Evaluating a Bayesian Network using "Sensitivity to Findings": How useful is it?

Ann E. Nicholson, David Albrecht,
Lucas Azzola, Michael Gill and Stuart Lloyd
Monash University

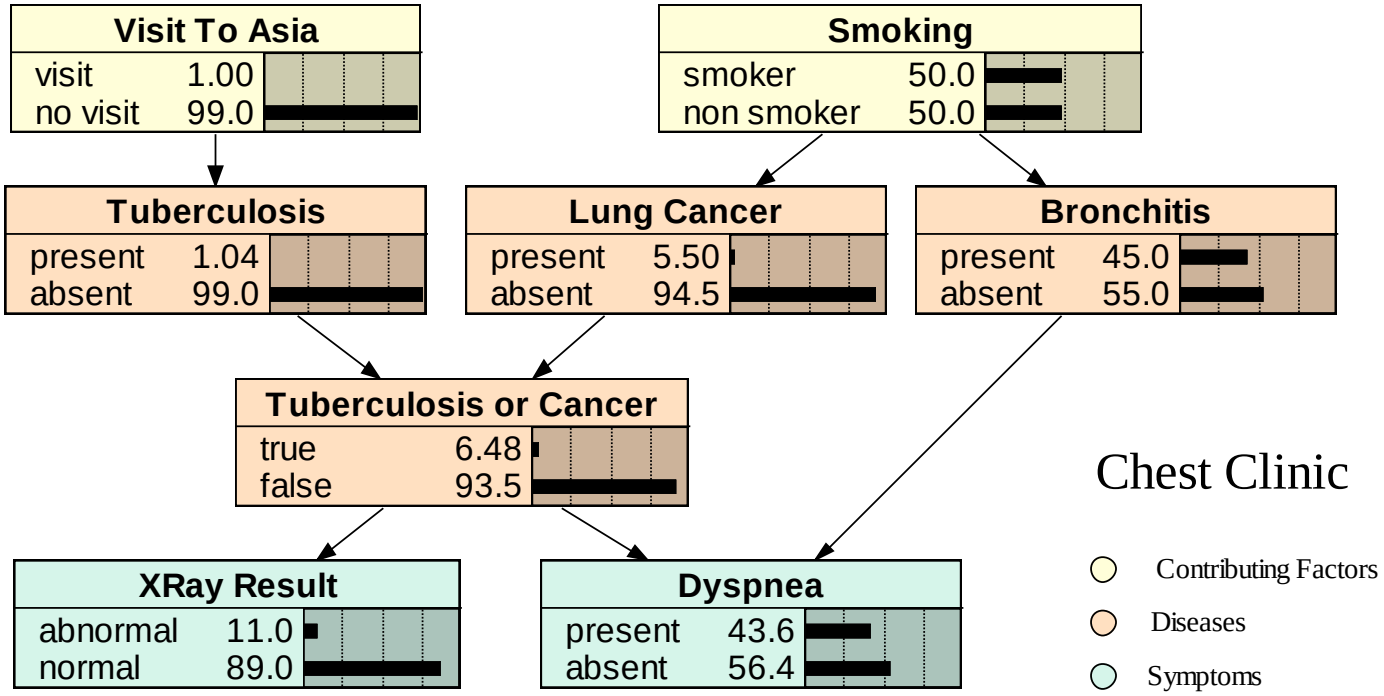
Overview

- Sensitivity Analysis
- Mutual Information
- A Simple Metric
- Examples

Sensitivity Analysis

- Evidence
- Parameters
- Structure
- Decisions

