University of Zurich[UZH]

http://r-bayesian-networks.org/

GILLES KRATZER, PHD

gilles.kratzer@gmail.com

PROF. REINHARD FURRER

reinhard.furrer@math.uzh.ch

USER! CONFERENCE, ONLINE 07.07.2021

ADVANCED BAYESIAN NETWORK MODELLING

ADVANCED METHODS WITH BN MODELING

Outline of the talk

- Mixed models – correction for grouped data
- Heuristic search
- MCMC over structures

- *Other advanced methods/features:*
  - *Scoring system*
  - *Tunable parameter prior*
  - *Structural prior*
  - *Data separation*
  - *Covariate adjustment*
  - *Likelihood contribution*

CORRECTION FOR CLUSTERING

Correction for grouped data

- The way the data were collected has a clear **grouping aspect**
- Then potential for **non-independence** between data points
- Lead to analyses which "are" **over-optimistic**
- As the **true level of variation** in the data is **under-estimated**
- Could impact study result … or not!
- **Good practice to check!**

In practice:

- Random effect
- **GLM** -> **GLMM** for each node
- Fit the DAG and check the posterior distribution (**widening**)
- If needed one can **incorporate random effect** in the **scoring** scheme

Pitfalls:

- Hugh computational complexity!

ADVANCED METHODS WITH BN MODELING

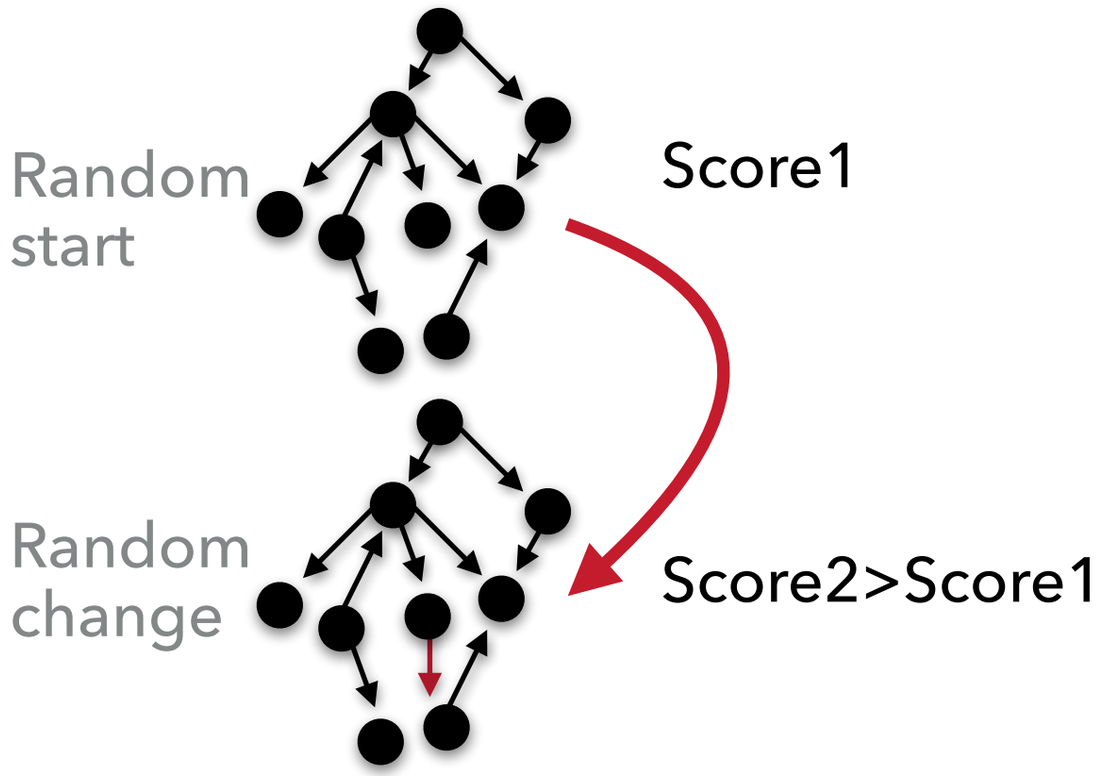
Find maximum a posteriori score

- ✓ Exact search
- ▸ Heuristic search
- ▸ MCMC over structures

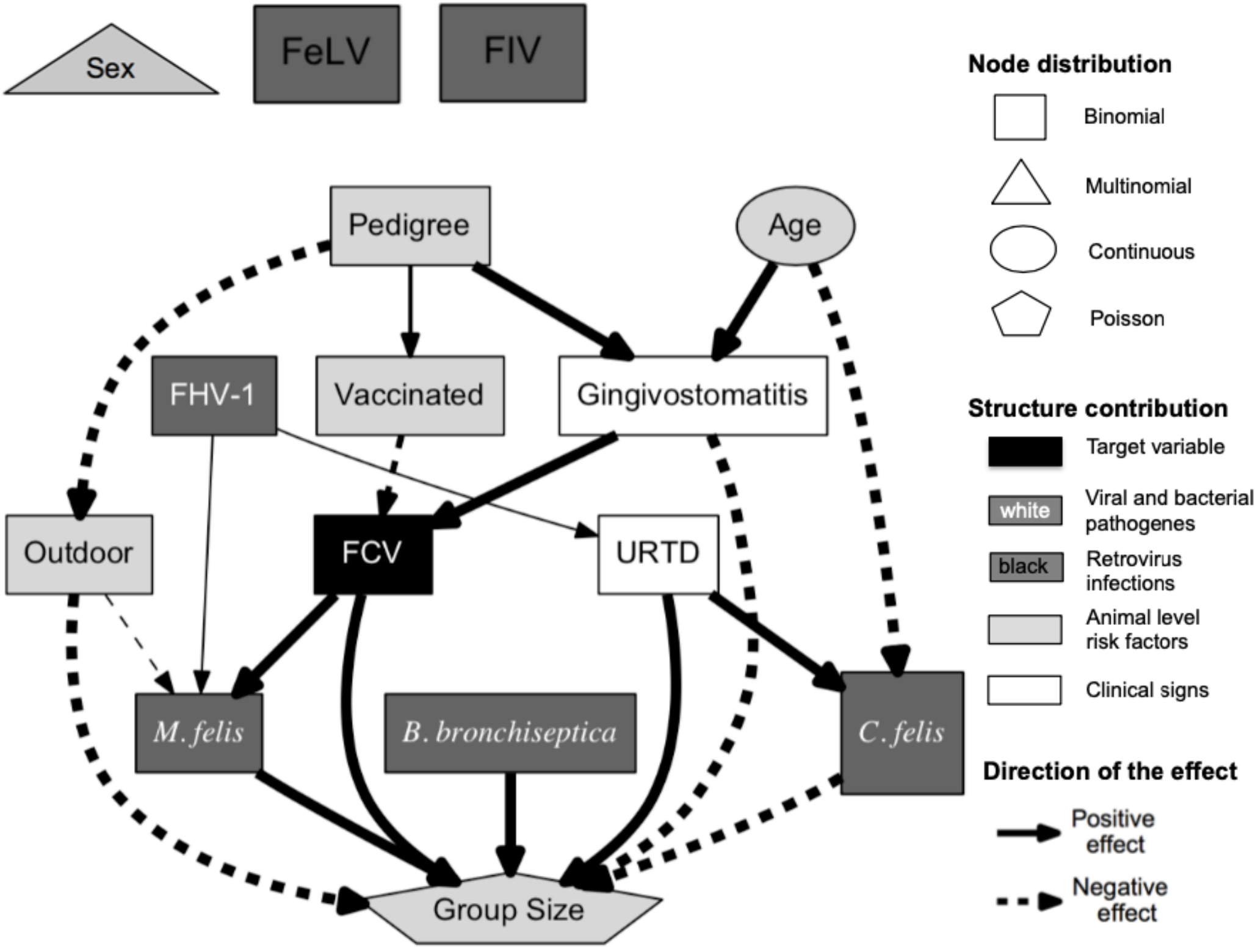
HEURISTIC SEARCH

Heuristic search: Greedy Hill-Climbing

- Simplest heuristic local search
  - Start with a given network
    - empty network
    - best tree
    - a random network
  - At each iteration
    - Evaluate all possible changes
      - Apply change that leads to best improvement in score
      - Reiterate
  - Stop when no modification improves score
- *Pitfalls*:
  - Local Maxima
  - Plateaus
- *Solution*:
  - Tabu
  - Random restart
  - Simulated annealing

