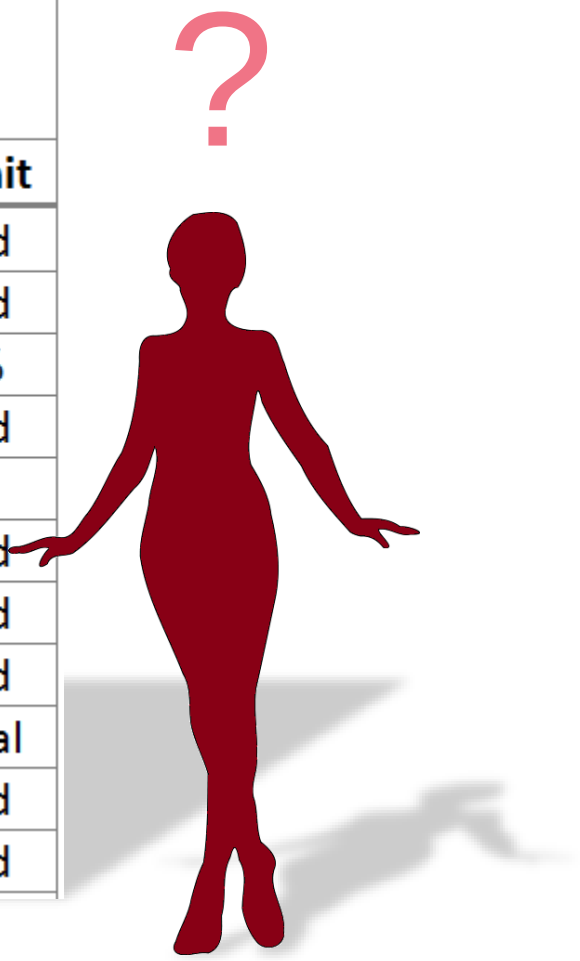
Modeling the Effects of Nutrition with Mixed-Effect Bayesian Network