Chest Clinic Network

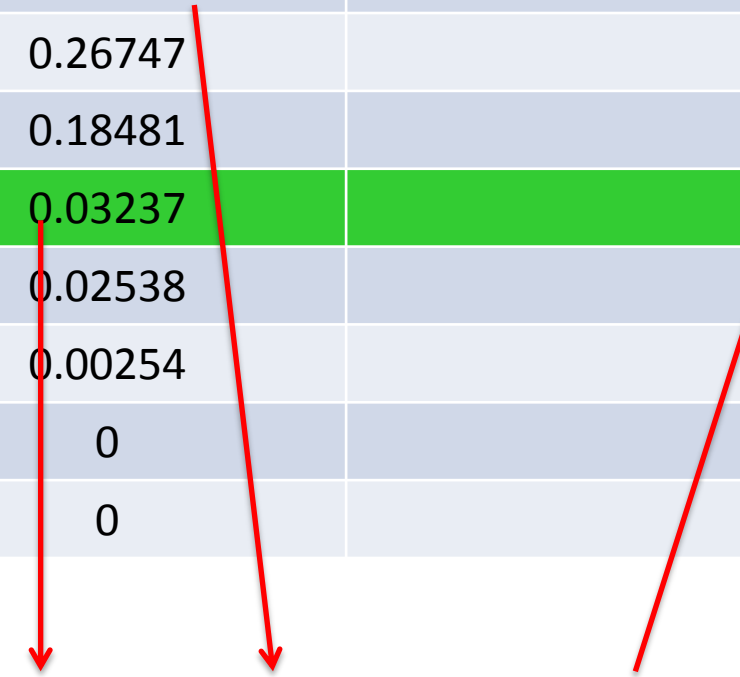


Sensitivity of "Lung Cancer"

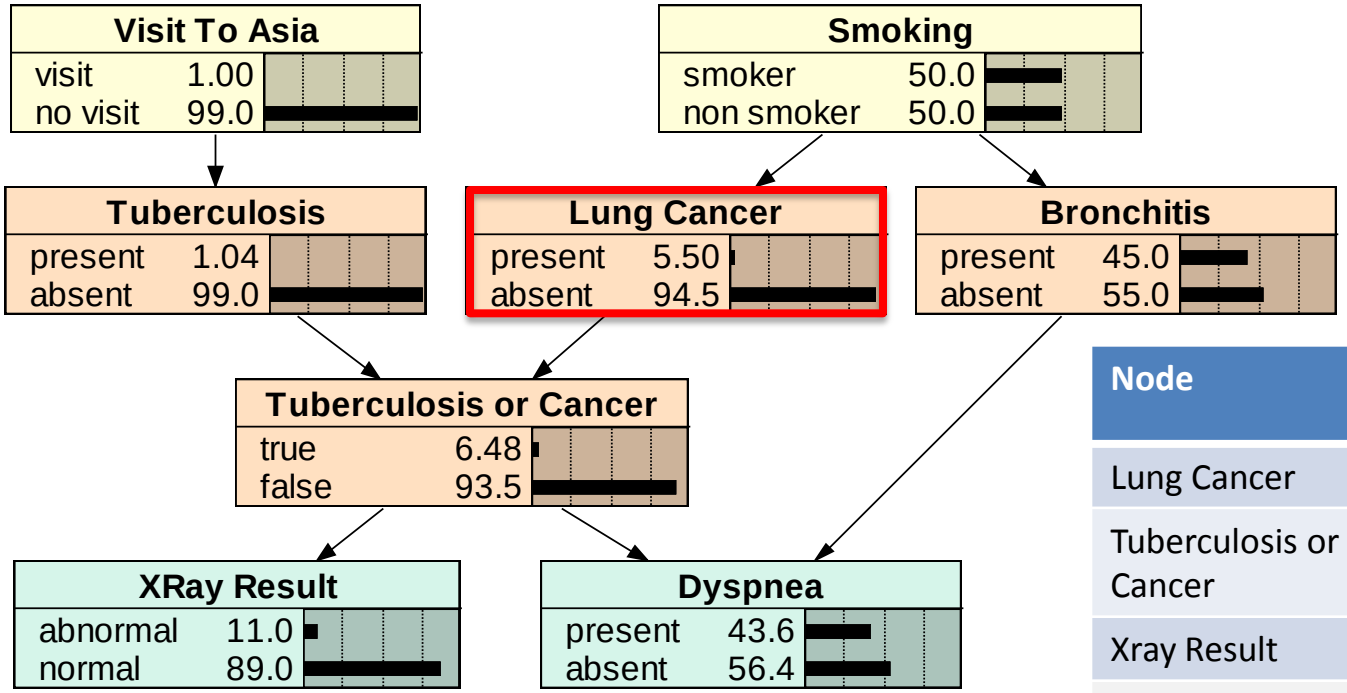
Node	Mutual Information	Percentage
Lung Cancer	0.30727	100
Tuberculosis or Cancer	0.26747	87
Xray Result	0.18481	60.1
Smoking	0.03237	10.5
Dyspnea	0.02538	8.26
Bronchitis	0.00254	0.827
Visit to Asia	0	0
Tuberculosis	0	0

Sensitivity of "Lung Cancer"