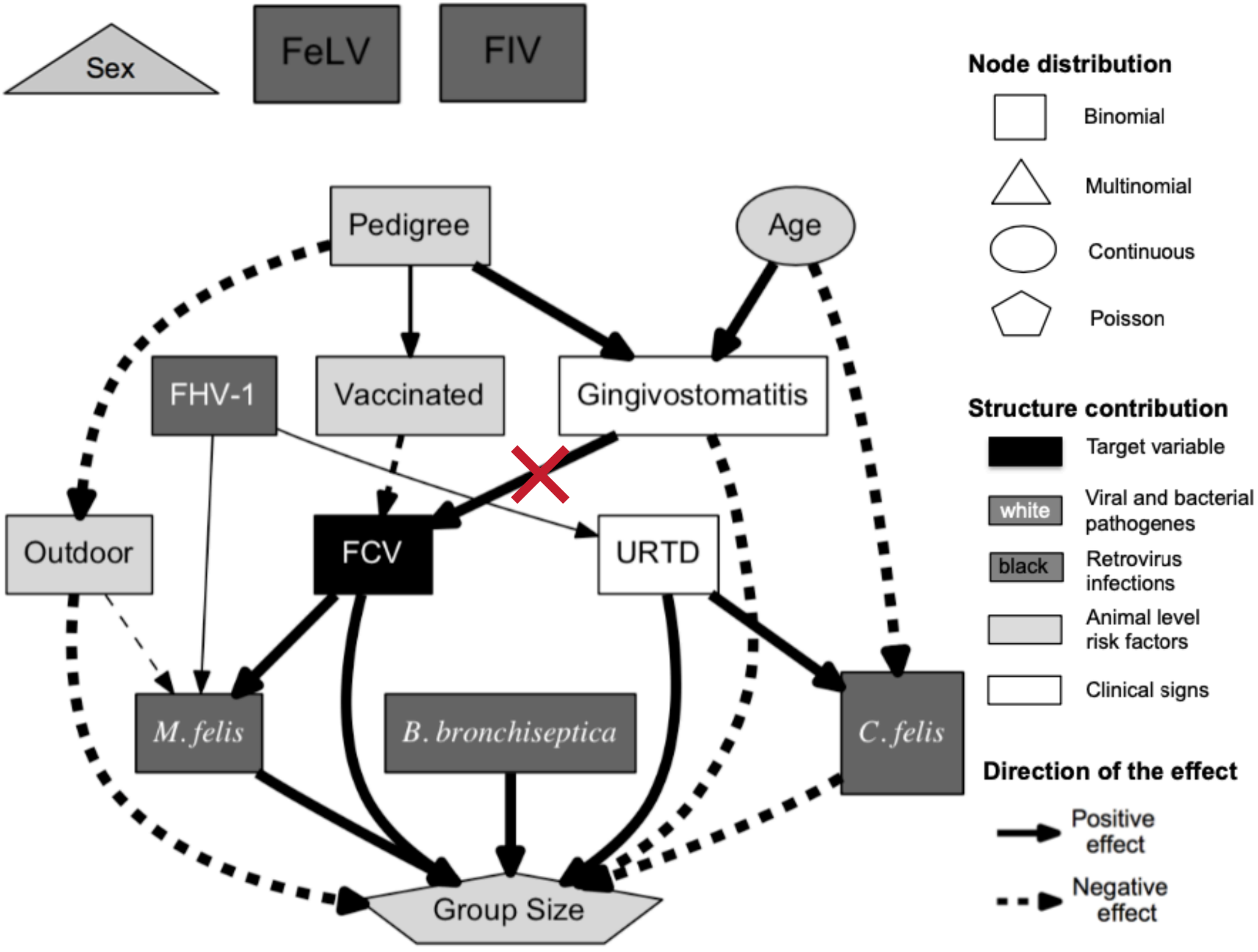


BAYESIAN NETWORK OF THE “CAT FLU“



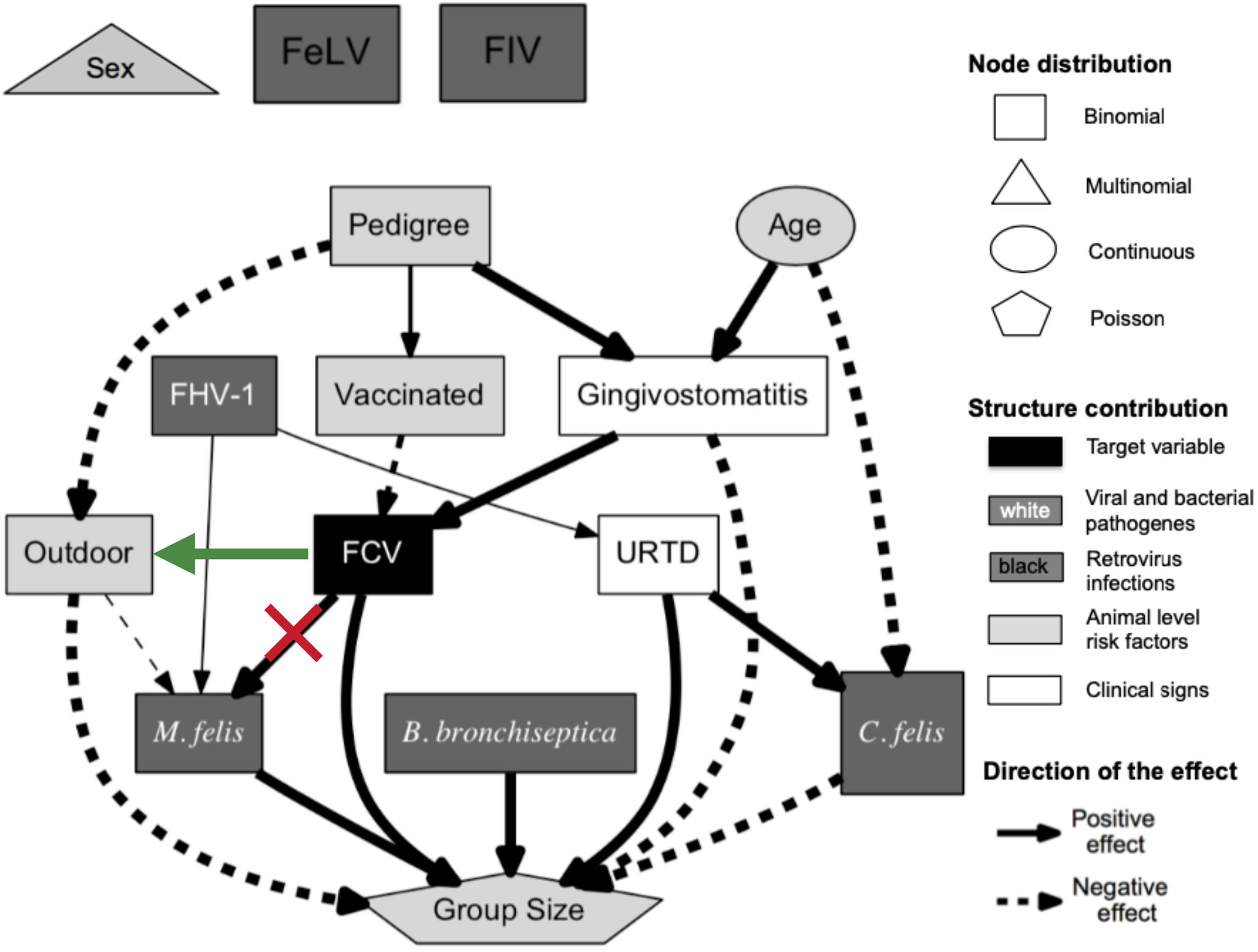
Kratzer and al. (2020) in Front. Vet. Sci.

