Bayesian Networks for Risk Assessment

Society of Information Risk Analysts
14 November 2014

Norman Fenton
Queen Mary University of London
and
Agena Ltd

Outline

Overview of Bayes and Bayesian networks

Why Bayesian networks are needed for risk assessment

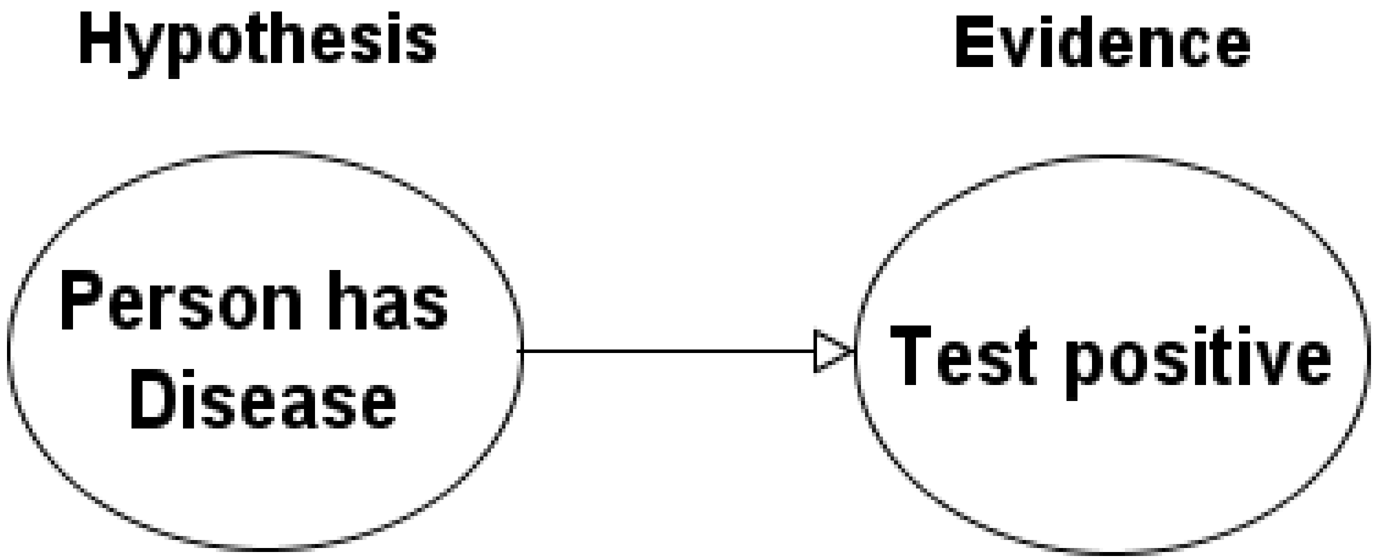
The challenges

Applications

www.BayesianRisk.com

www.AgenaRisk.com

A typical probability problem



1 in a 1000

False	0.999
True	0.0010

100% accurate for those with disease; 95% accurate for those without

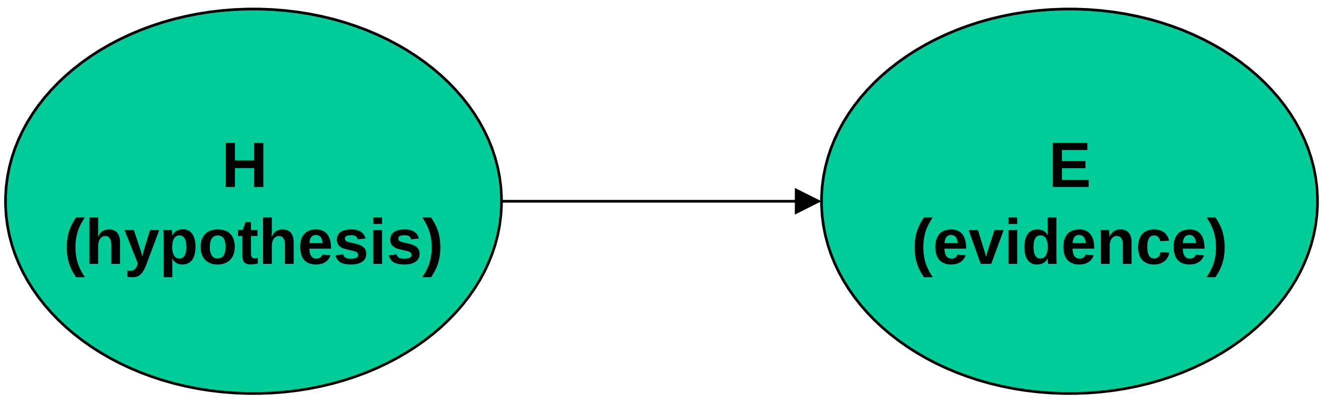
Person has Disease	False	True
False	0.95	0.0
True	0.05	1.0

What is the probability a person has the disease if they test positive?

Bayes Theorem

Have a prior P(H) ("person has disease")

Now get some evidence E ("test result positive")