Node	Mutual Information	Percentage
Lung Cancer	0.30727	100
Tuberculosis or Cancer	0.26747	87
Xray Result	0.18481	60.1
Smoking	0.03237	10.5
Dyspnea	0.02538	8.26
Bronchitis	0.00254	0.827
Visit to Asia	0	0
Tuberculosis	0	0



Sensitivity of "Lung Cancer"



Entropy

Node	MI
Lung Cancer	0.30727
Tuberculosis or Cancer	0.26747
Xray Result	0.18481
Smoking	0.03237
Dyspnea	0.02538
Bronchitis	0.00254
Visit to Asia	0
Tuberculosis	0

Independent of Lung Cancer

Mutual Information

X

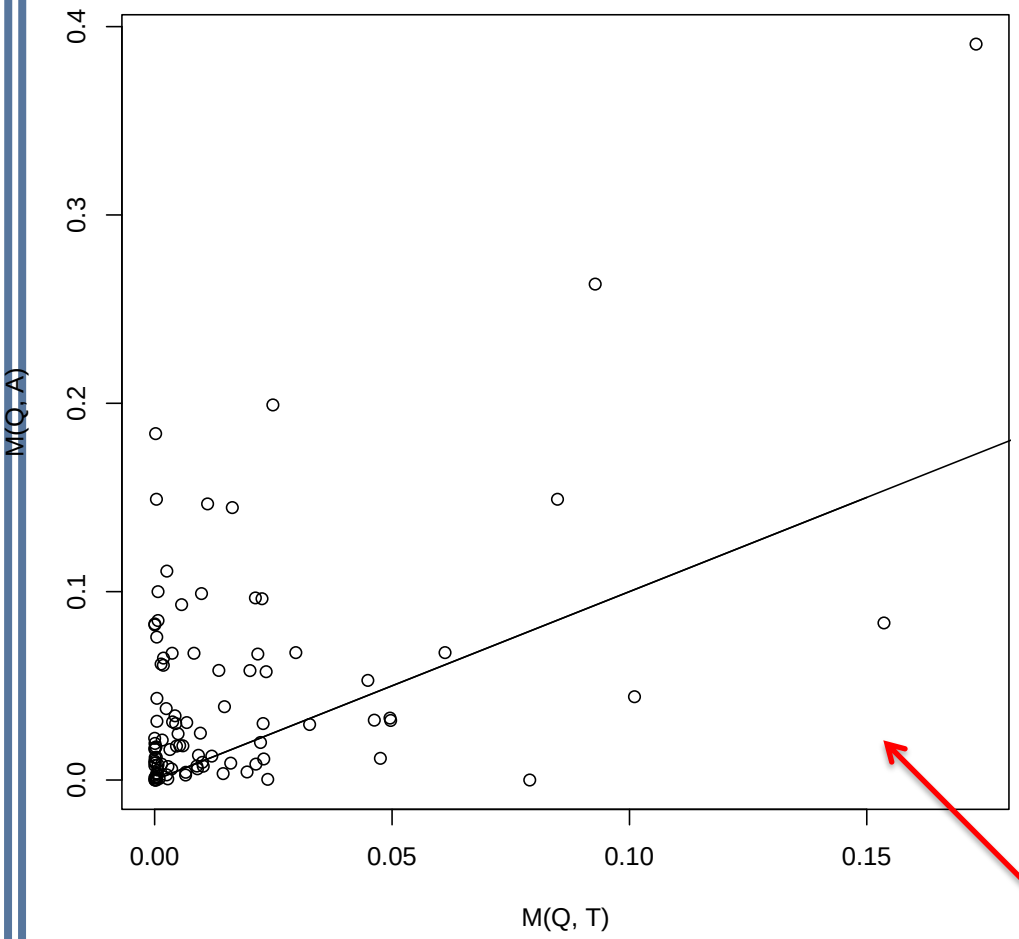
Y

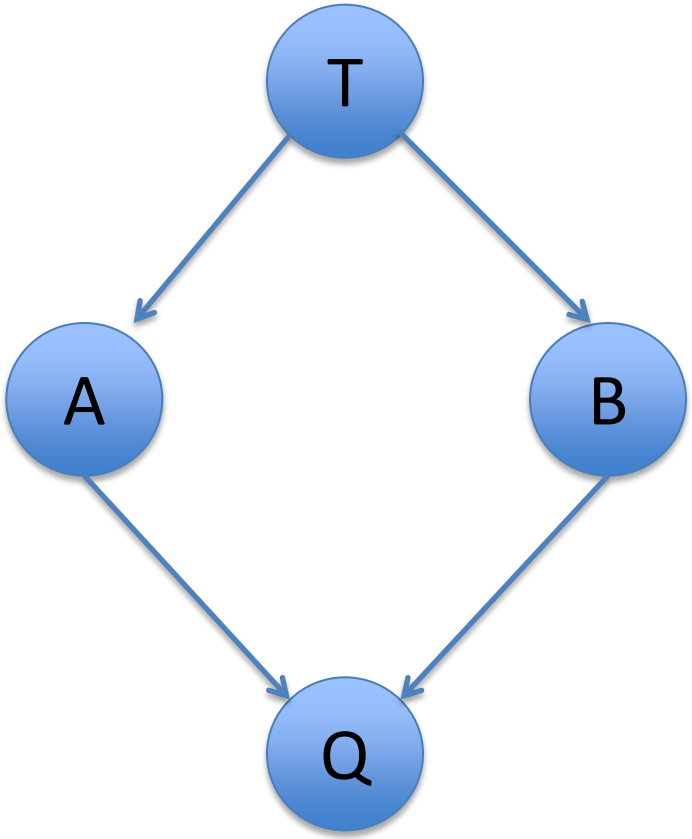
	x_1	…	x_n
y_1	$P(x_1,y_1)$	…	$P(x_n,y_1)$
:	:		:
y_m	$P(x_1,y_m)$	…	$P(x_n,y_m)$

