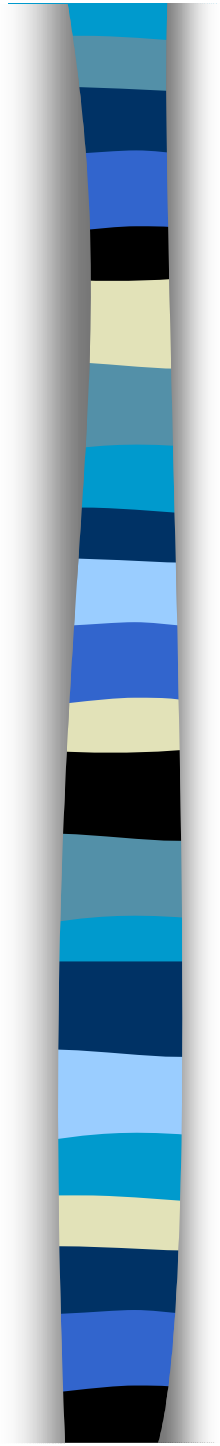
Prediction of Insulin Sensitivity Index using Bayesian Networks

Susanne G. Bøttcher
Claus Dethlefsen

Aalborg University
and
Novo Nordisk A/S



Insulin Sensitivity Index

- Insulin Sensitivity Index (S_I) measures the fractional increase in glucose clearance rate during an IVGTT

- A low S_I indicate a risk of developing type 2 diabetes [Martin et al]

Aim

- Estimate insulin sensitivity index based on measurements of plasma glucose and serum insulin levels during an oral glucose tolerance test in individuals with normal glucose tolerance

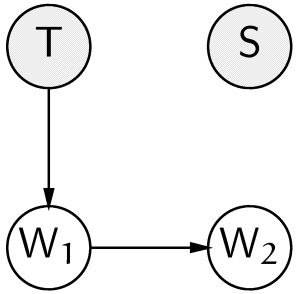
Method

- 187 subjects without known diabetes
- IVGTT determines insulin sensitivity index
- OGTT with measurements of plasma glucose and serum insulin levels at time points 0, 30, 60, 105, 180, 240

Bayesian Networks

- A Directed Acyclic Graph (DAG) $D = (V, E)$.
- To each node v with parents $\mathrm{pa}(v)$, there is attached a local probability distribution, $p(x_v|x_{\mathrm{pa}(v)})$.
- Bayes' rule: $p(\text{Unobserved}|\text{observed}) \propto p(\text{Unobserved})p(\text{Observed}|\text{Unobserved})$

DAG:

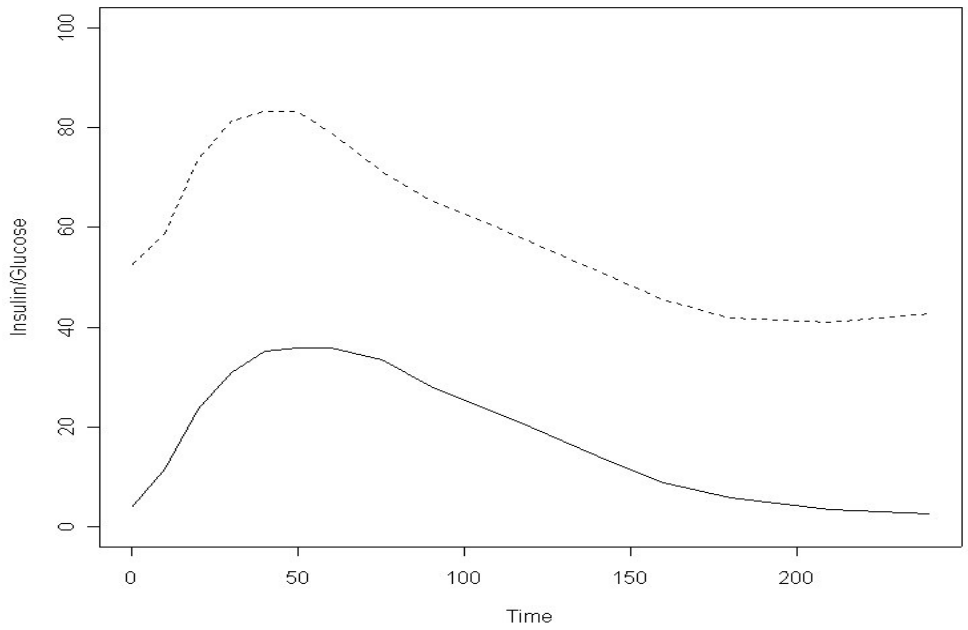


Joint probability distribution:

$$p(S, T, W_1, W_2) = p(S)p(T)p(W_1|T)p(W_2|W_1)$$

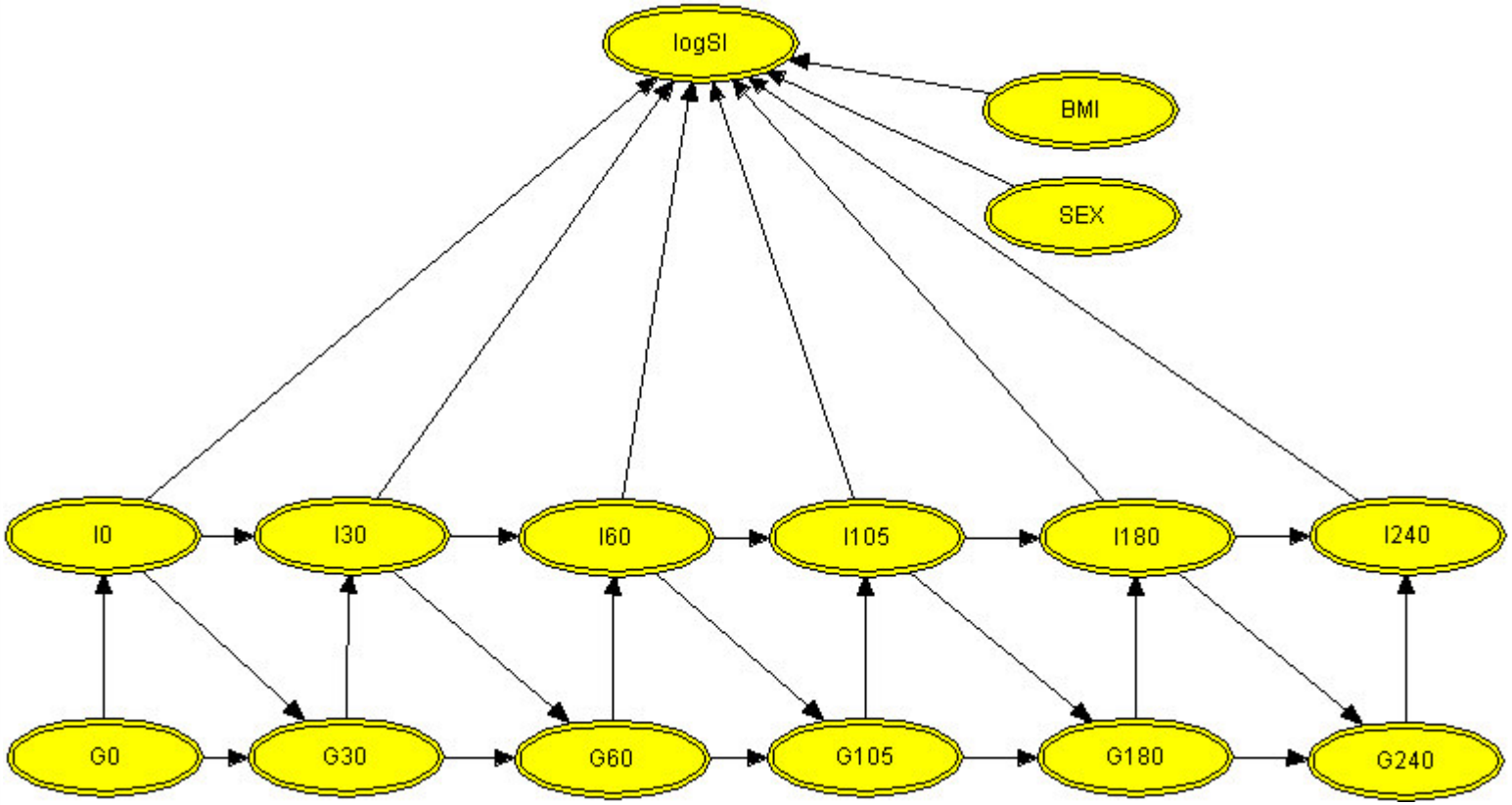
Data

- 140 subjects as training data
- 47 validation data



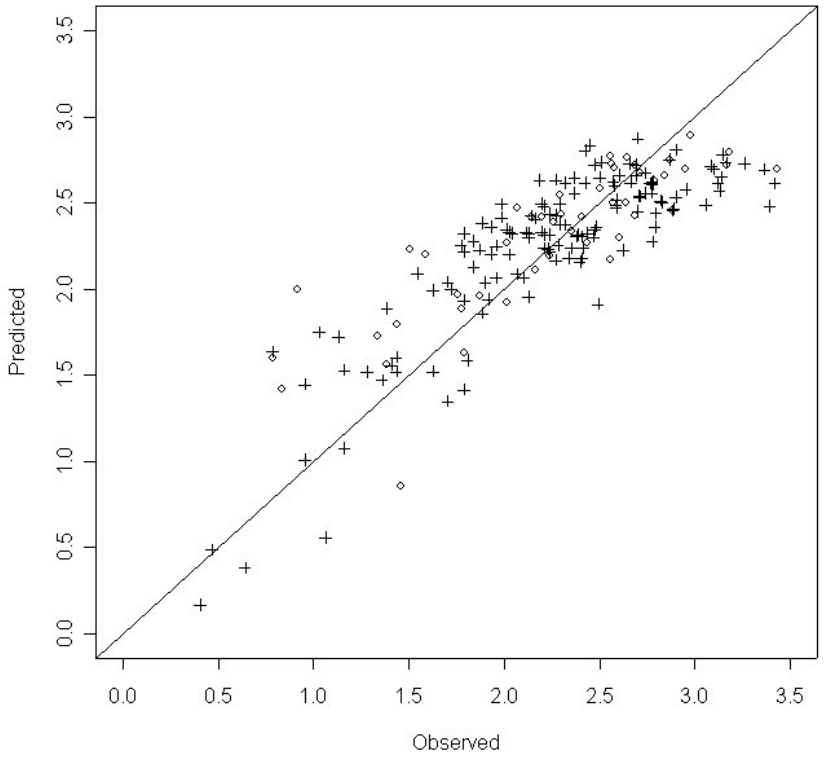
Dashed:
avg. blood glucose
Solid:
avg. insulin

Bayesian Network



Result

- Bayesian network learned with training dataset using DEAL
- Predictions for training data (+) and validation data (o) obtained using Hugin
- SD = 0.36