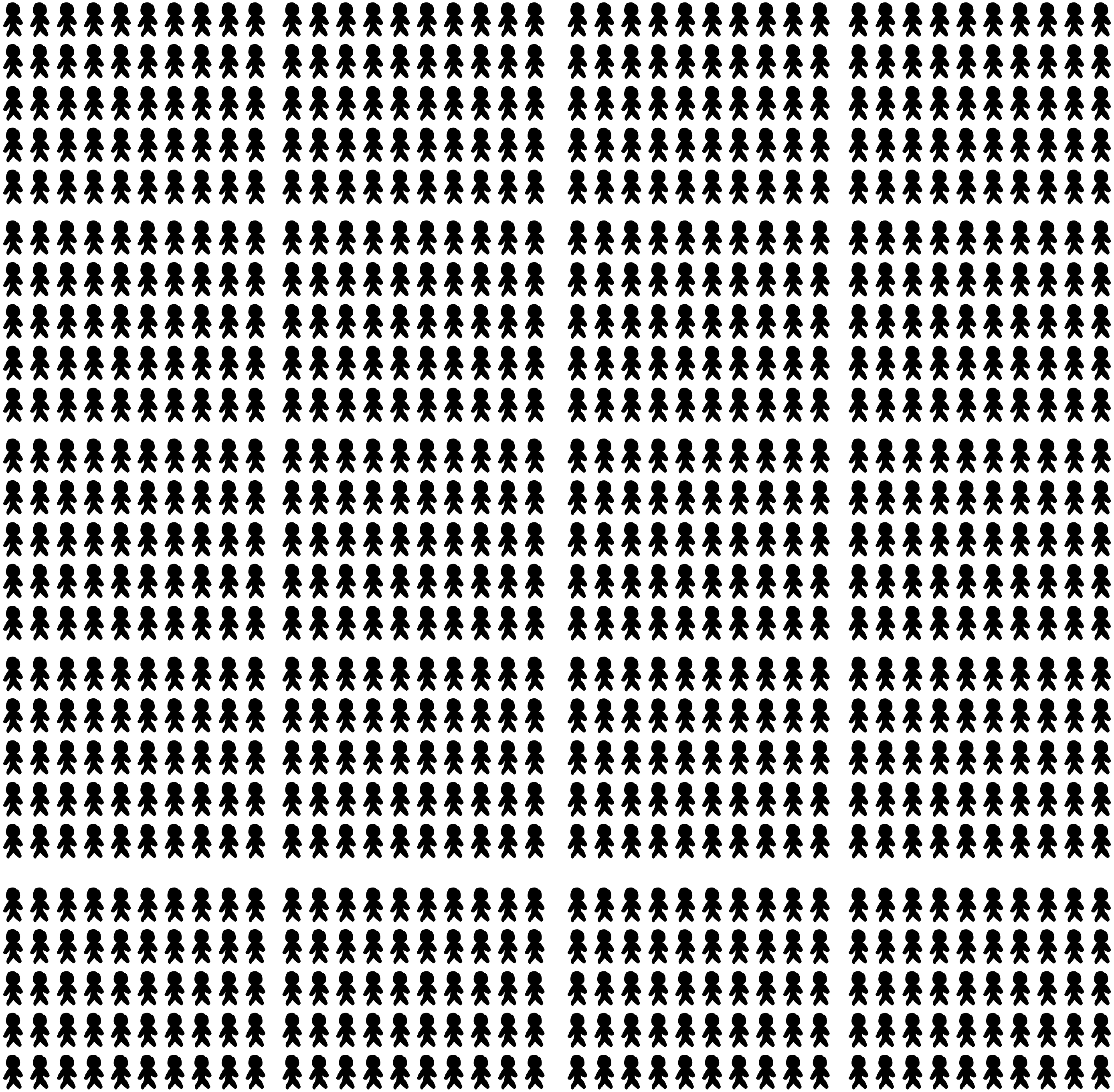


We know P(E|H)

But we want the posterior P(H|E)

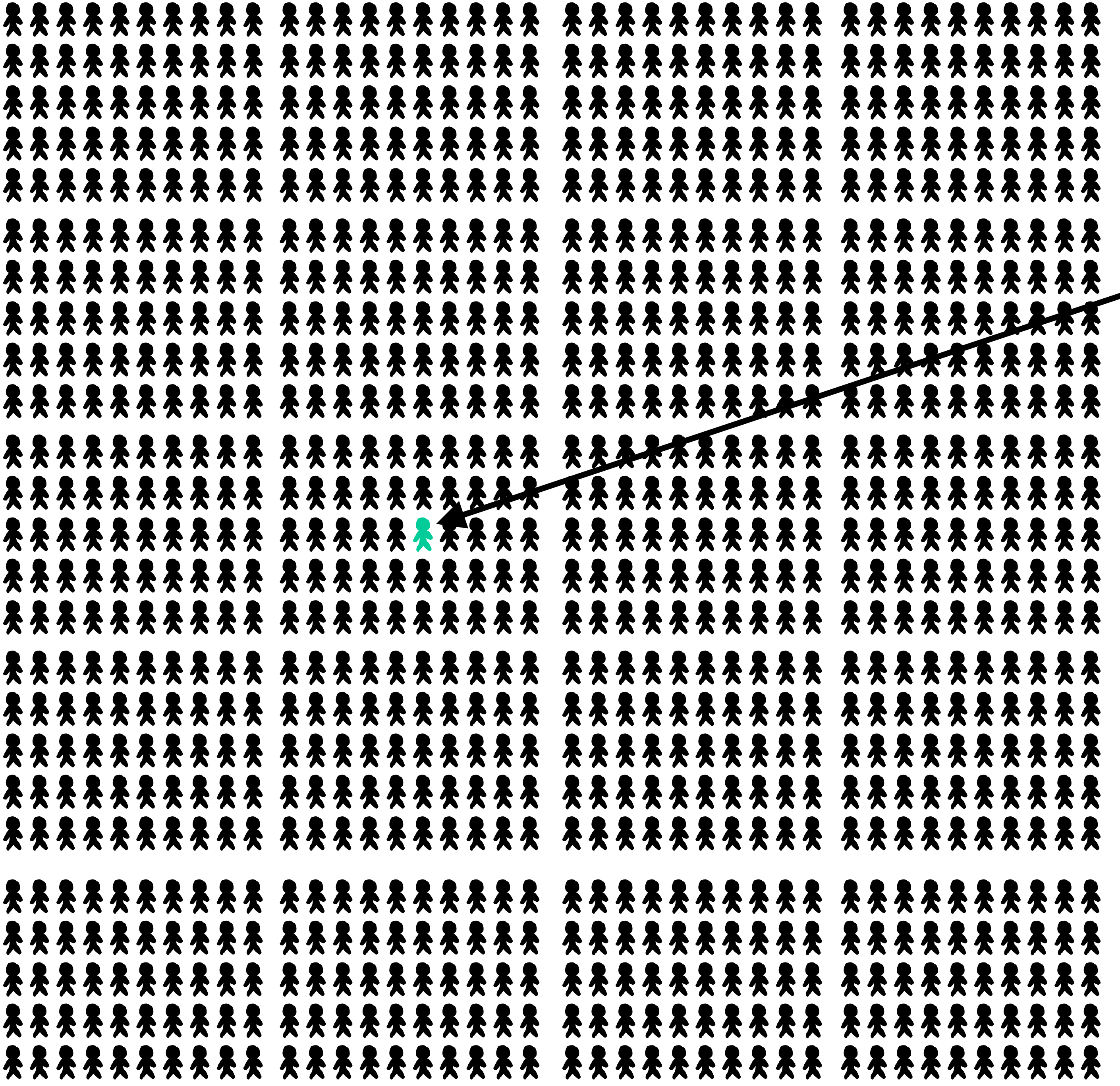
$$P(H|E) = \frac{P(E|H)*P(H)}{P(E)} = \frac{P(E|H)*P(H)}{P(E|H)*P(H) + P(E|\text{not } H)*P(\text{not } H)}$$

$$P(H|E) = \frac{1*0.001}{1*0.001 + 0.05*0.999} = \frac{0.001}{0.5005} \approx 0.02$$