BAYESIAN NETWORK OF THE “CAT FLU“



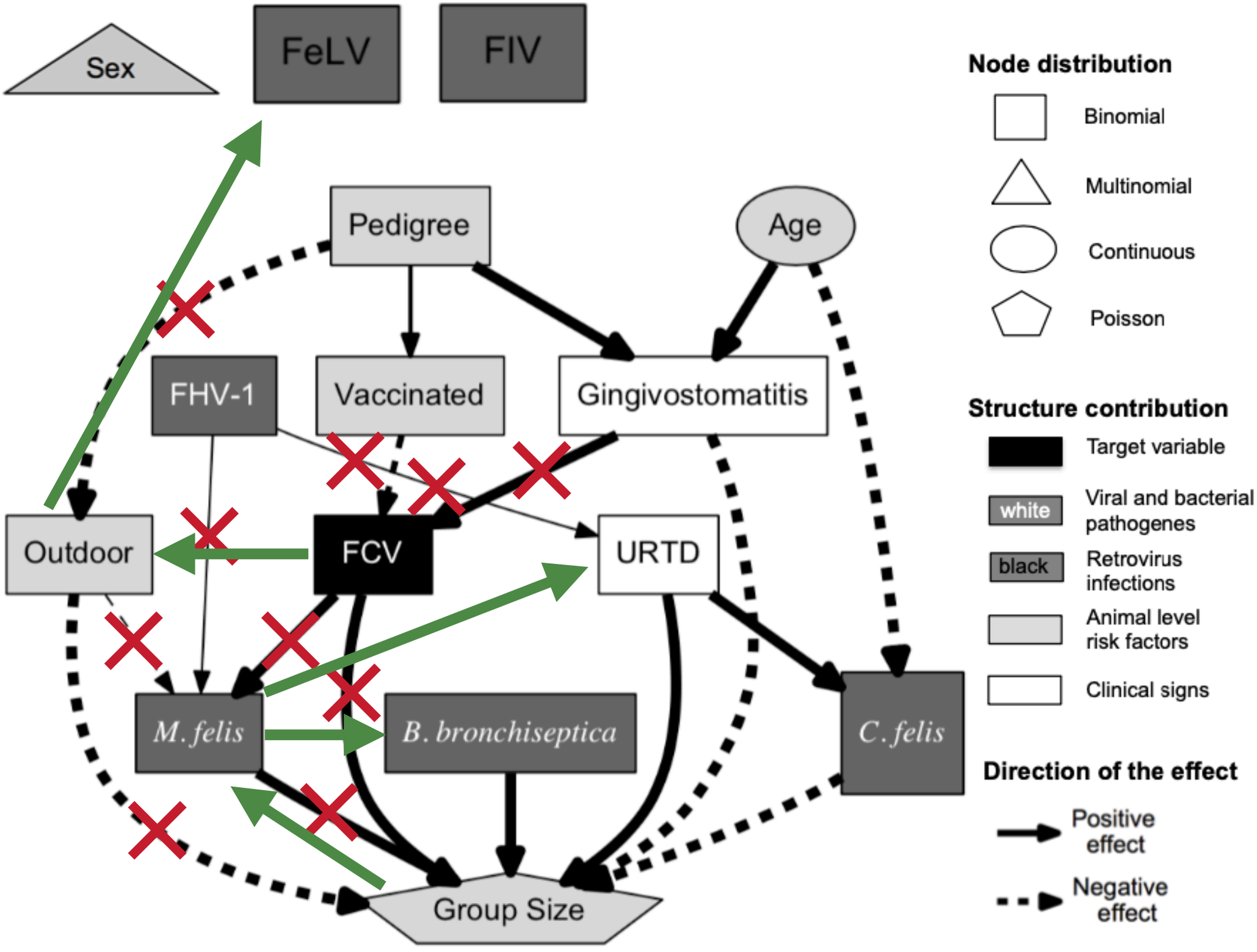
Kratzer and al. (2020) in Front. Vet. Sci.

BAYESIAN NETWORK OF THE “CAT FLU“



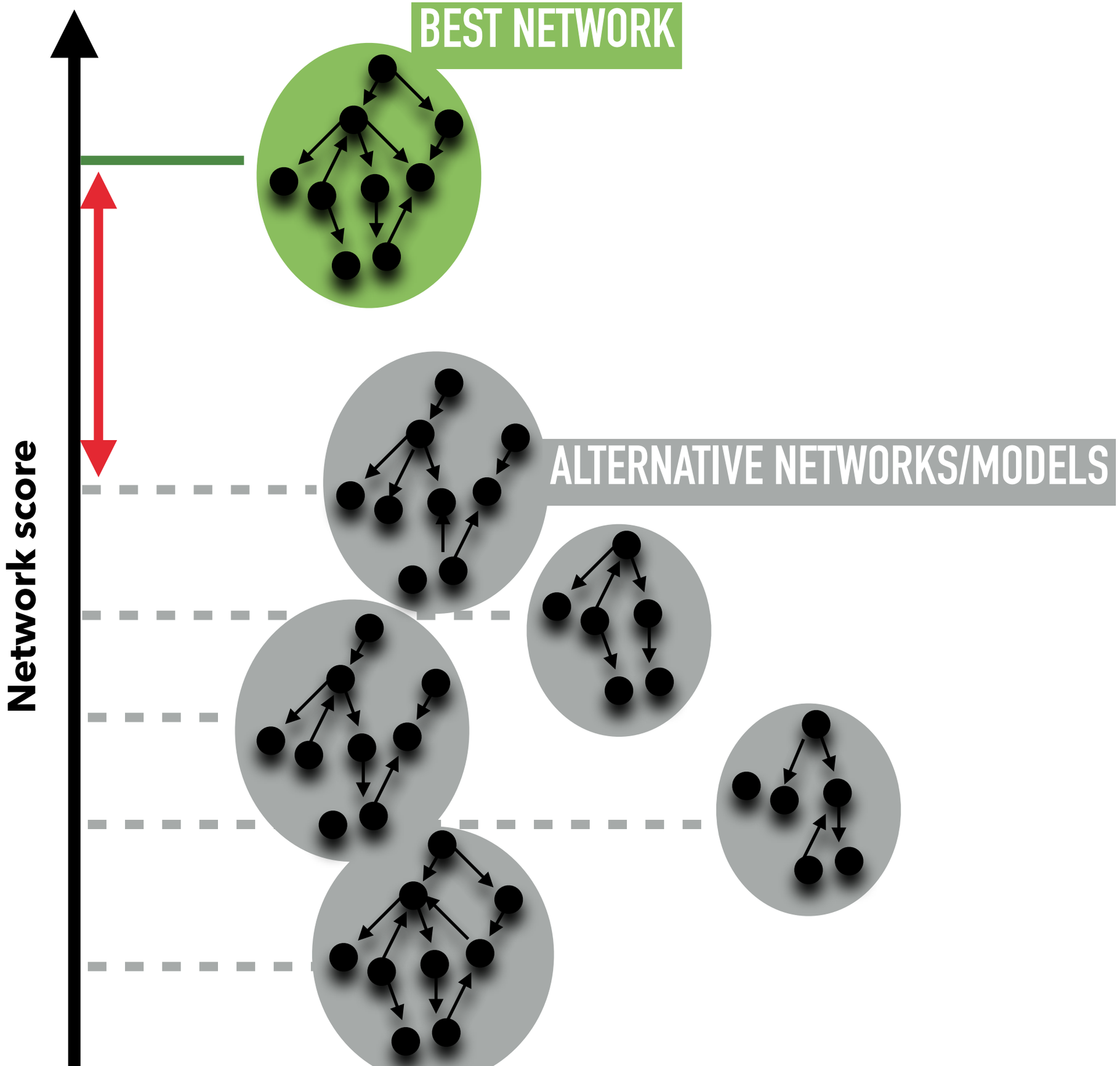
Kratzer and al. (2020) in Front. Vet. Sci.