$$M(X,Y)=\sum_{i=1}^{n}\sum_{j=1}^{m}P(x_i,y_j)\log\left(\frac{P(x_i)P(y_j)}{P(x_i,y_j)}\right)$$

Some Properties

- $M(X, Y) = M(Y, X)$
- $0 \leq M(X, Y) \leq M(X, X)$
- If $X \rightarrow Y \rightarrow Z$ then $M(X, Y) \geq M(X, Z)$
- If $X \leftarrow Y \rightarrow Z$ then $M(X, Y) \geq M(X, Z)$
- If $X \rightarrow Y \leftarrow Z$ then $M(X, Y) \geq M(X, Z)$





$M(Q,A) \leq M(Q,T)$

Structure Metrics

- Measure the component of sensitivity of evidence that depends on structure.
- Depends on the number of:
  - Nodes in a path
  - Paths between nodes
- Investigated a simple metric:
  - Distance Weighted Influence (DWI)

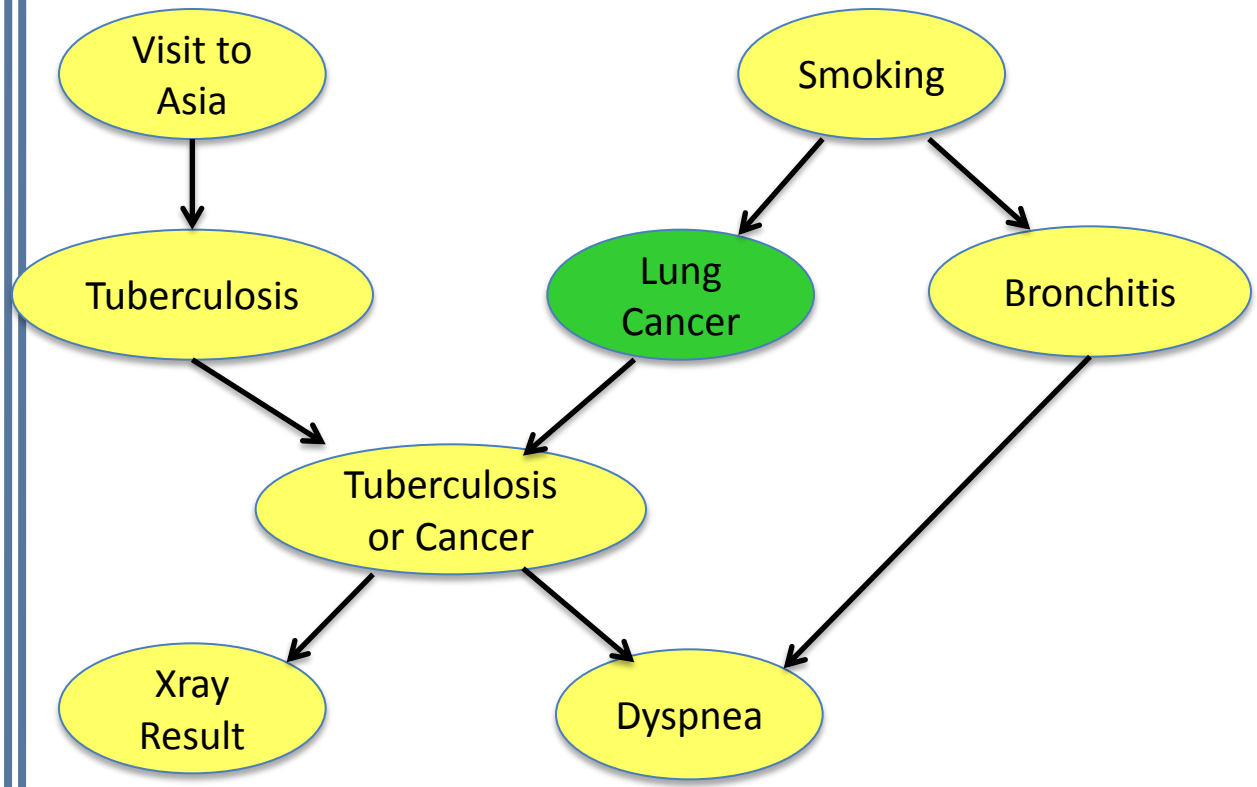
Distance Weighted Influence (DWI)