Imagine 100,000 people

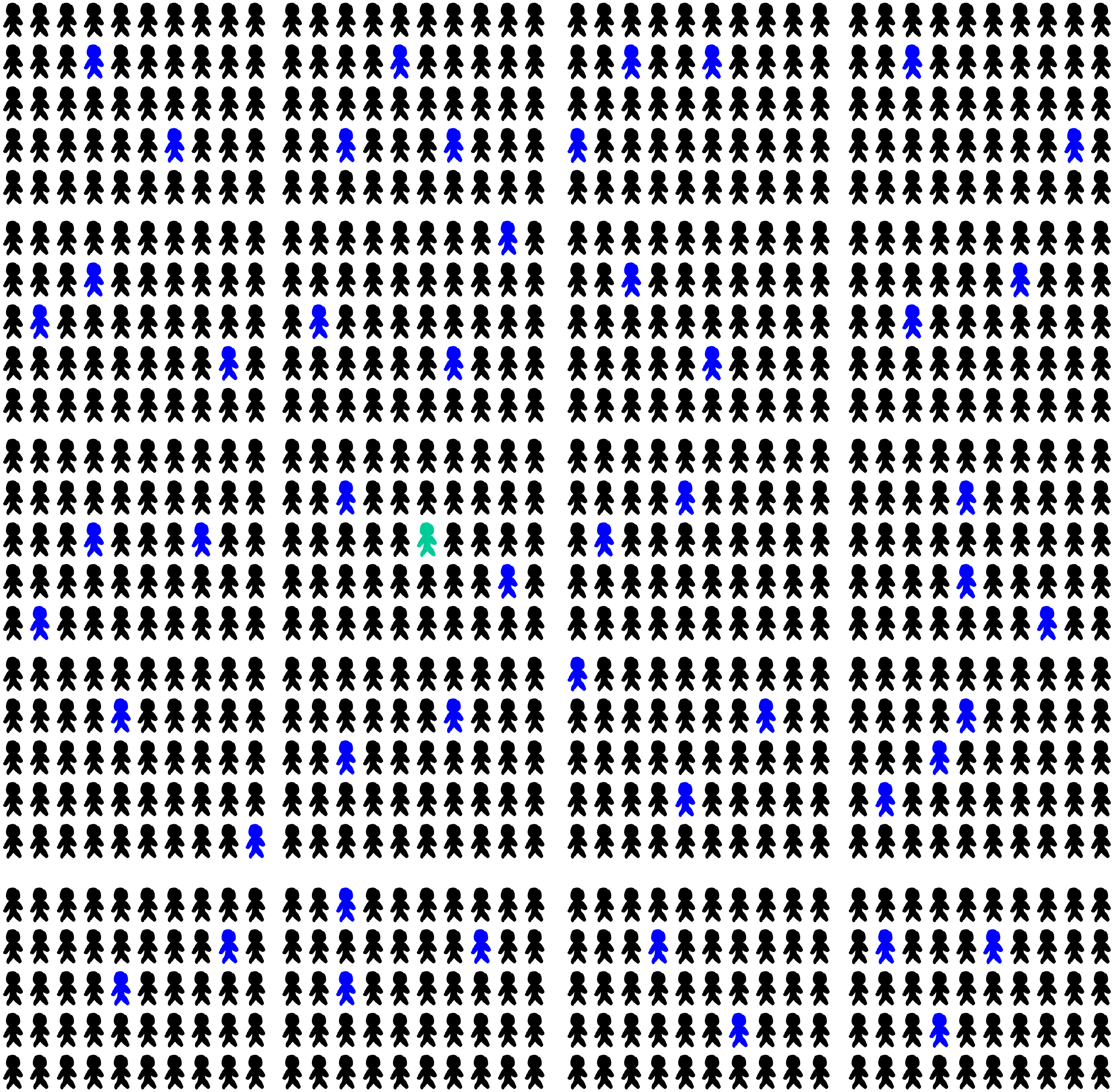
Slide 6



Out of whom 10 has the disease

Slide 7

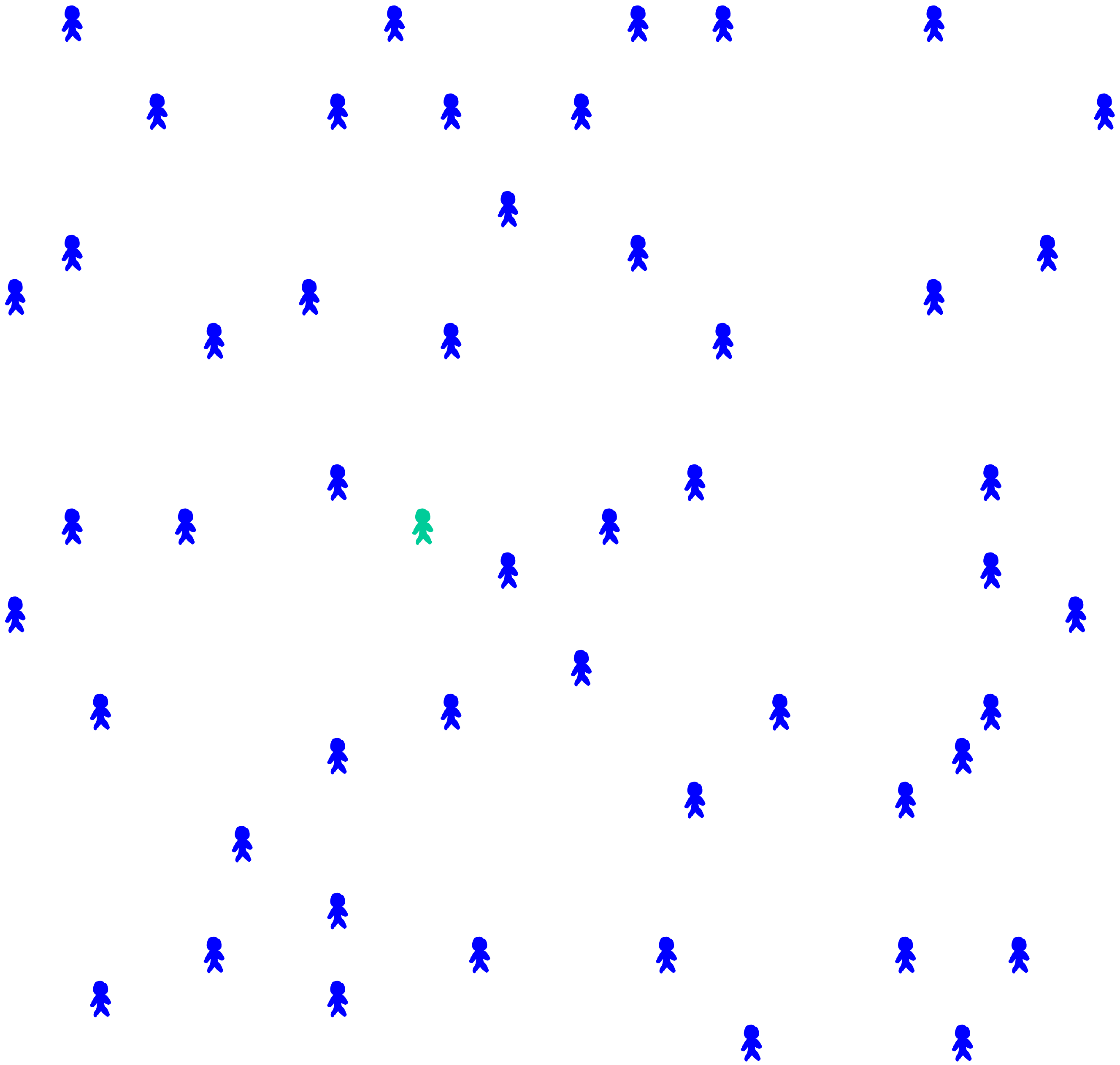
Slide 7