Jari Turkia
CGI, Next
University of Eastern Finland, School of Computing

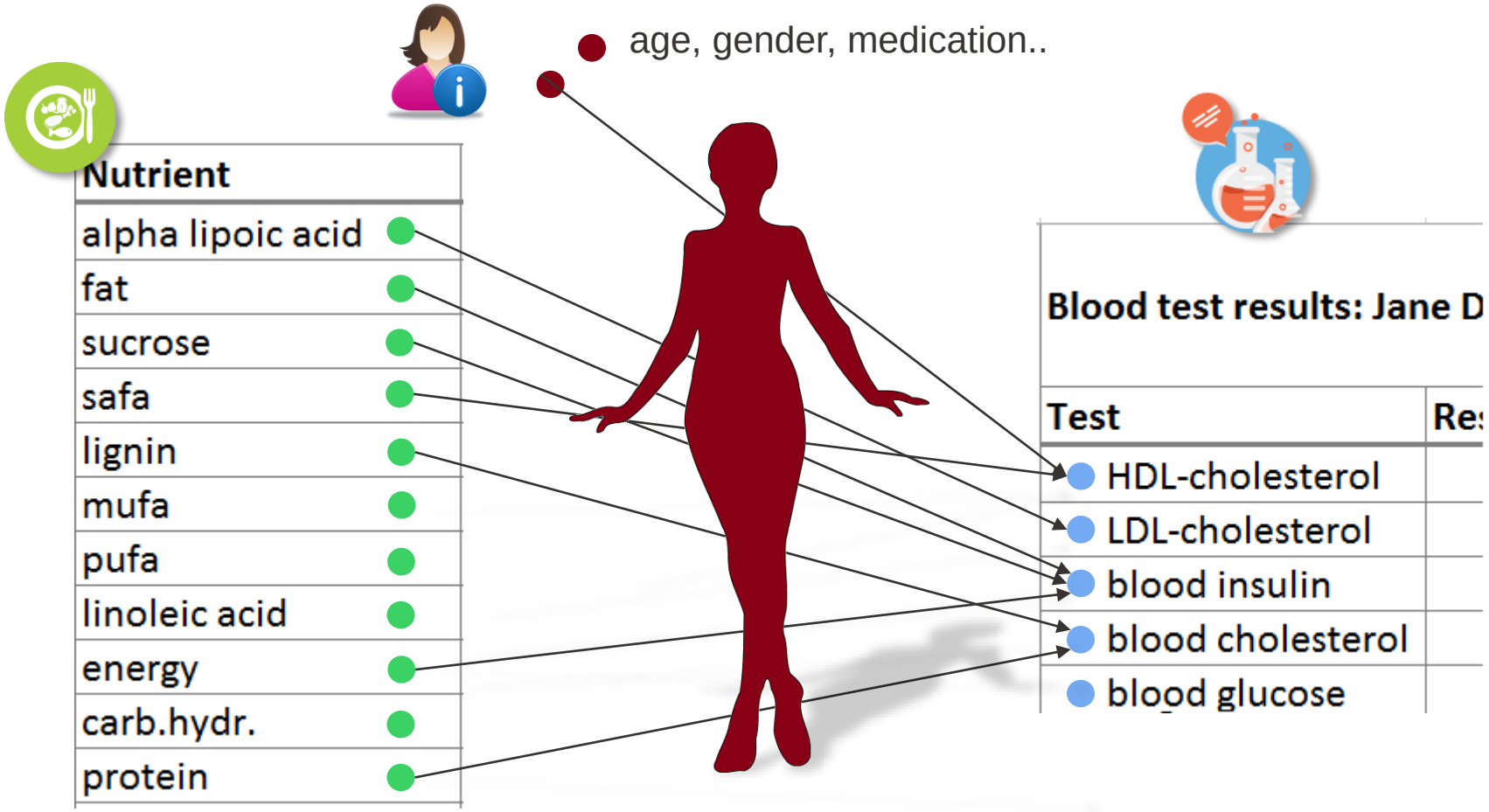

Do eggs raise cholesterol levels?
Why my insulin levels are high?
Can I have more coffee?
What if you could eat in
a way your body works?

General guidelines won't answer to these!

Recommendation: VRN-14/daily: Women 31-60 yrs			
Nutrient	**Lower limit**	**Upper limit**	**Unit**
alpha lipoic acid	36,00	45,00	g/d
fat	59,00	106,00	g/d
sucrose		10,00	E%
safa	23,00	53,00	g/d
lignin			
mufa	12,00	26,00	g/d
pufa	12,00	26,00	g/d
linoleic acid	0,71	1,53	g/d
energy	2110,00	2380,00	kcal
carb.hydr.	237,00	357,00	g/d
protein	53,00	119,00	g/d



Towards personal understanding

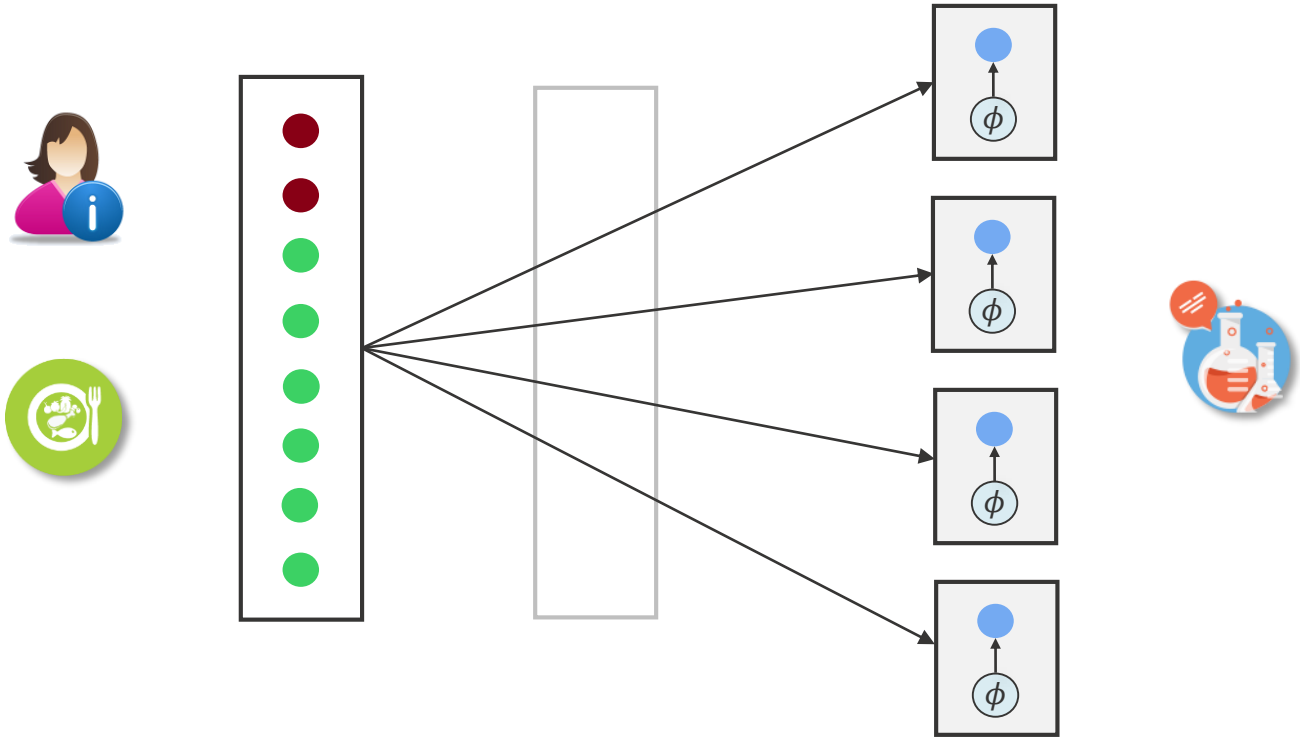


So, it's a graph! Which graph describes the reactions best?