BAYESIAN NETWORK OF THE “CAT FLU“

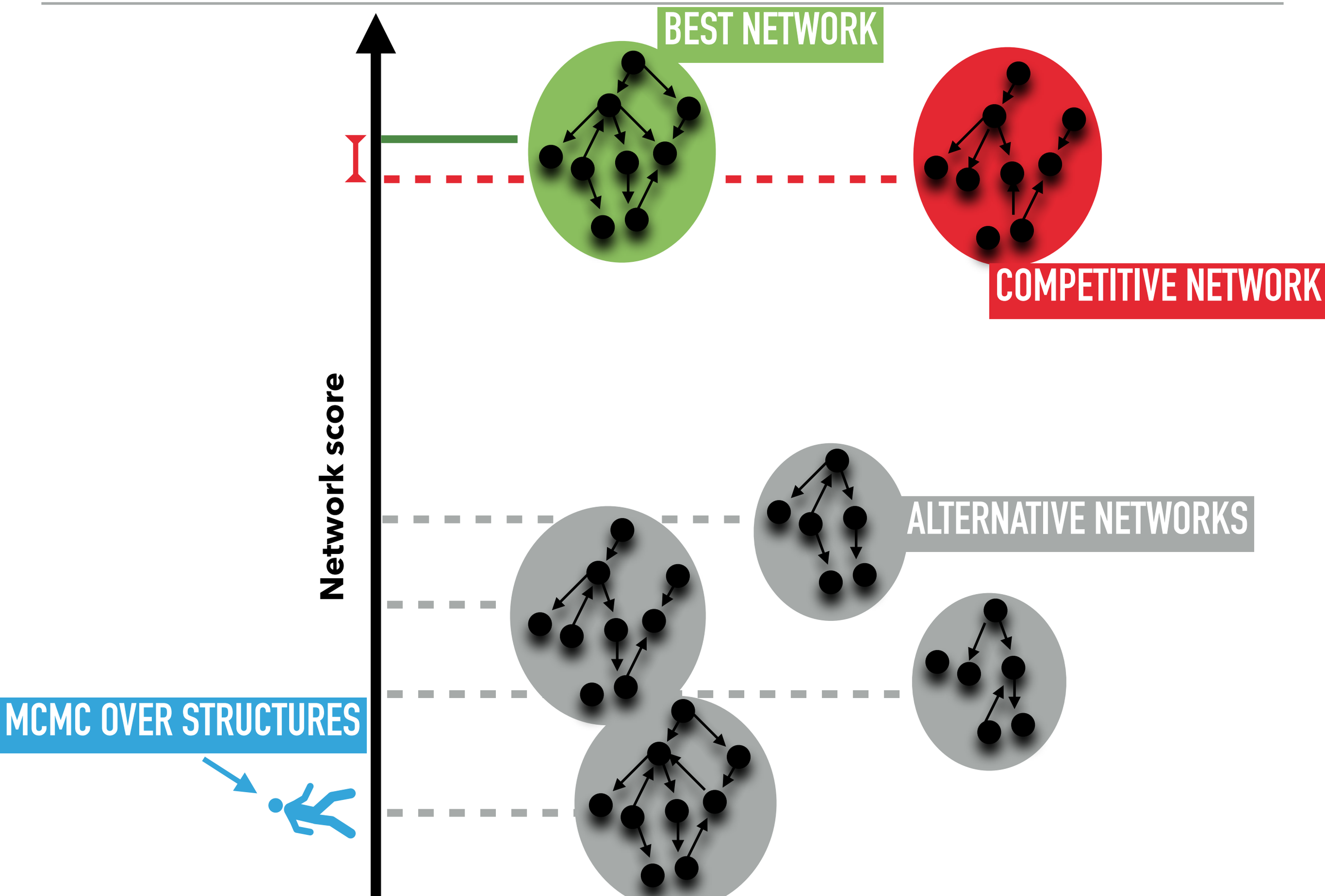


Kratzer and al. (2020) in Front. Vet. Sci.

MCMCABN - HOW TO CHOOSE THE BEST NETWORK?



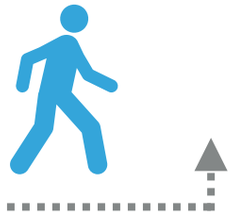
MCMCABN - HOW TO CHOOSE THE BEST NETWORK?

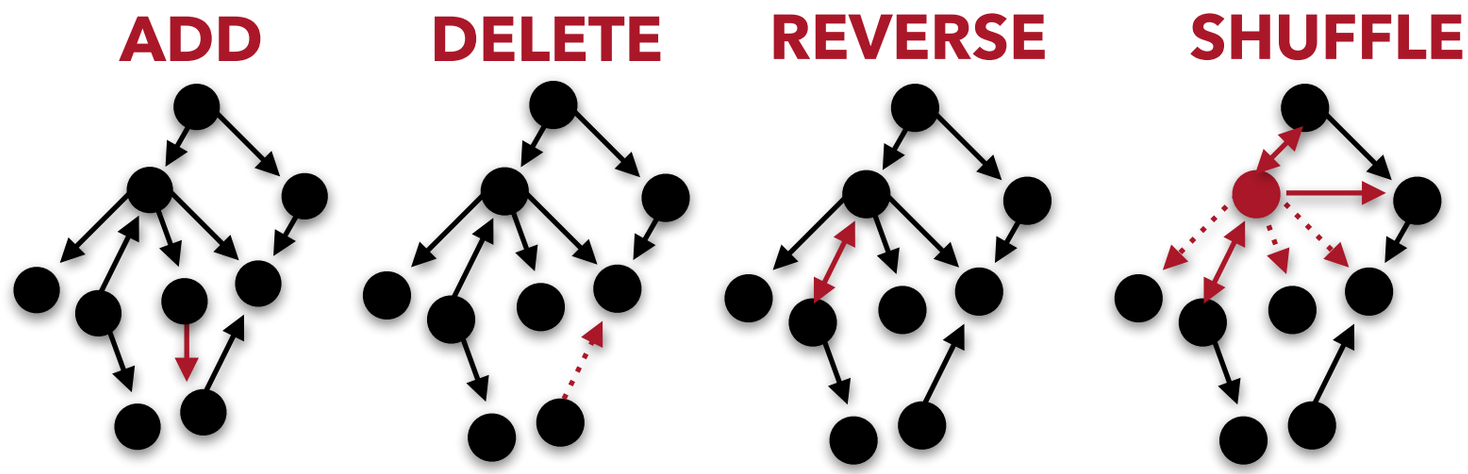


MCMC TO COMPUTE THE POSTERIOR DISTRIBUTION OF NETWORK

University of Zurich[UZH]