But about 500 of the Remaining 9990 people without the disease test positive

Slide 8

Slide 8



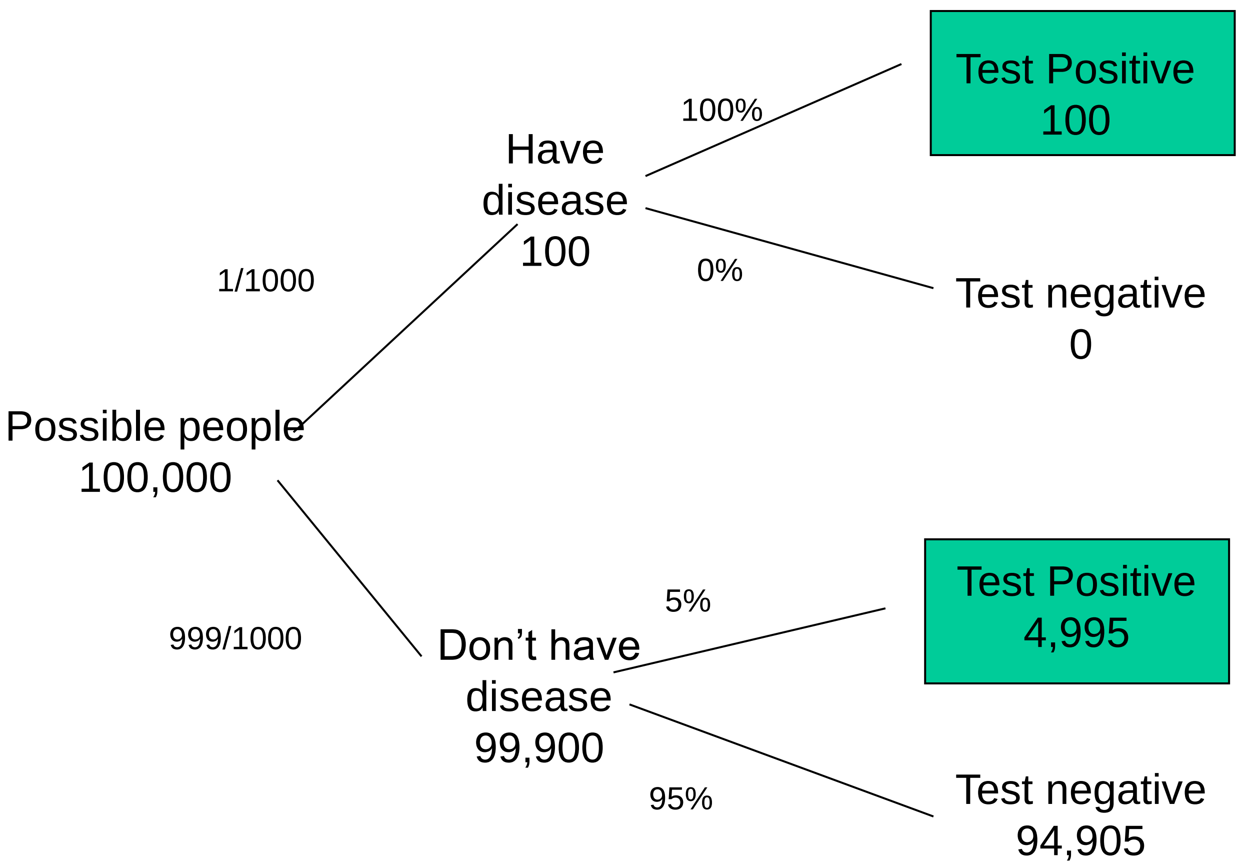
So 10 out of 510 who test positive actually have the disease