$$I(Q,T) = \sum_{path} I(path),$$

where the summation is over all unblocked paths between Q and T, and

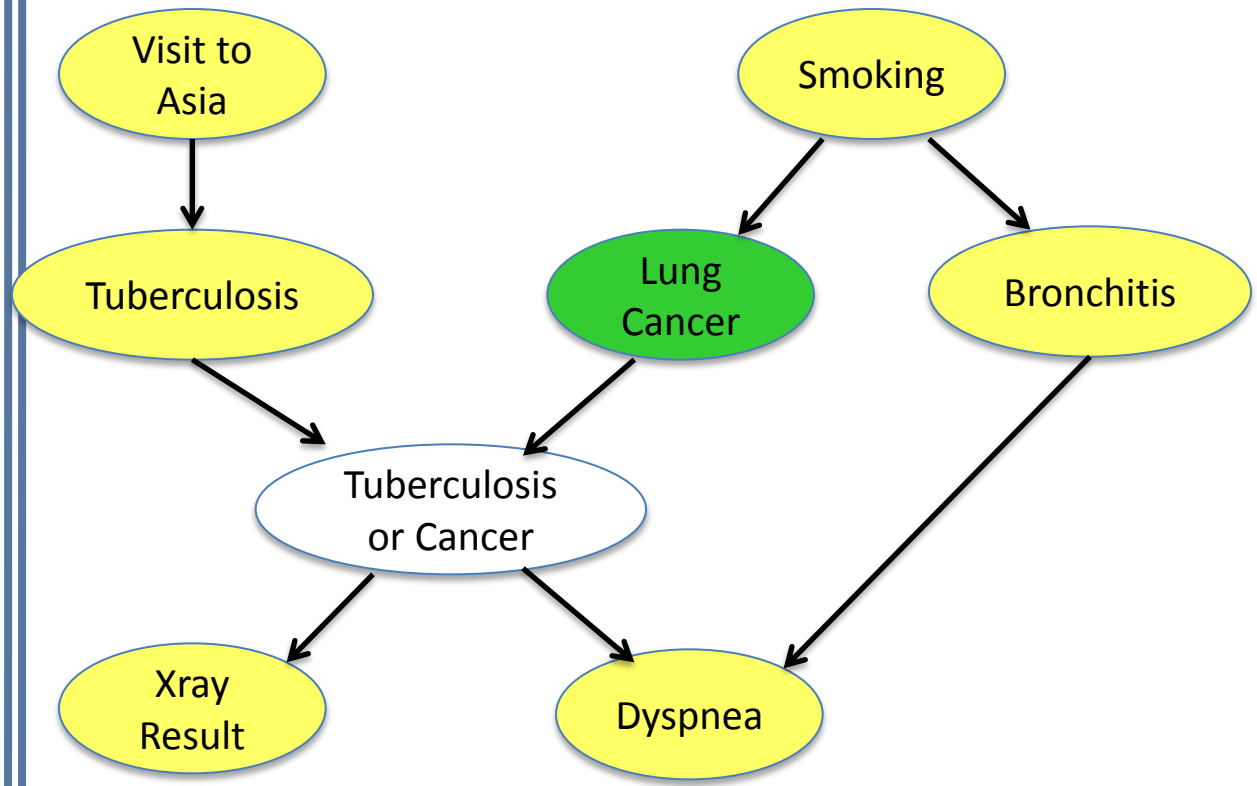
$$I(path) = w^{length(path)}$$

Clinic Example