MCMC jumps =





Network score

Moderate downhill steps are usually accepted

Steep downhill steps are almost never accepted

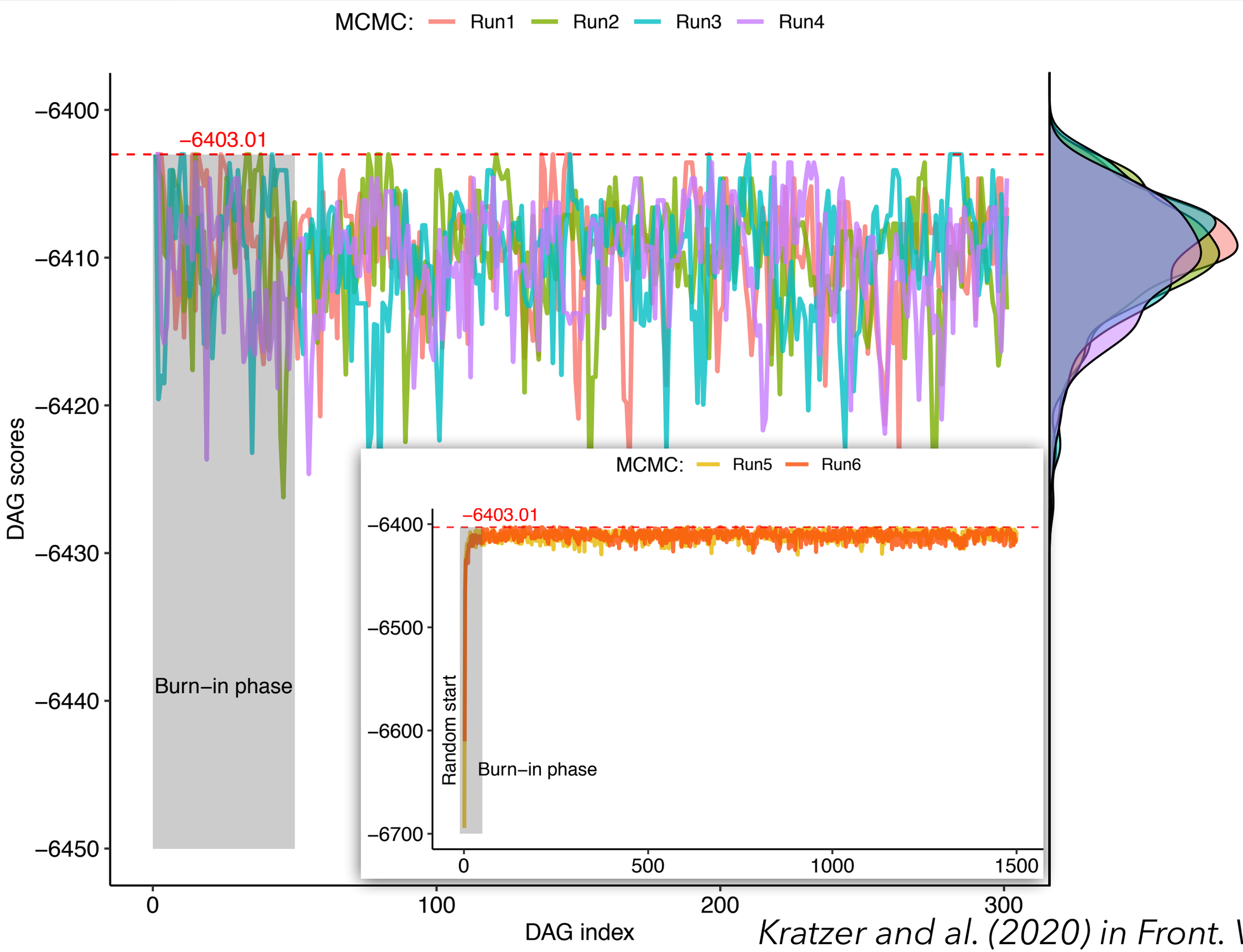
Uphill steps are always accepted

Following these rules, the workers tend to stay near the peak

Network index

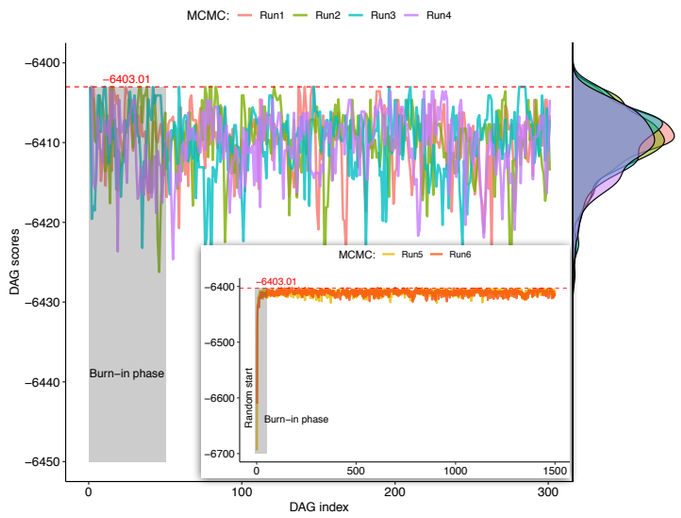
Adapted from Lewis (2011)

MCMC OVER STRUCTURE - POSTERIOR DISTRIBUTION



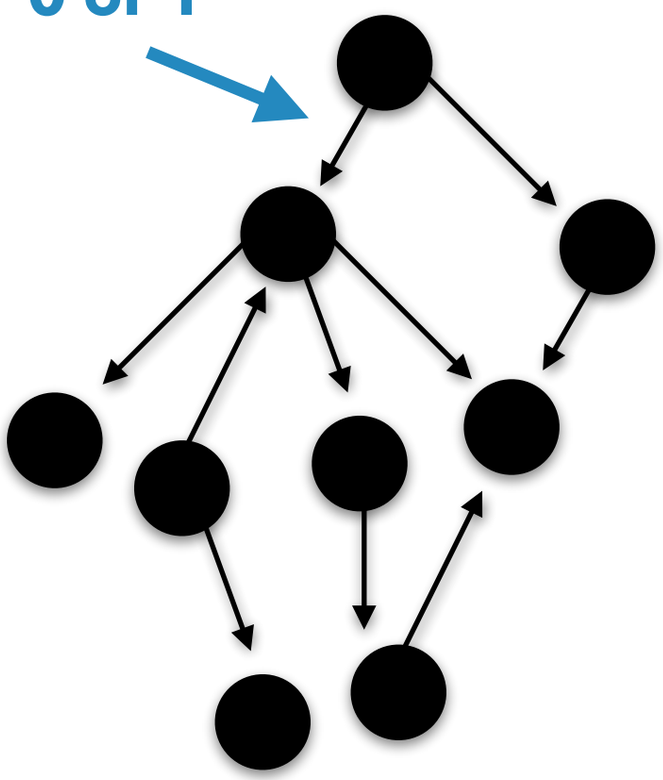
Kratzer and al. (2020) in Front. Vet. Sci.

WHY DO WE PERFORM MCMC OVER STRUCTURES?