That’s just under 2%

That’s very different from the 95% assumed by most medics

Slide 9

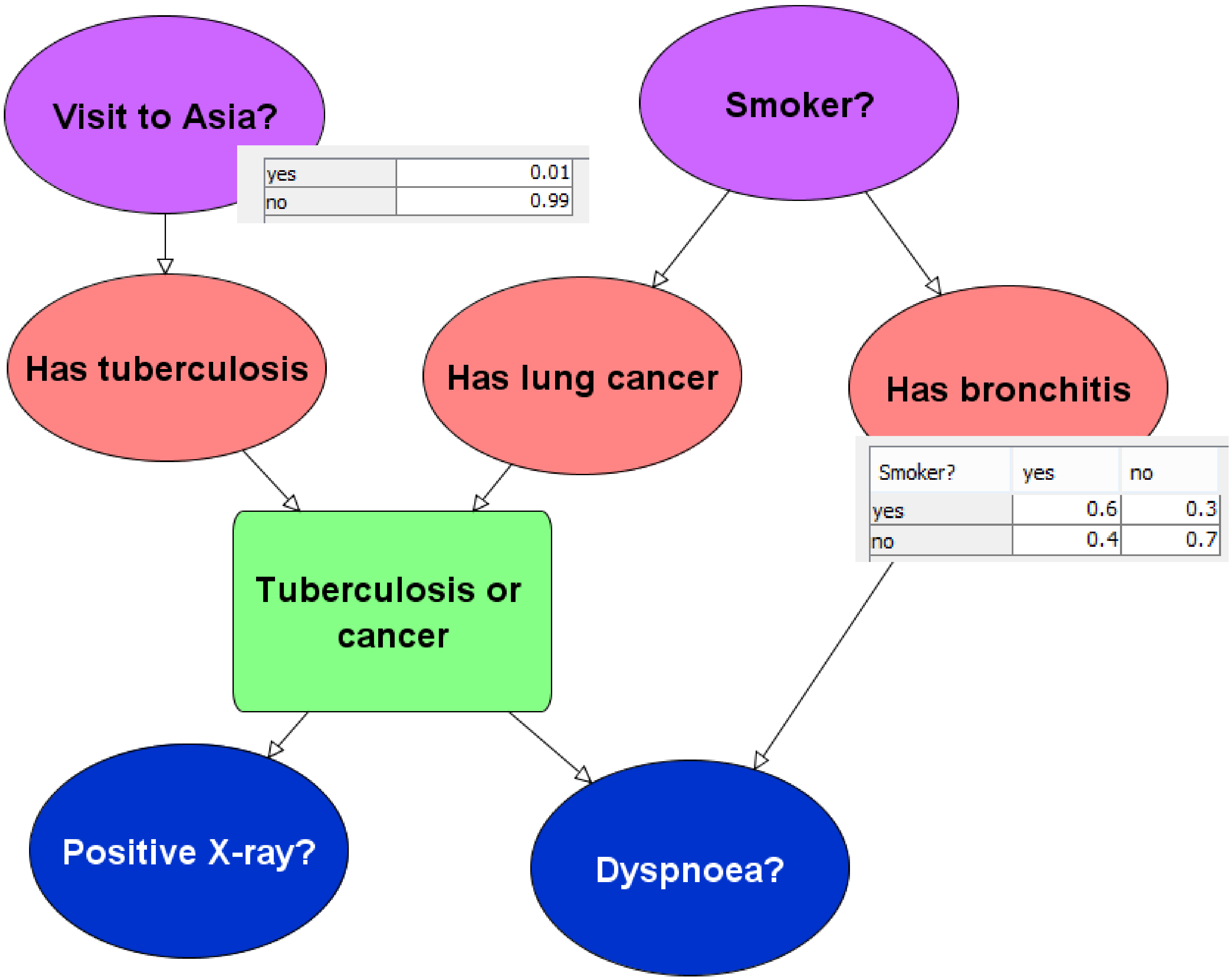
An alternative visual explanation



So 100 out of 5,095 who test positive match actually have the disease, I.e. Under 2%

Slide 10

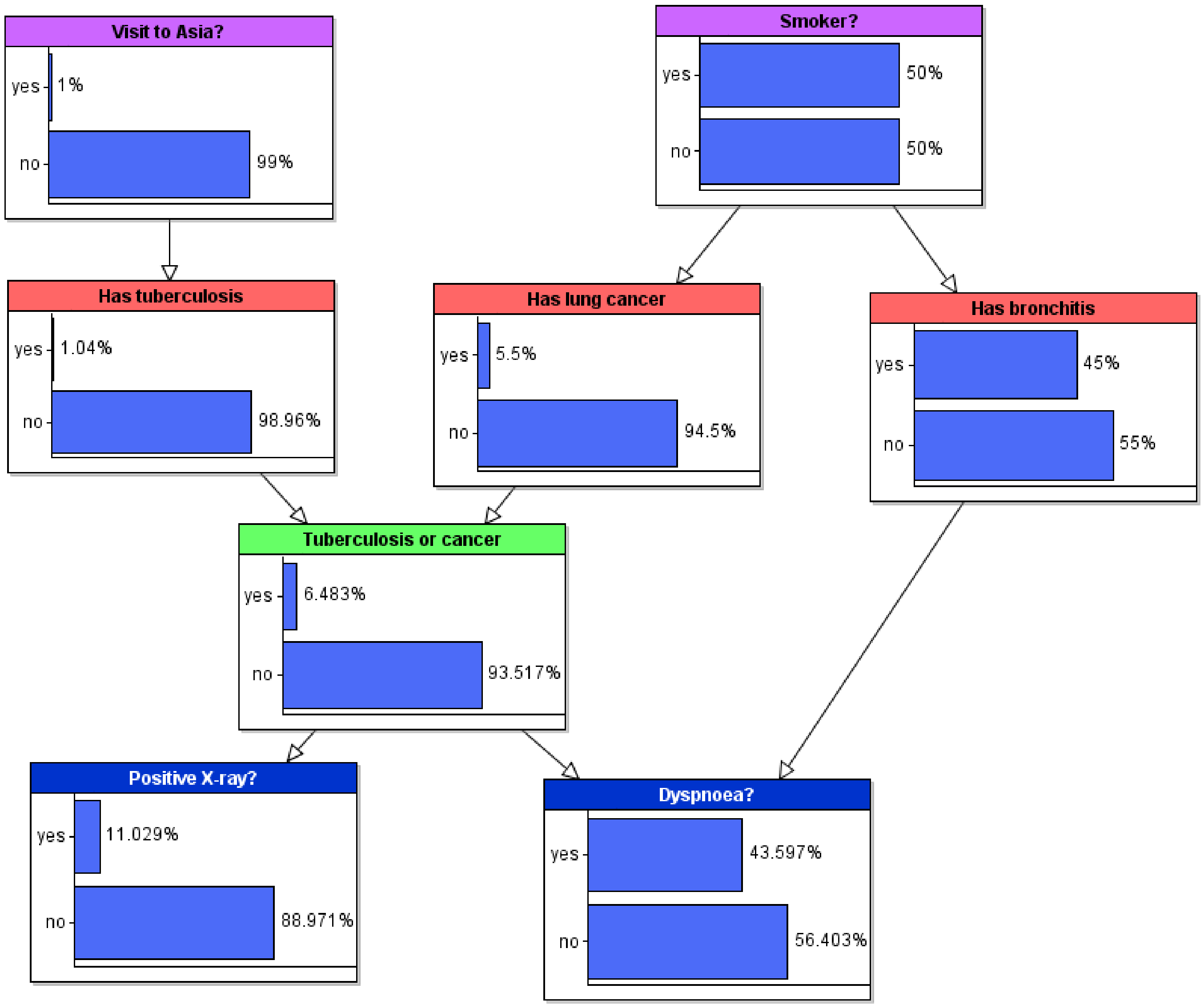
A Classic BN


Visit to Asia?
yes 0.01
no 0.99
Smoker?
Has tuberculosis
Has lung cancer
Has bronchitis
Smoker? yes no
yes 0.6 0.3
no 0.4 0.7
Tuberculosis or cancer
Positive X-ray?
Dyspnoea?