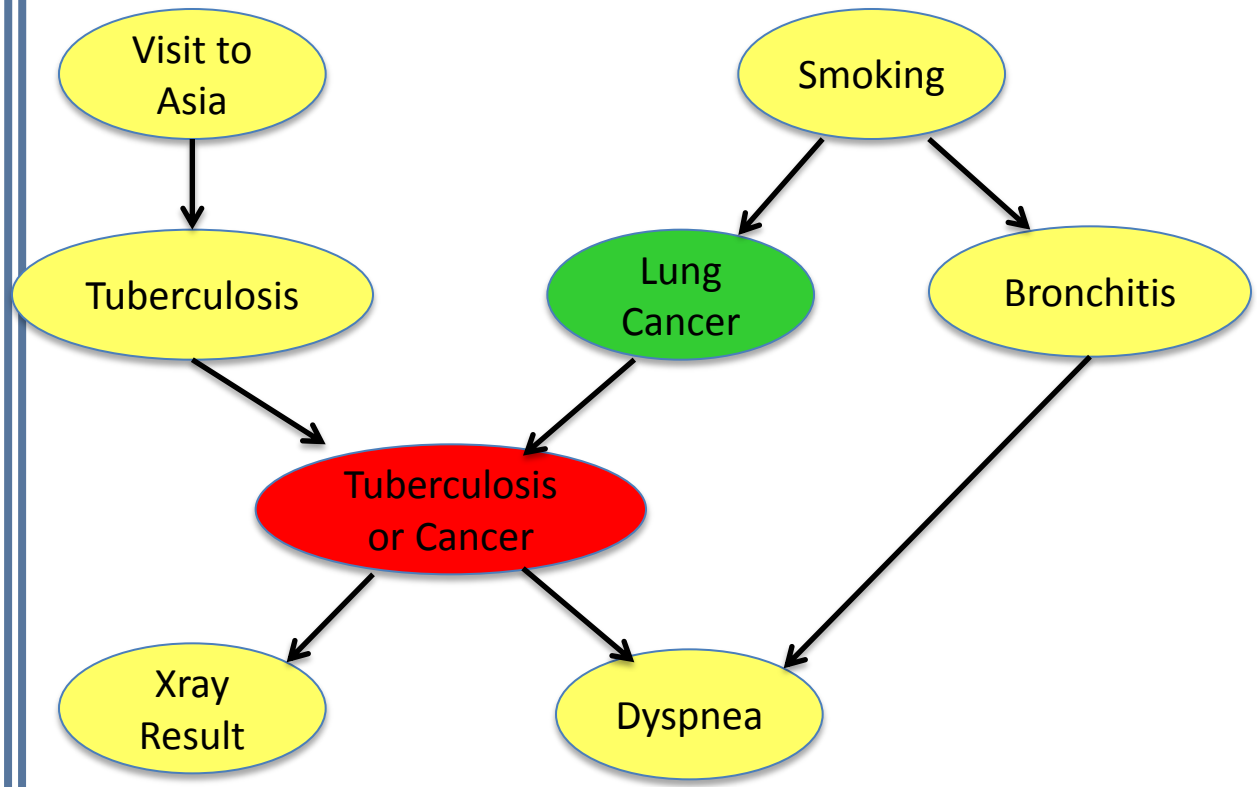
Node	DWI	W=0.5
Lung Cancer	1	1
Tuberculosis or Cancer	w	0.5
Xray Result	w^2	0.25
Smoking	w	0.5
Dyspnea	w^2+w^3	0.375
Bronchitis	w^2	0.25
Visit to Asia	0	0
Tuberculosis	0	0