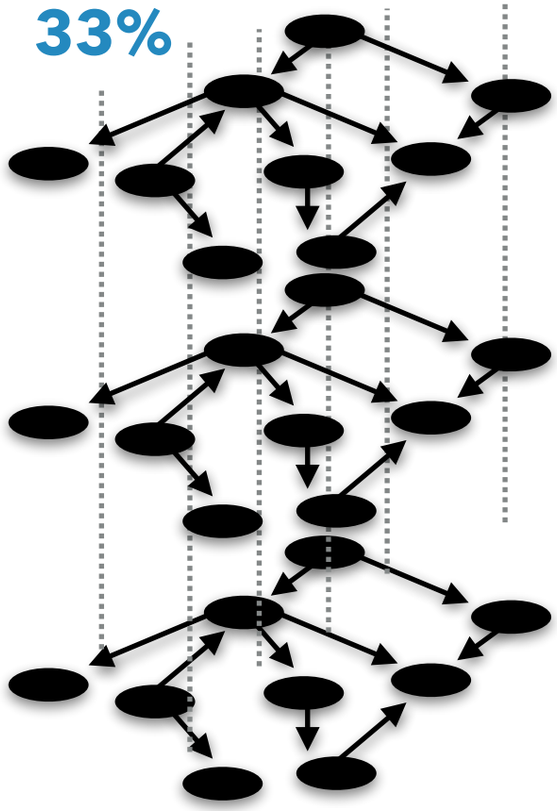
Best Unique Bayesian Network

0 or 1



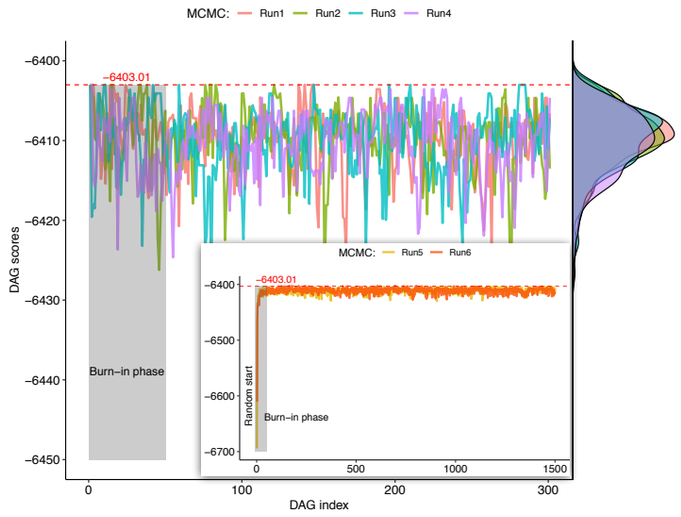
Counting prevalence of each ARC

33% ... 69% 73%



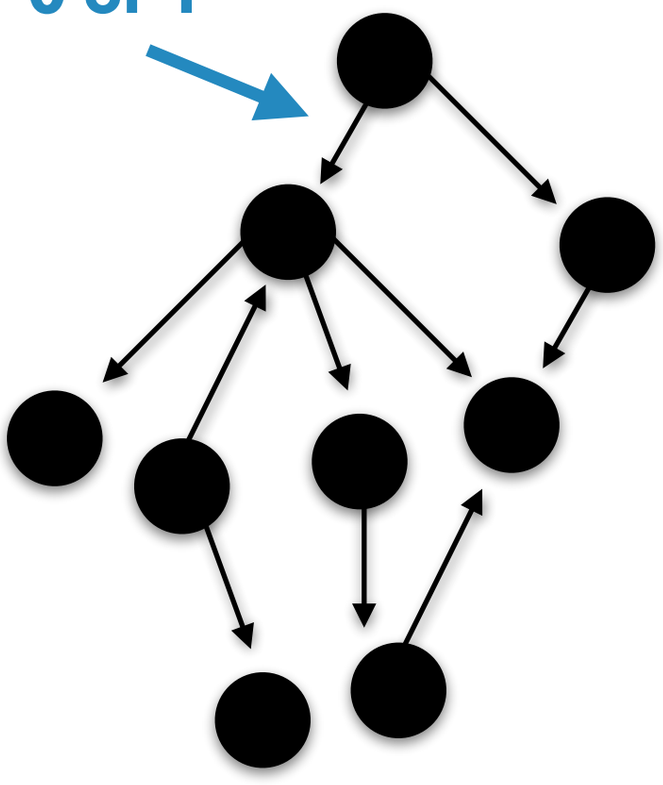
WHY DO WE PERFORM MCMC OVER STRUCTURES?

University of Zurich[UZH]