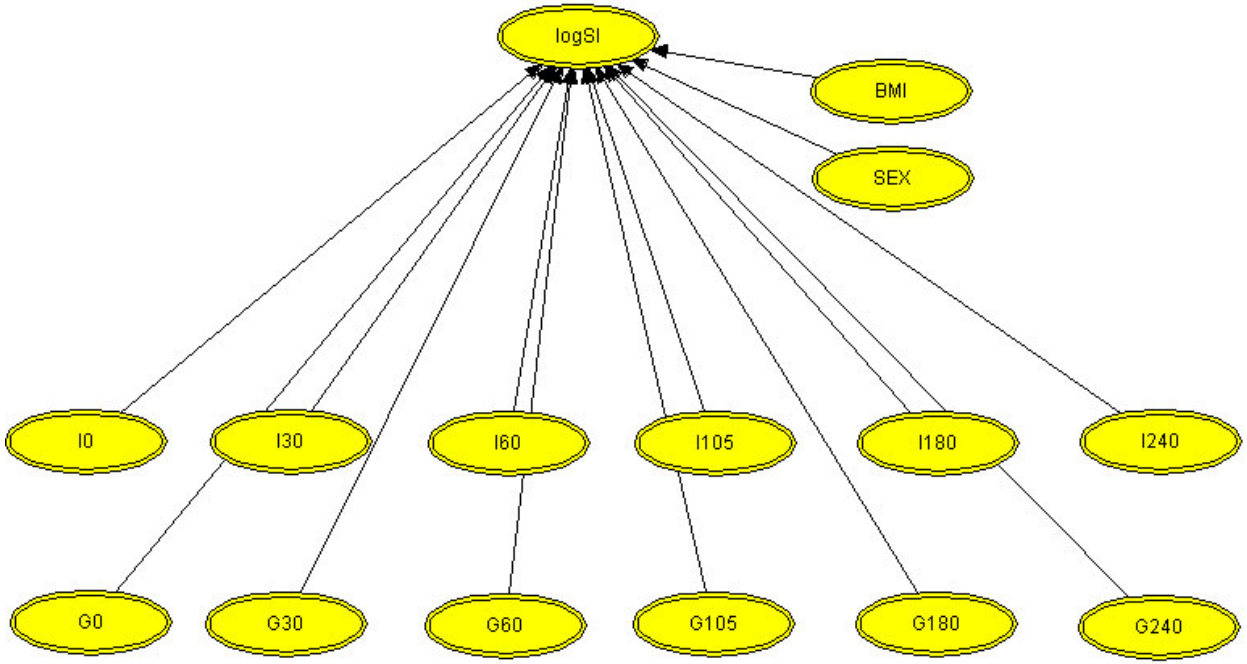


Comparison with previous result

Hansen et al used a multiple regression analysis

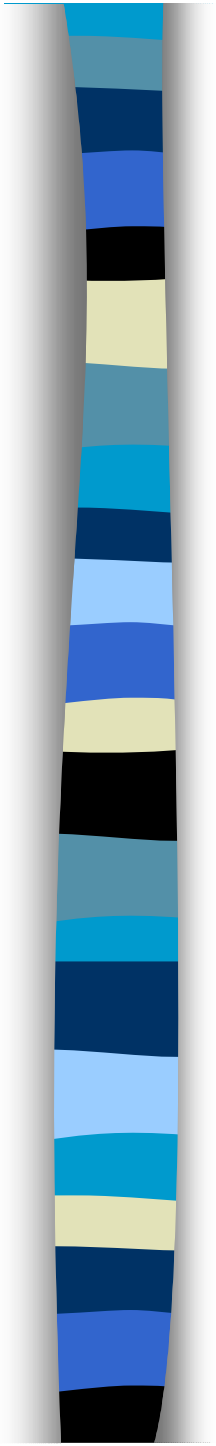
Log(S.I) ~ BMI + SEX + G0 + I0 + G30 + I30 + G60 + I60 + G105 + I105 + G180 + I180 + G240 + I240

SD = 0.39 for the validation data after using the training data for estimation of parameters



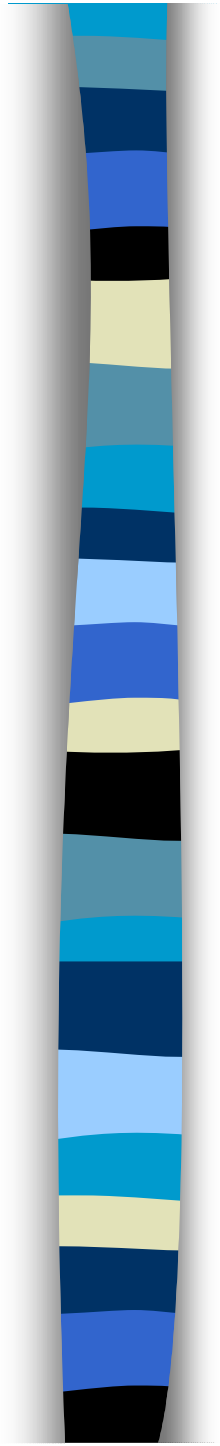
Advantages using Bayesian networks

- Qualitative representation of knowledge. Correlation of all variables is modelled in a complex dependency structure
- Prior information is combined with data in the learning process
- Observations at all timepoints are not needed



Software

- Hugin: http://www.hugin.com
  Prediction in Bayesian networks
- R: Free software http://sunsite.dk/R
  Statistical software
- Deal: http://www.math.auc.dk/novo/deal
  Package for R
  Learning of parameters and structure



What can DEAL do for us?

General setup of Bayesian Networks with mixed variables

- Graphical interface.
- Setup prior networks.
- Master prior procedure for determining parameter priors for all networks.
- Learning parameter posteriors from data.
- Calculation of network score.
- Structure learning using heuristic search.
- Supports time series data.
- Simulation of datasets.

Acknowledgements

- The work is supported by Novo Nordisk A/S
- The idea to the study and collection of data were done by Hansen, Drivsholm, Urhammer, Borch-Johnsen, Pedersen

Bibliography

- Bøttcher and Dethlefsen, 2002. Deal: A package for learning Bayesian networks.
- Cowell, Dawid, Lauritzen, Spiegelhalter, 1999. Probabilistic Networks and Expert Systems.
- Hansen, Drivsholm, Urhammer, Borch-Johnsen, Pedersen, 2002. ”Assessment of insulin sensitivity index and acute insulin reponse from an oral glucose tolerance test in subjects with normal glucose tolerance, unpublished paper
- Martin, Warram, Krolewski, Bergman, Soeldner, Kahn, 1992. ”Role of glucose and insulin resistance in development of type 2 diabetes mellitus: results of a 25-year follow-up study. The Lancet.