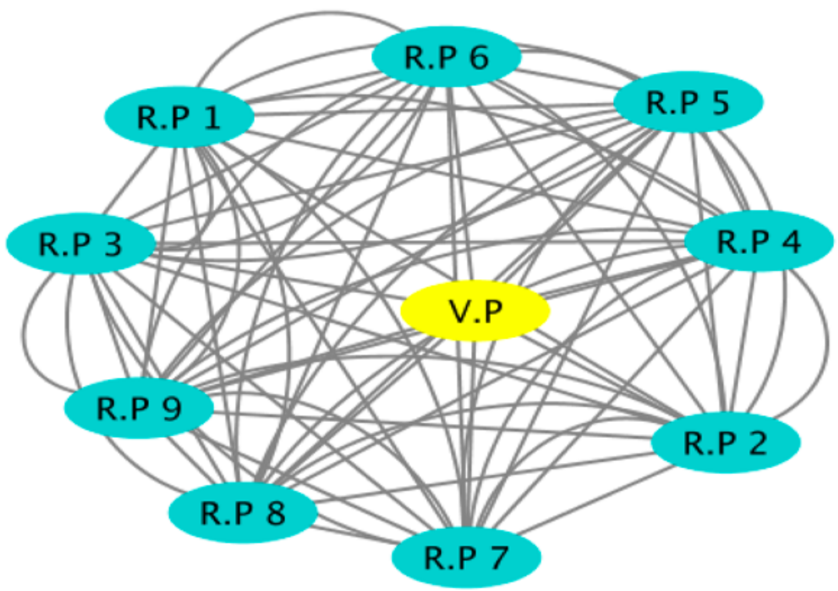
Fraunhofer SCAI

Virtual Patient Simulation

1

Comparison based on the distance measure

- Gower's generated dissimilarity matrix
- k nearest neighbours of each virtual patient determined (k = 9)*



- **Develop a classifier**

Conclusion

- Bayesian networks can be used for obtaining inter-dependencies among multivariate features

- Predict values of features at future time points from previous time points

- Quantitative evidence, can be further developed for data-driven virtual clinical trial and creation of virtual patients

Acknowledgement

- Akrishta Sahay
- Reagon Karki
- Prof. Dr. Holger Fröhlich

Thank you