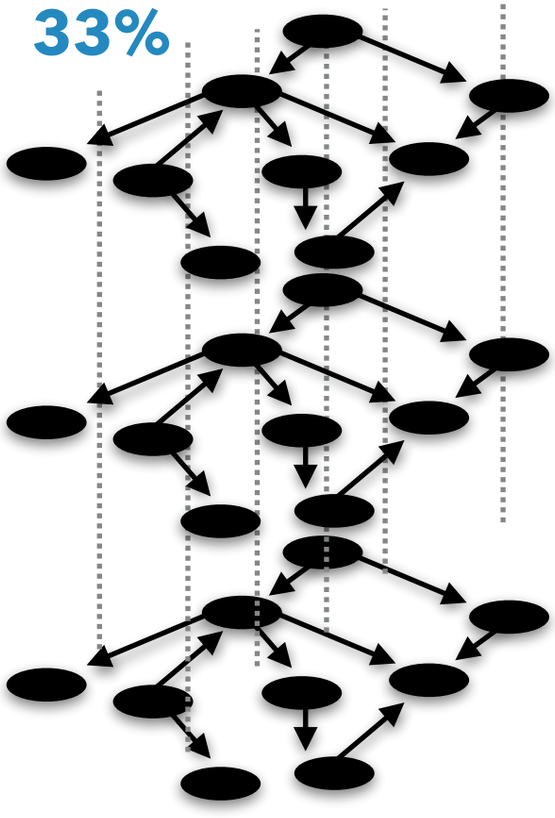
Best Unique Bayesian Network

0 or 1

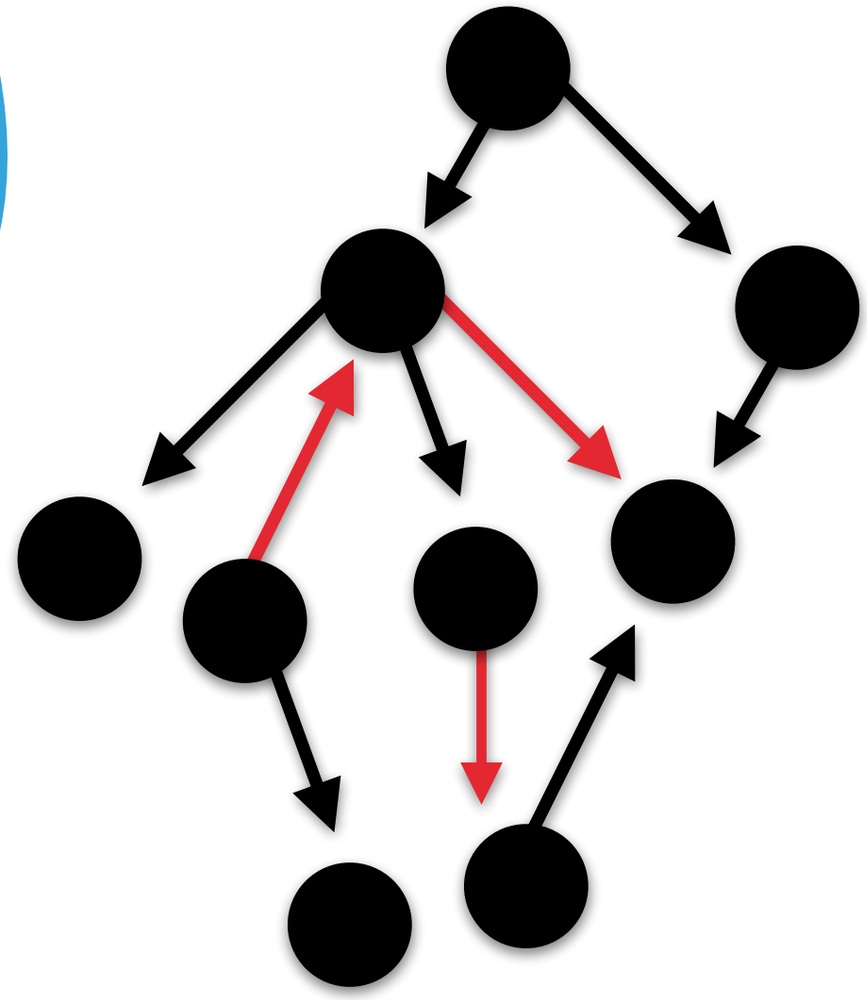


Counting prevalence of each ARC

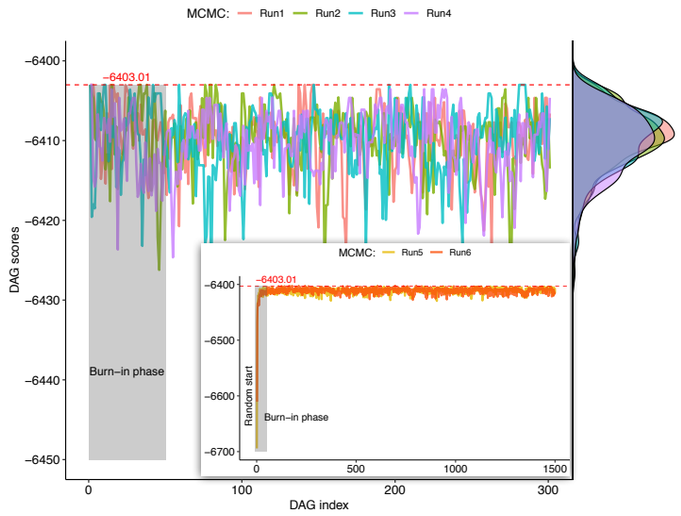
33% ... 69% 73%



Structural queries

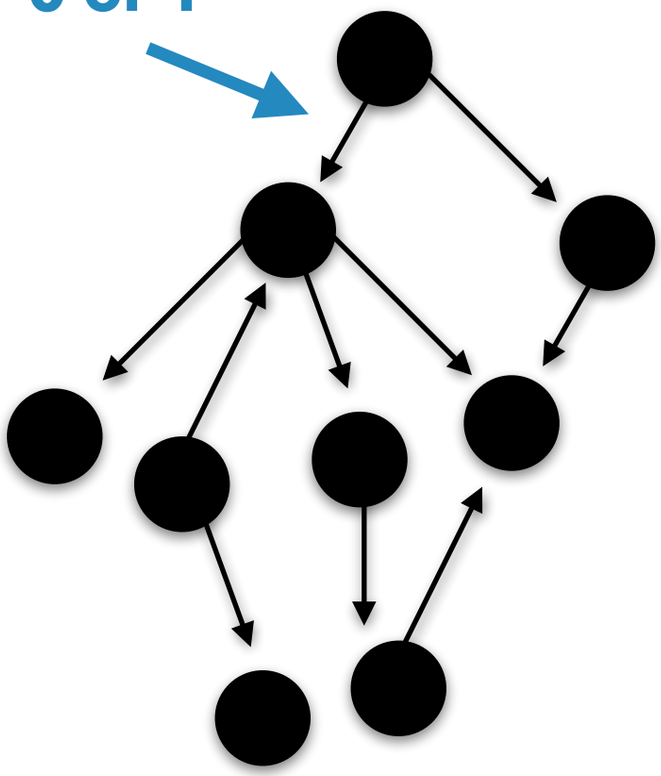


WHY DO WE PERFORM MCMC OVER STRUCTURES?