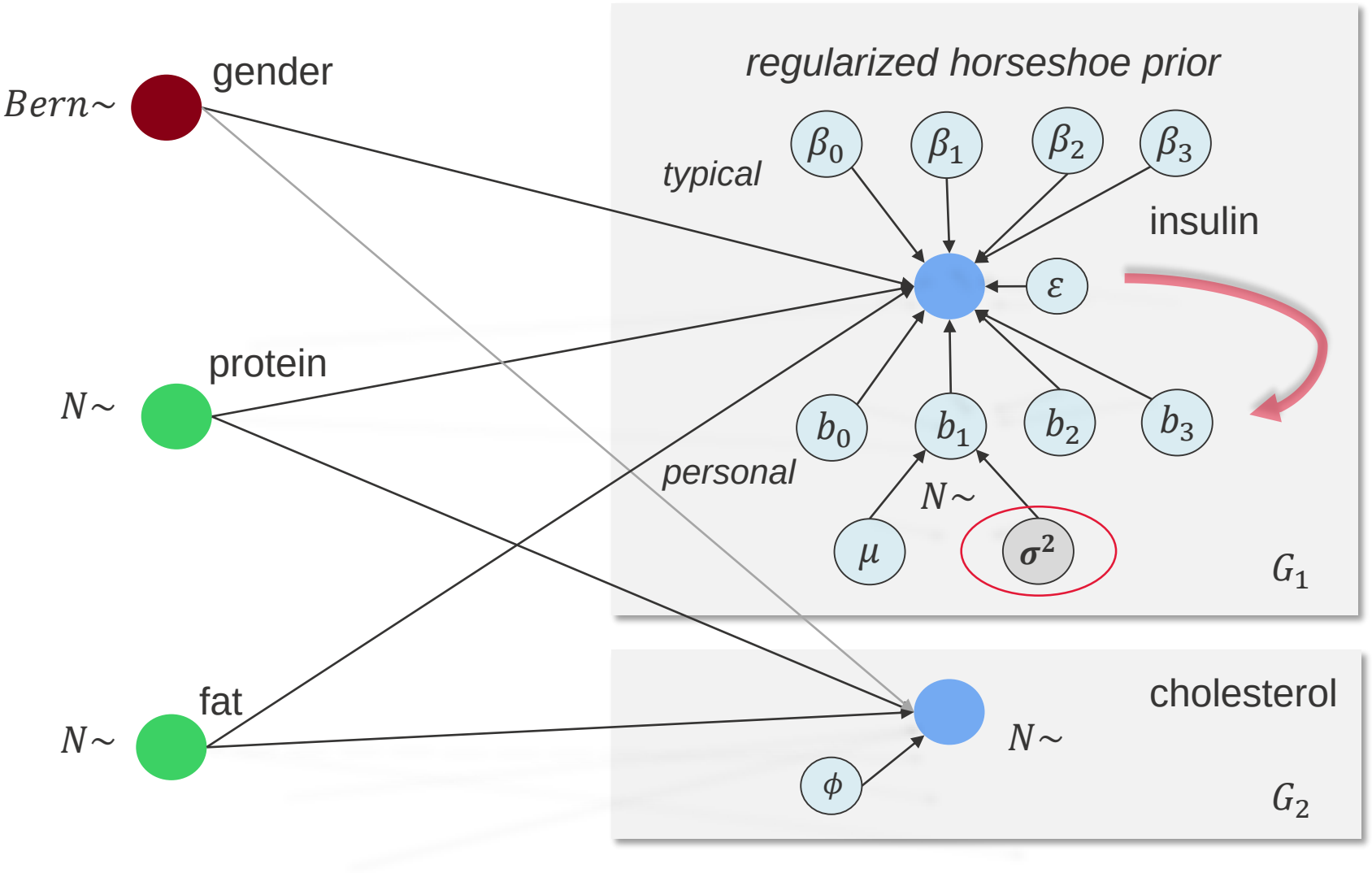
From the graph to a graphical model

- Joint distribution factorizes into local distributions by Markov boundaries
- Finding an optimal graph directly is hard. It is easier to search for optimum.
- This search can be narrowed down with prior information