Clinic Example with Determistic Nodes



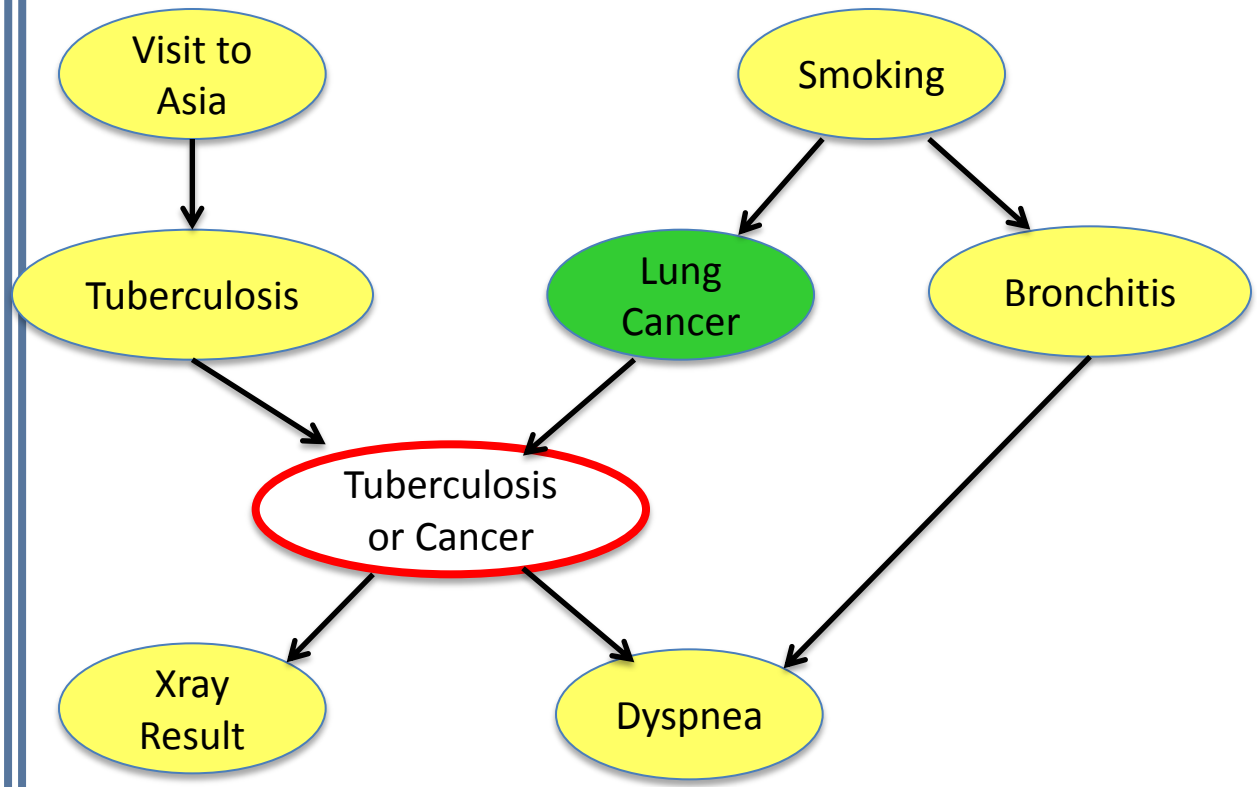
Node	DWI	W=0.5
Lung Cancer	1	1
Tuberculosis or Cancer	1	1
Xray Result	w	0.5
Smoking	w	0.5
Dyspnea	$w+w^3$	0.625
Bronchitis	w^2	0.25
Visit to Asia	0	0
Tuberculosis	0	0