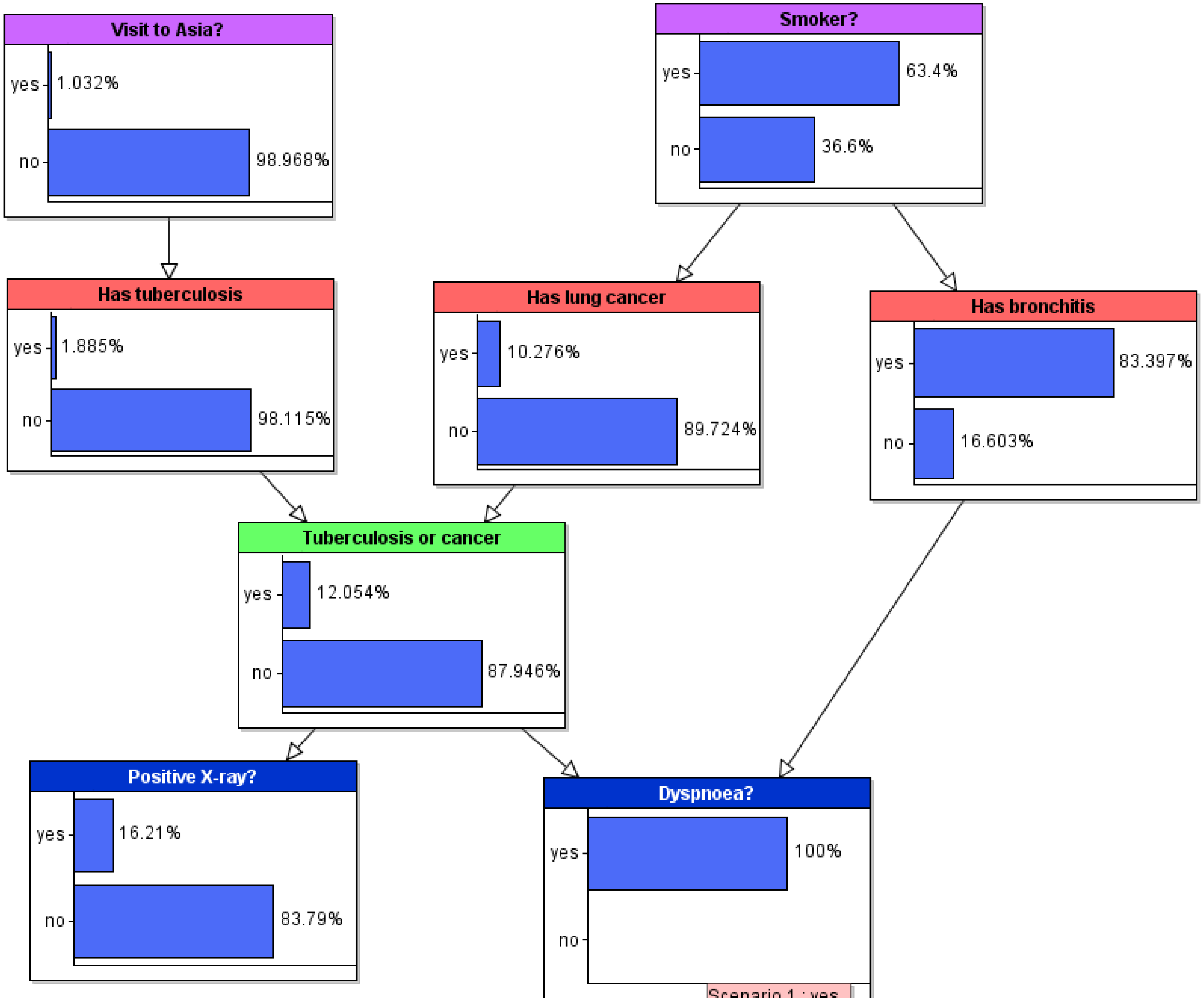
Bayesian Propagation

- Applying Bayes theorem to update all probabilities when new evidence is entered
- Intractable even for small BNs
- Breakthrough in late 1980s - fast algorithms
- Tools implement efficient propagation