$$P(G|D) \propto P(D|G)P(G) = \prod P(X_i|pa(X_i), G_i)P(G_i)$$



Hierarchical local distributions - $P(D|G_i)$



The data: Sysdimet study

106 patients

originally RCT of Nordic diet

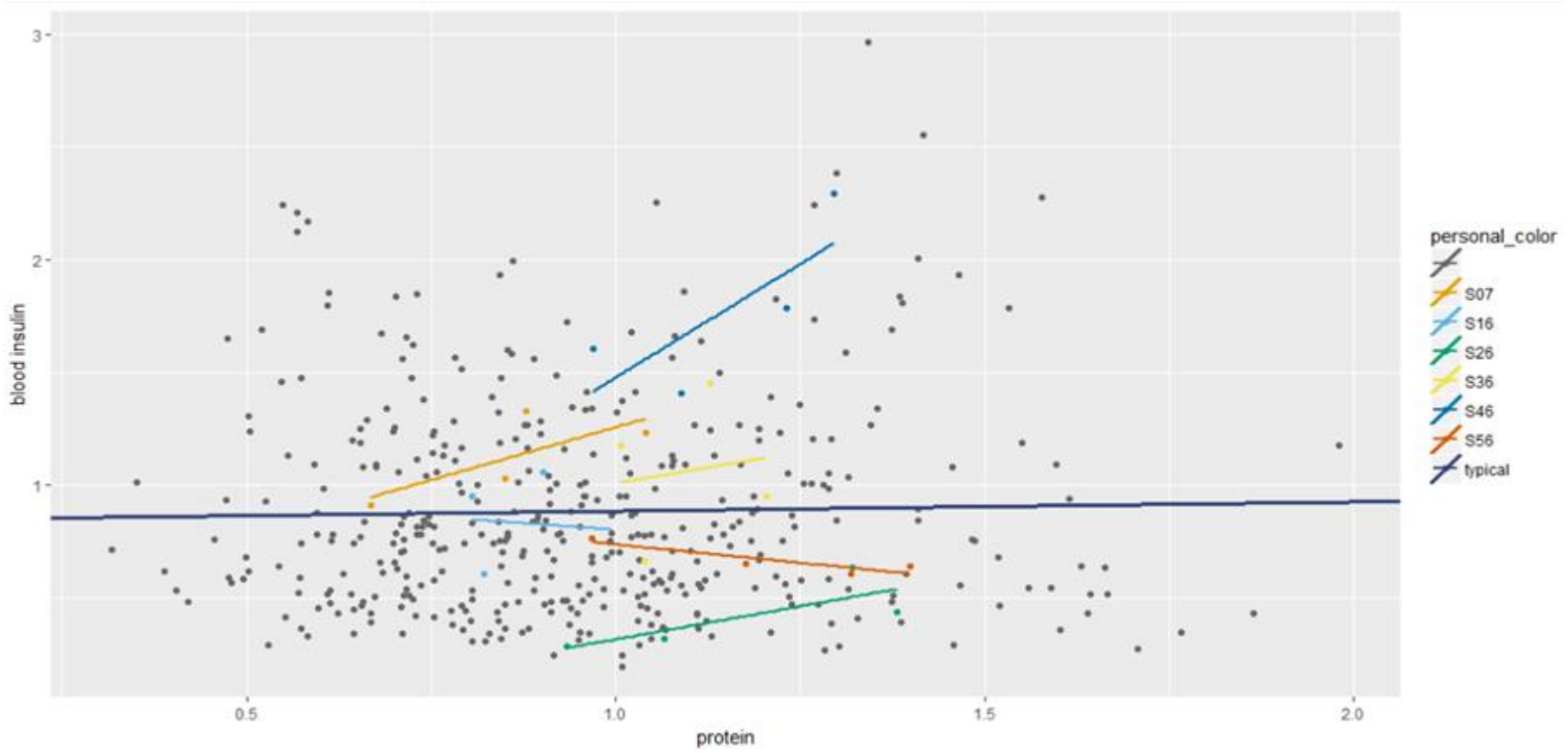
4 repeated measurements during 12 weeks

food diaries and blood tests

17 nutrients, 5 blood tests and personal info

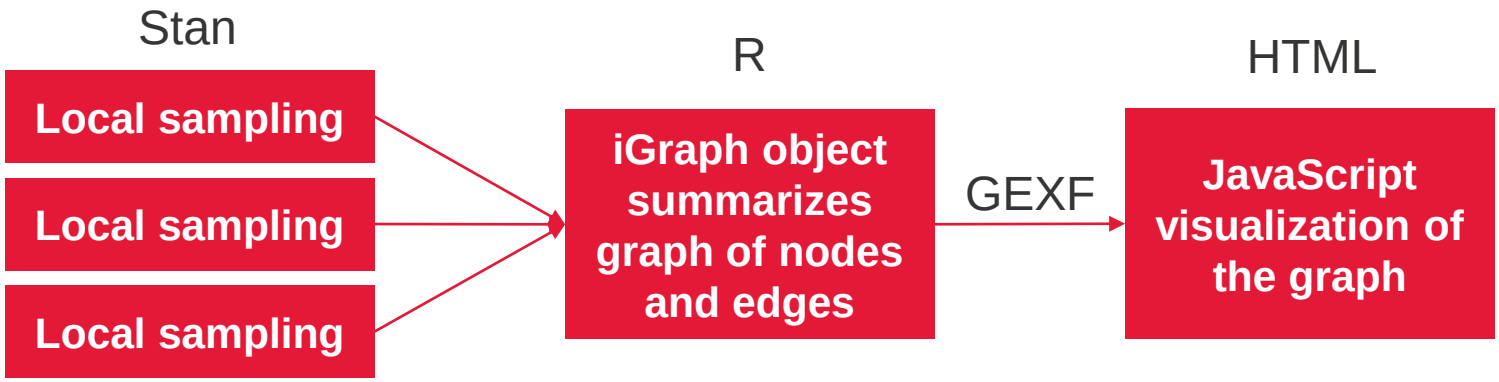