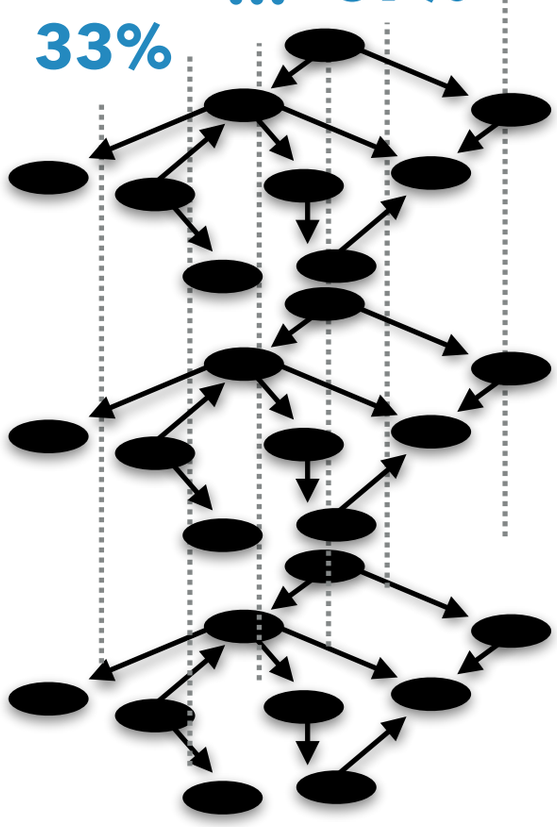
Best Unique Bayesian Network

0 or 1

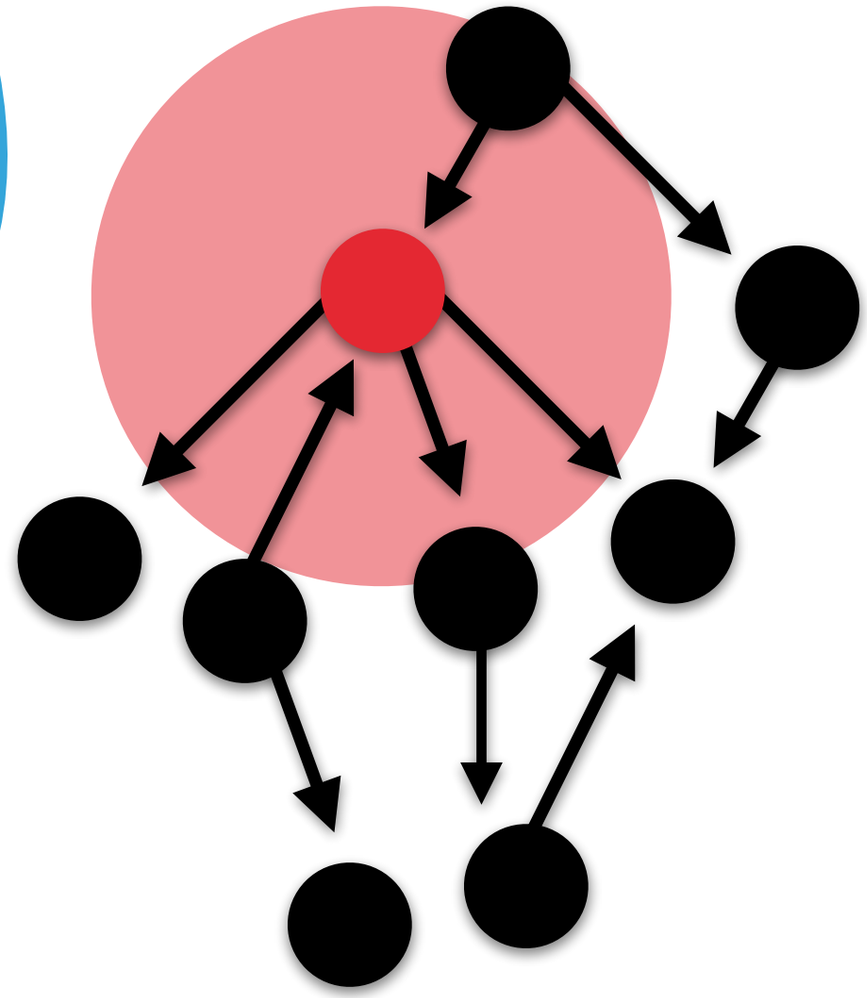


Counting prevalence of each ARC

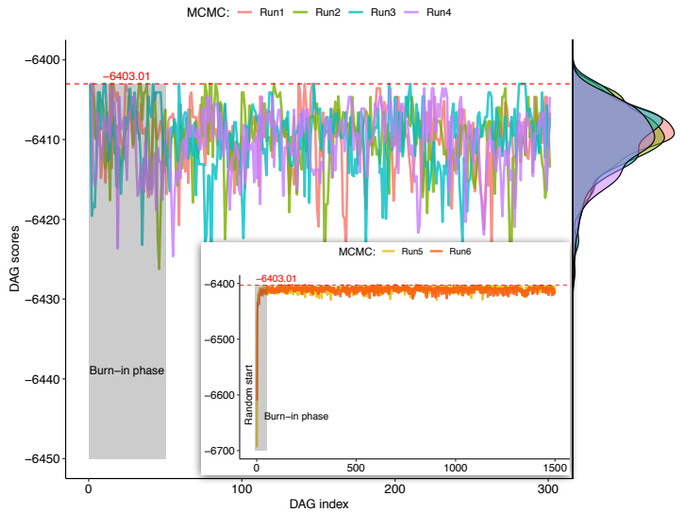
73%
69%
...
33%



Structural queries

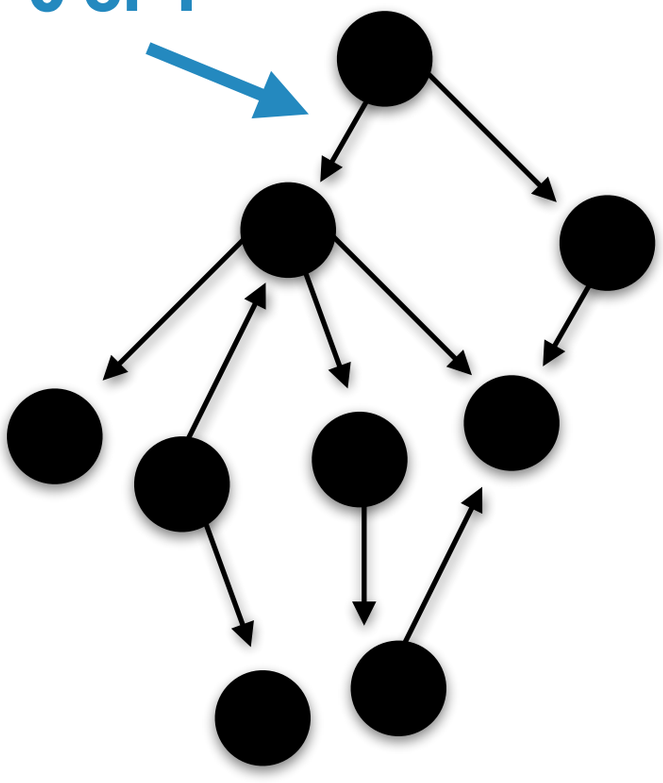


WHY DO WE PERFORM MCMC OVER STRUCTURES?



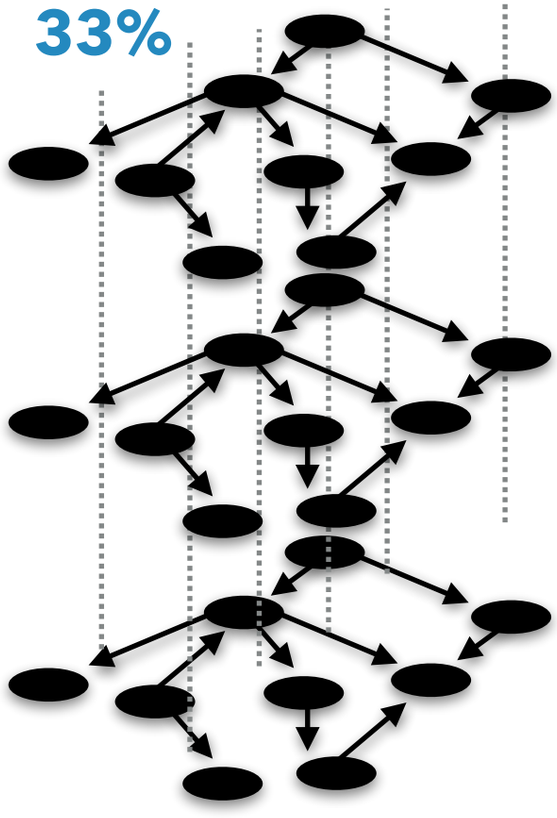
Best Unique Bayesian Network

0 or 1

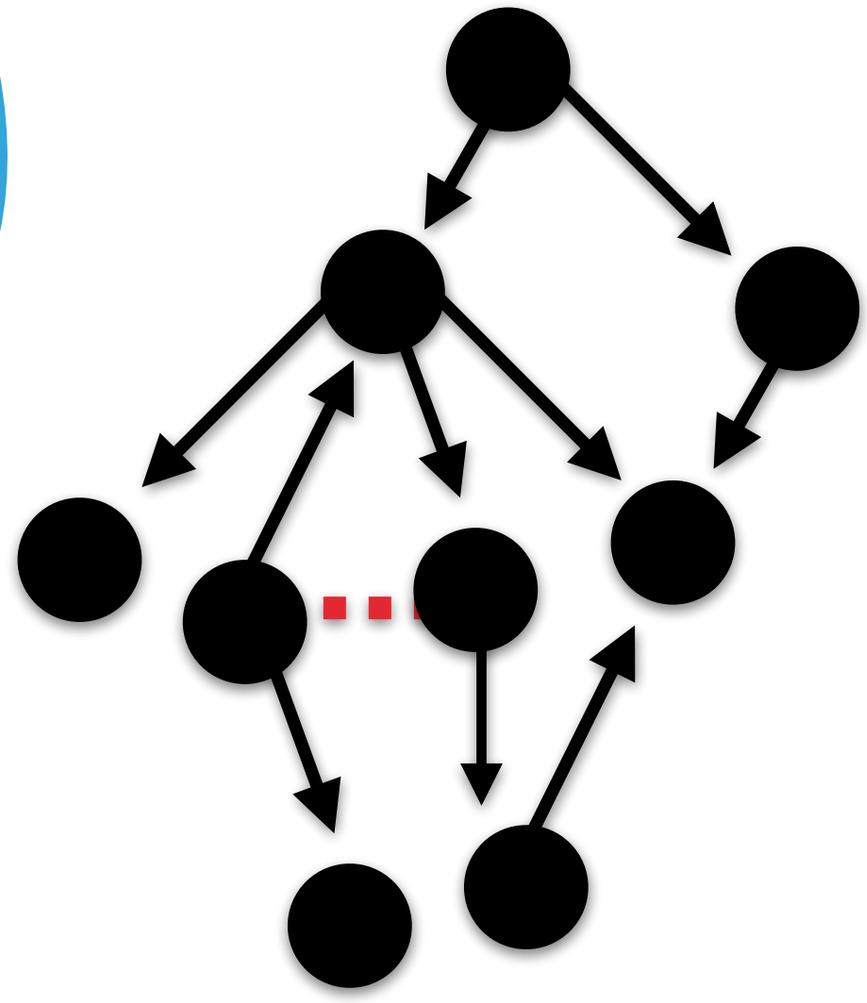


Counting prevalence of each ARC

33% ... 69% 73%



Structural queries



MCMC OVER STRUCTURES

MCMC over structures

- Selecting the most probable structure
- Controlling for overfitting
- Sampling the landscape of high scoring structures
  - In applied perspective avoid reducing the richness of BN modelling to only **one** structure
  - Quantify the marginal impact of relationships by marginalising out over structures

MCMC OVER STRUCTURES

MCMC over structures

- Selecting the most probable structure
- Controlling for overfitting
- Sampling the landscape of high scoring structures
  - In applied perspective avoid reducing the richness of BN modelling to only **one** structure
  - Quantify the marginal impact of relationships by marginalising out over structures

Question?

- Why in a typical regression analysis (GLM) based on observational data we usually present only the best fitting model?

HANDS-ON EXERCICES

HEURISTIC SEARCH

```
num.searches <- 200
max.steps <- 150
heur.res <- quiet(search.heuristic(score.cache = mycache,
                                   score = "mlik",
                                   data.dists = dist,
                                   max.parents = 4,
                                   start.dag = "random",
                                   num.searches = num.searches,
                                   max.steps = max.steps,
                                   seed = 3213,
                                   verbose = TRUE,
                                   algo = "hc"))
```