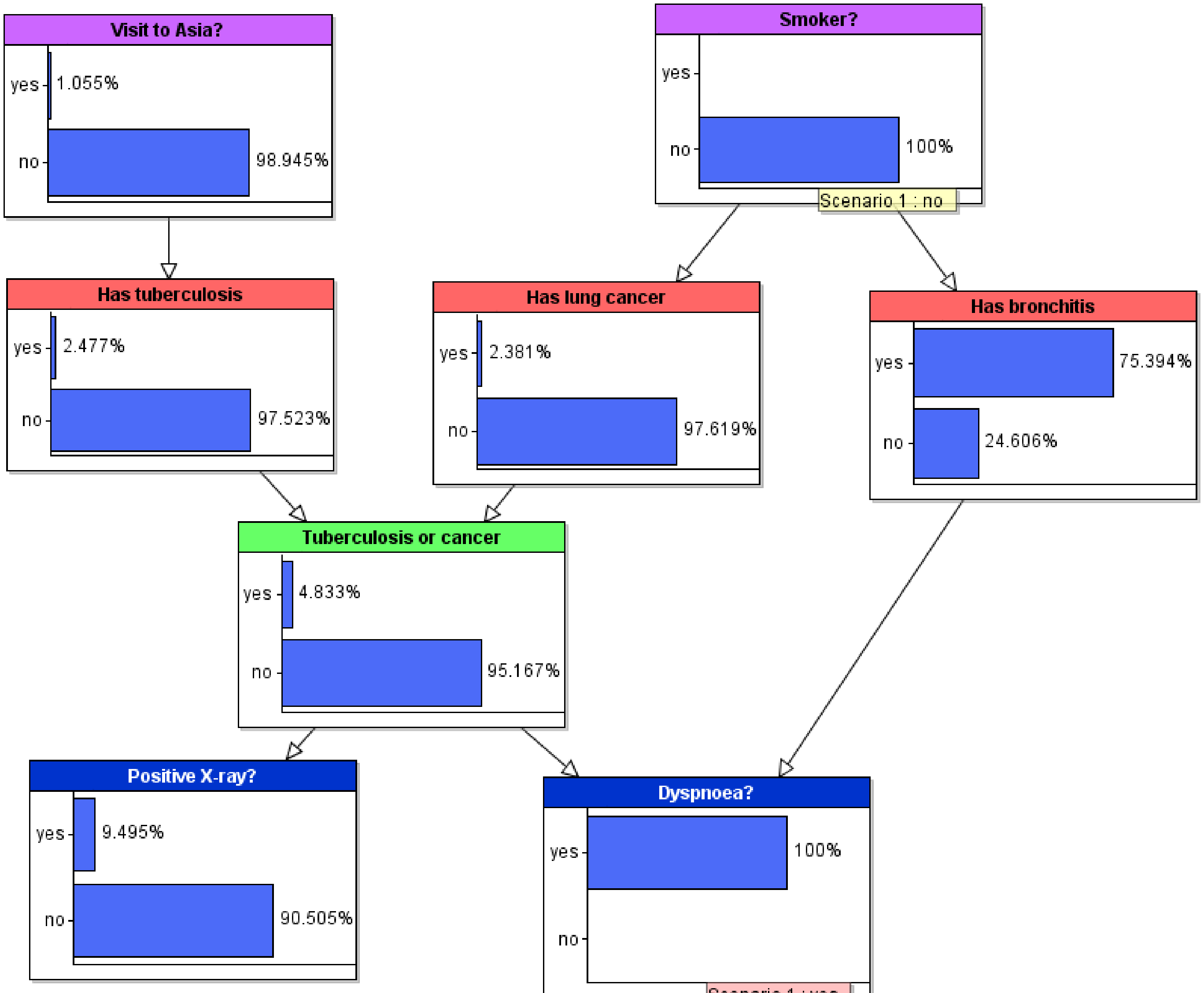
A Classic BN: Marginals



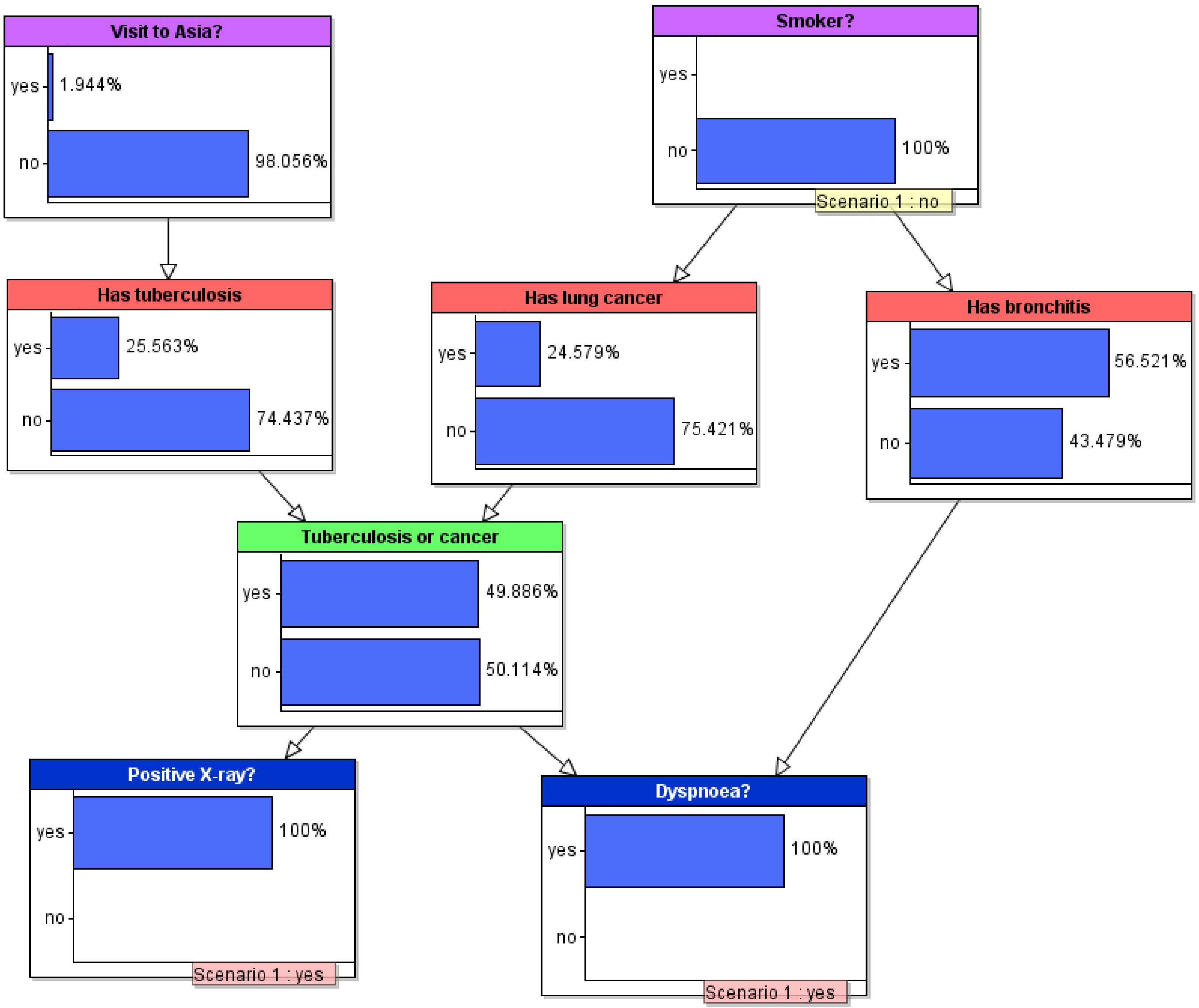
Dyspnoea observed



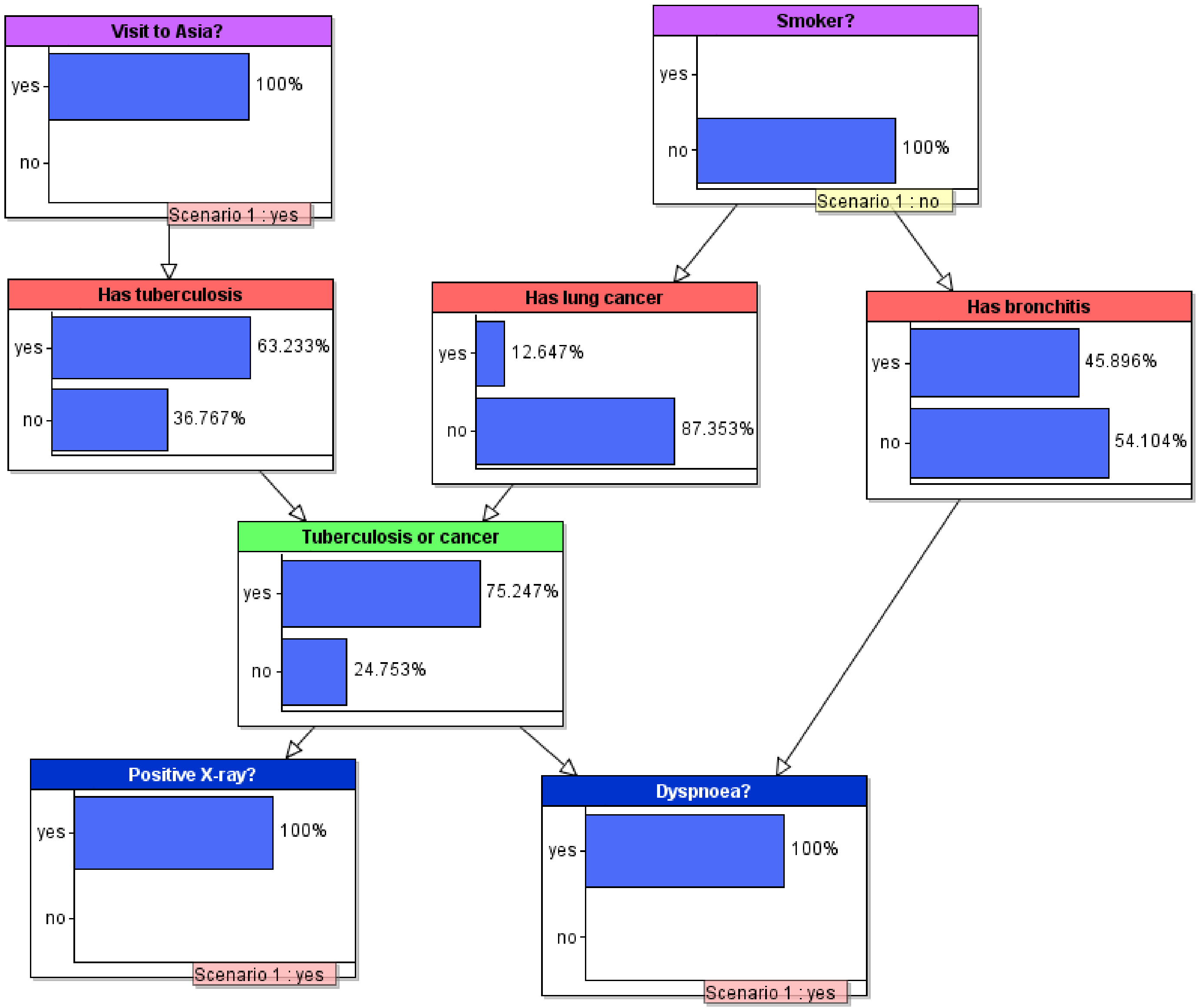
Also non-smoker



Positive x-ray



..but recent visit to Asia