1 week response time before test

Example effect: Protein's effect to blood insulin



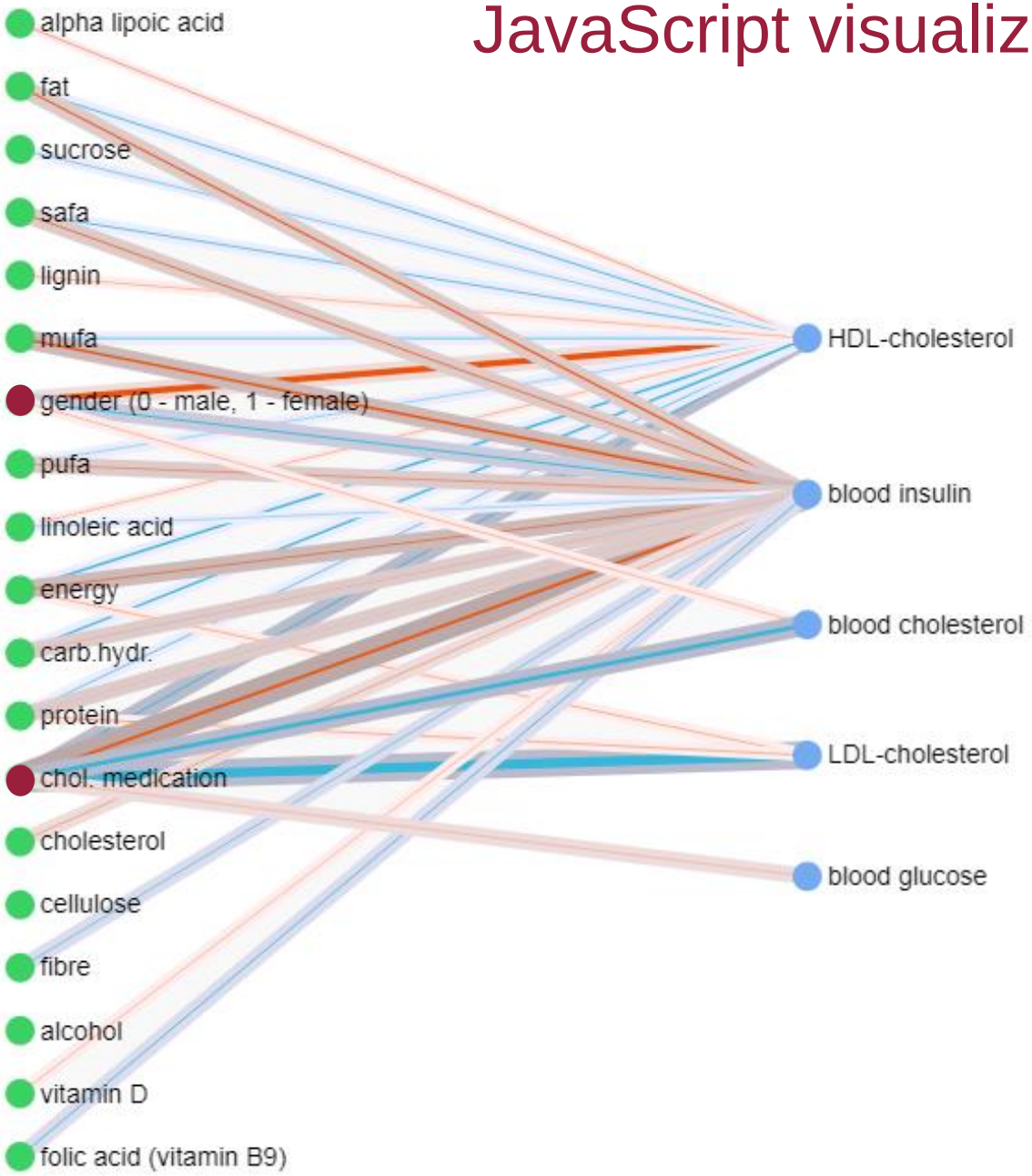
- Linear model for single effects is considered adequate since we are more interested in the system as a whole
- Modeling can reveal if these personal differences are just random noise or possibility for personalization