Clinic Example with Evidence

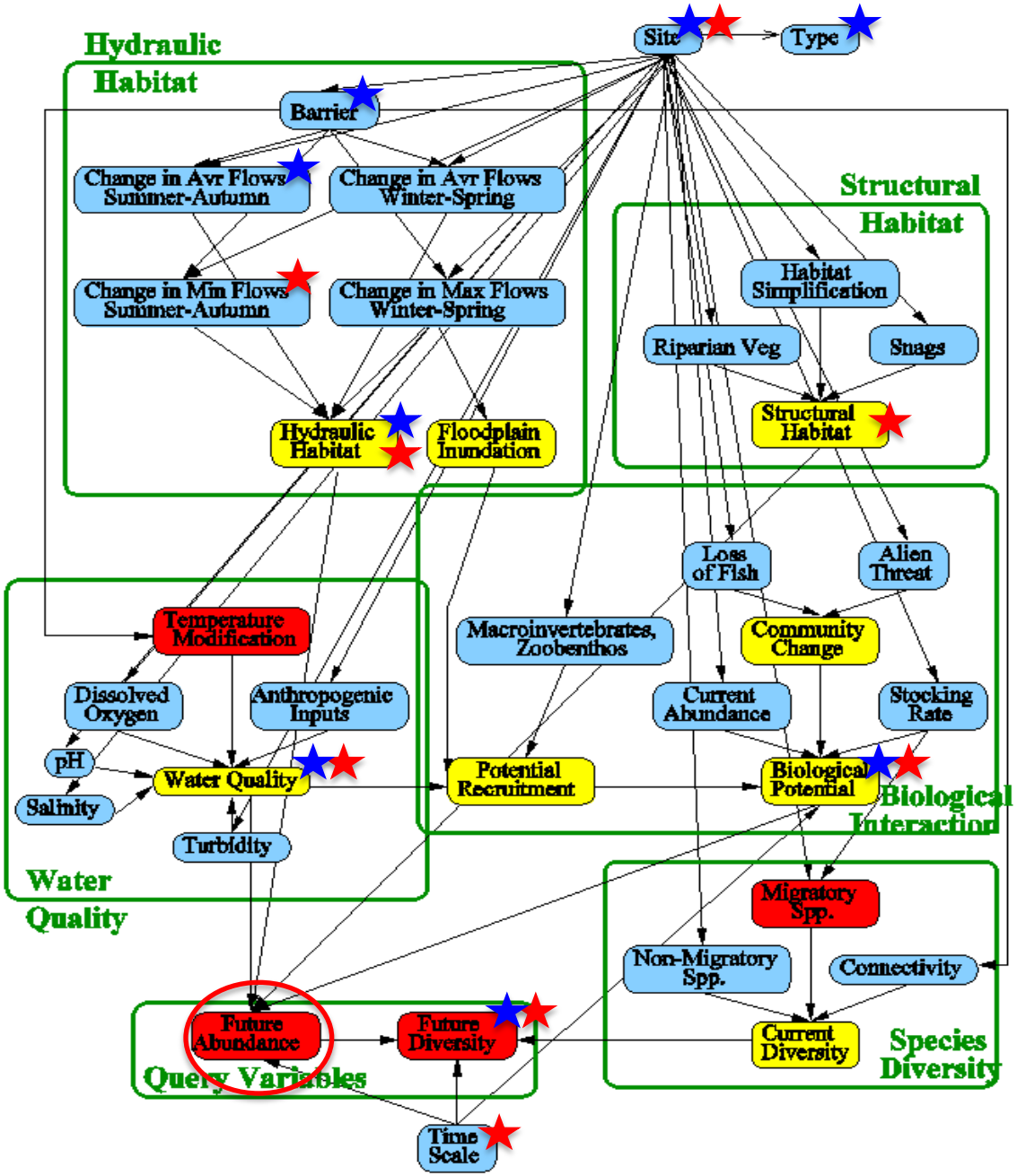


Node	DWI	W=0.5
Lung Cancer	1	1
Tuberculosis	w^2	0.25
Smoking	w	0.5
Bronchitis	w^3	0.125
Visit to Asia	w^3	0.125
Dyspnea	w^3	0.125
Xray Result	0	0
Tuberculosis or Cancer	0	0

Clinic Example with Evidence and Determisitic Node



Node	DWI	W=0.5
Lung Cancer	1	1
Tuberculosis	w	0.5
Smoking	w	0.5
Bronchitis	w^2	0.25
Visit to Asia	w^2	0.25
Dyspnea	w^3	0.125
Xray Result	0	0
Tuberculosis or Cancer	0	0


Hydraulic Habitat
Site
Type
Barrier
Change in Avr Flows Summer-Autumn
Change in Avr Flows Winter-Spring
Change in Min Flows Summer-Autumn
Change in Max Flows Winter-Spring
Hydraulic Habitat
Floodplain Inundation
Structural Habitat
Habitat Simplification
Riparian Veg
Snags
Structural Habitat
Loss of Fish
Alien Threat
Temperature Modification
Macroinvertebrates, Zoobenthos
Community Change
Dissolved Oxygen
Anthropogenic Inputs
Current Abundance
Stocking Rate
pH
Water Quality
Potential Recruitment
Biological Potential
Salinity
Turbidity
Biological Interaction
Water Quality
Migratory Spp.
Non-Migratory Spp.
Connectivity
Future Abundance
Future Diversity
Current Diversity
Species Diversity
Query Variables
Time Scale
DWI
MI

Conclusions

- DWI gives a measure of the sensitivity of the structure.
- DWI can be calculated before the CPTs are filled in.
- Differences between MI and DWI suggest nodes to be investigated.

Future Work

- Improve the algorithm for calculating DWI.
- Investigated alternative metrics.
- Look more closely at theoretical foundations of these metrics.
- Investigate methods to visual the ordering.
- Investigate methods to visual the differences between MI and these metrics.