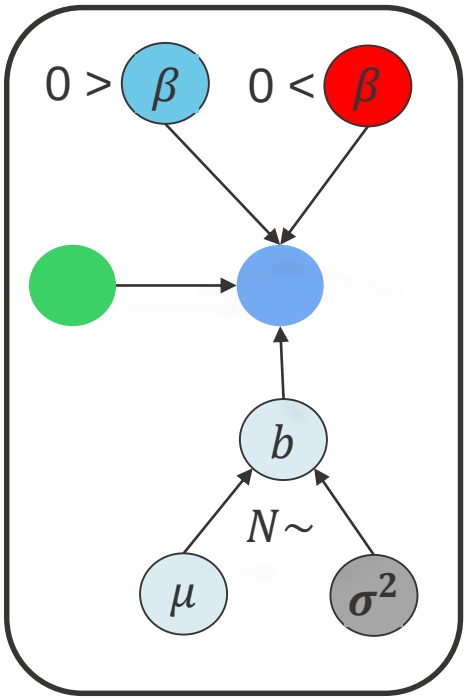
Implementation: Estimation with R and Stan

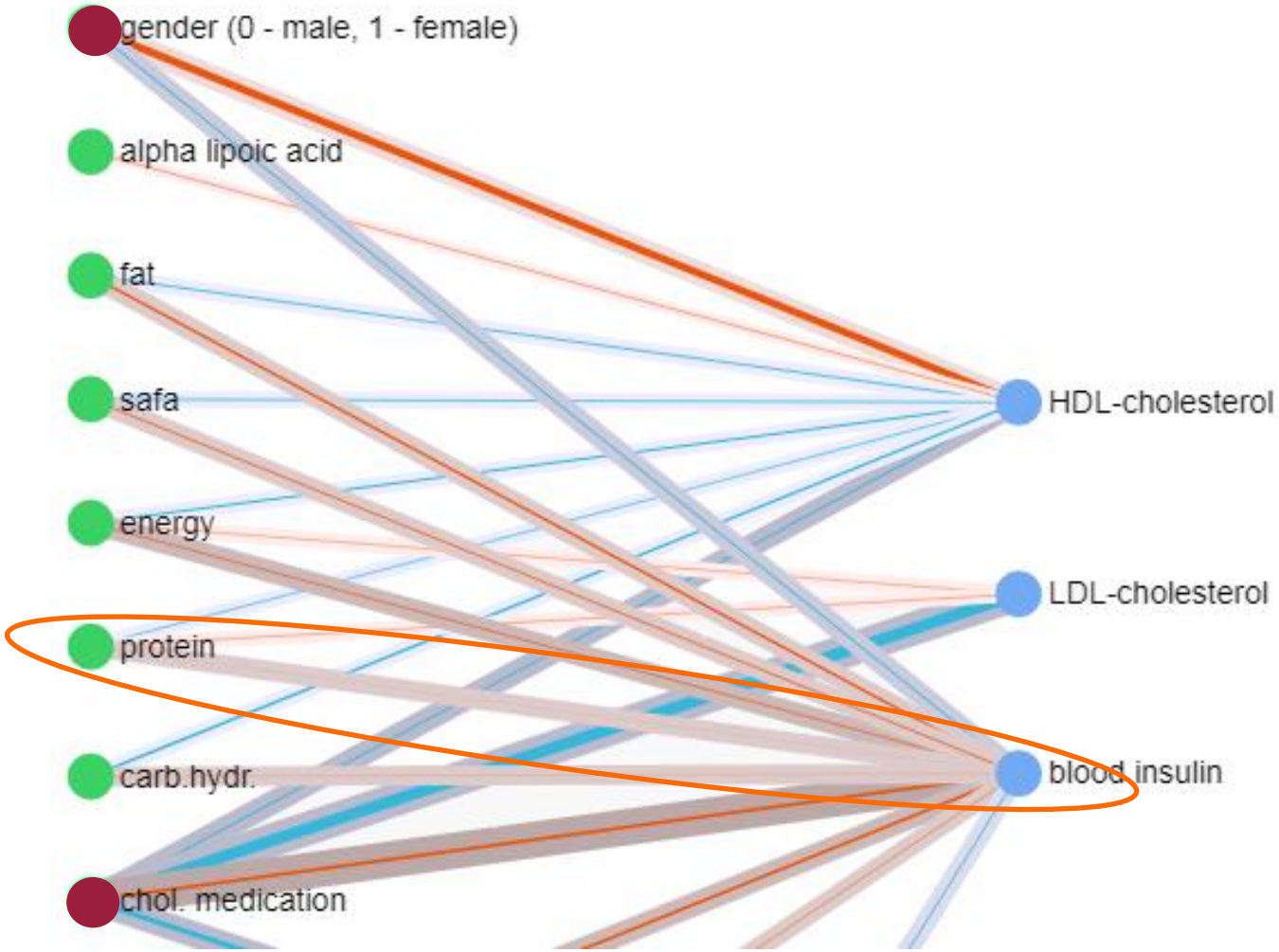


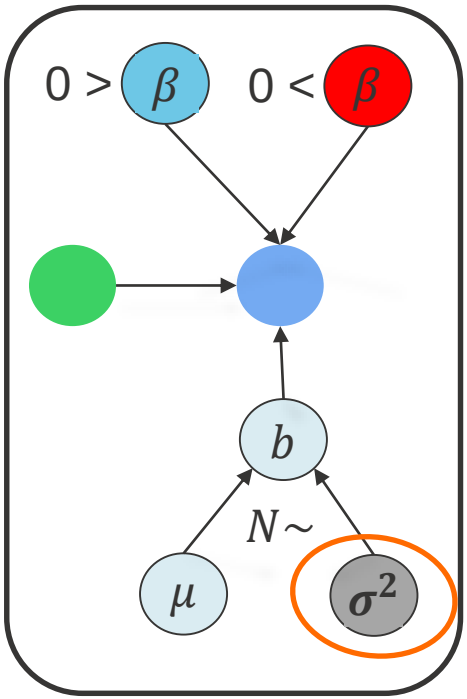
- The graph search is implemented by R code that estimates local distributions for every non-root node with hierarchical Stan-model, and gathers summaries to an iGraph object.
- The notebook includes also a responsive JavaScript visualization for the graph. Connections that are not typically or personally relevant are removed.
- The model can be accessed from all these three layers...

JavaScript visualization






gender (0 - male, 1 - female)
alpha lipoic acid
fat
safa
energy
protein
carb.hydr.
chol. medication
HDL-cholesterol
LDL-cholesterol
blood insulin


0 > β
0 < β
b
N ~
μ
σ²

Let's query the graph – for typical effects

From	To	Beta (normalized)
cholesterol medication	LDL-cholesterol	-0.026897912
cholesterol medication	blood cholesterol	-0.011823168
carbon hydrate	HDL-cholesterol	-0.006123904

From	To	Beta (normalized)
gender (female)	HDL-cholesterol	0.017239135
cholesterol medication	blood insulin	0.008108980
multisaturated fats (mufa)	blood insulin	0.007027948

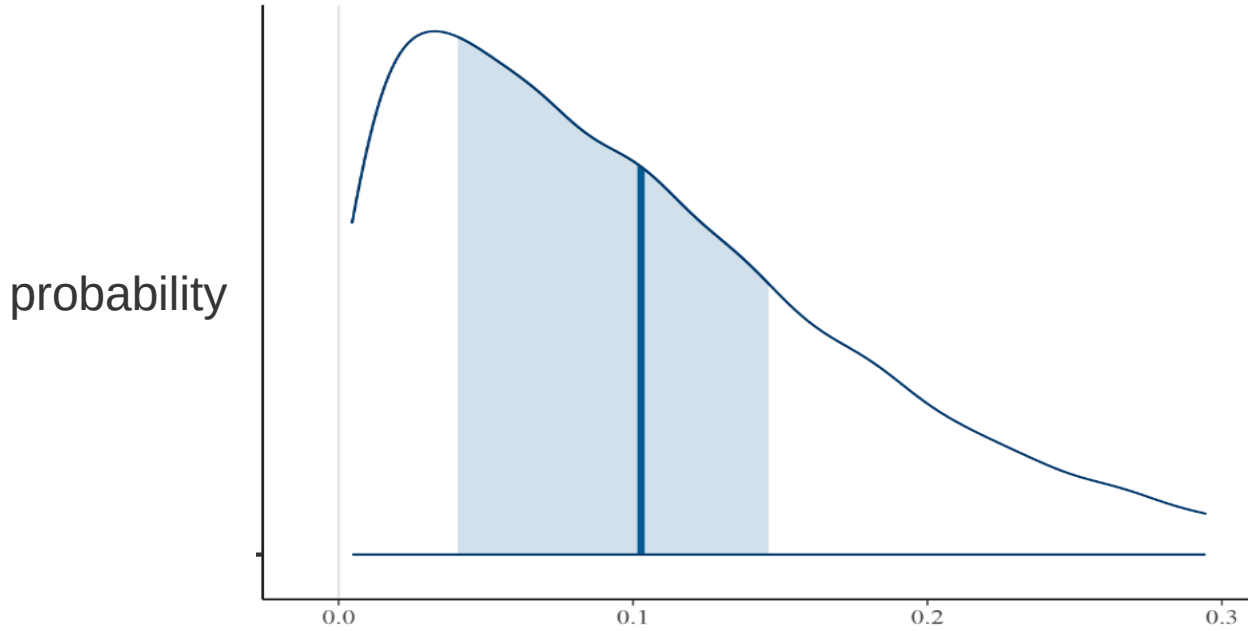
```
allnodes <- V(sysdimet_graph)
beta <- allnodes[allnodes$type=="beta"]
largest_typical_negative <- beta[order(beta$value),]
largest_typical_positive <- beta[order(-beta$value),]
```

… And for variances between persons

From	To	b_{σ^2}
energy	blood insulin	0.15997602
energy	LDL-cholesterol	0.10701078
energy	HDL-cholesterol	0.10680727
protein	**blood insulin**	**0.10266477**
pufa	blood insulin	0.10041444

```
allnodes <- V(sysdimet_graph)
b_sigma <- allnodes[allnodes$type=="b_sigma"]
largest_personal_variance <-
   b_sigma[order(-b_sigma$variance),]
```

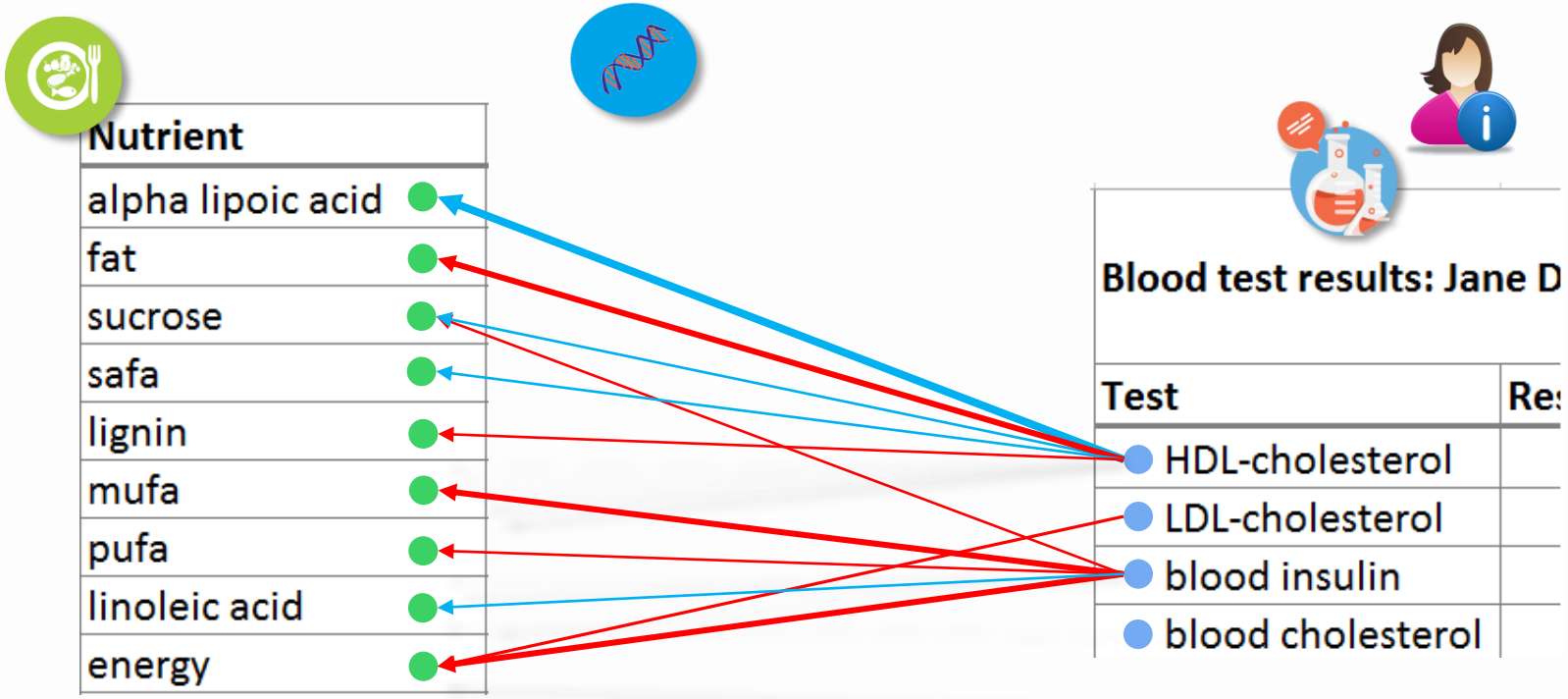
Querying the local distribution from Stan model: Personal variance in protein's effect to insulin



b_{σ^2}= 0.10266477

```
library("bayesplot")
fsins_posterior <- as.array(mebn.get_localfit("fsins"))
id <- match("prot", datadesc$Name)
mcmc_intervals(fsins_posterior, pars =
   c(paste0("sigma_b[",id,"]")), prob_outer = 0.95)..)
```

Future work: Personal graphs and diets



- With this model we can predict personal graphs for new patients
- Personal diet can be inferred by turning arrows at the personal graph and fixing the blood tests to their desirable levels
- Model is improved by adding the multi-response cases and ARMA
- Studying new dataset with kidney disease patients and new variables

Thank you!

For more details or collaboration, contact me at jari.turkia@cgi.com