The power of BNs

Explicitly model causal factors

Reason from effect to cause and vice versa

‘Explaining away’

Overturn previous beliefs

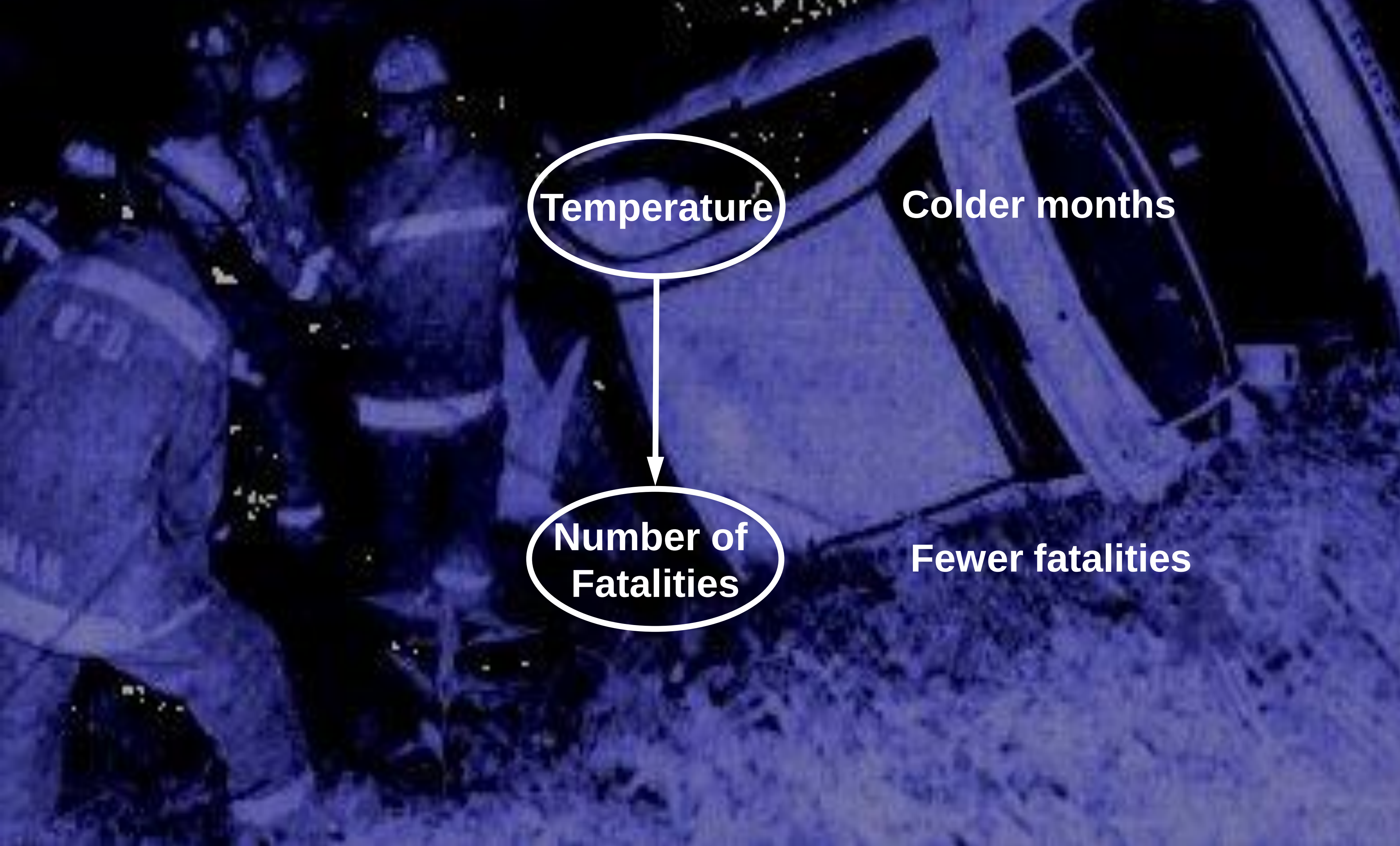
Make predictions with incomplete data

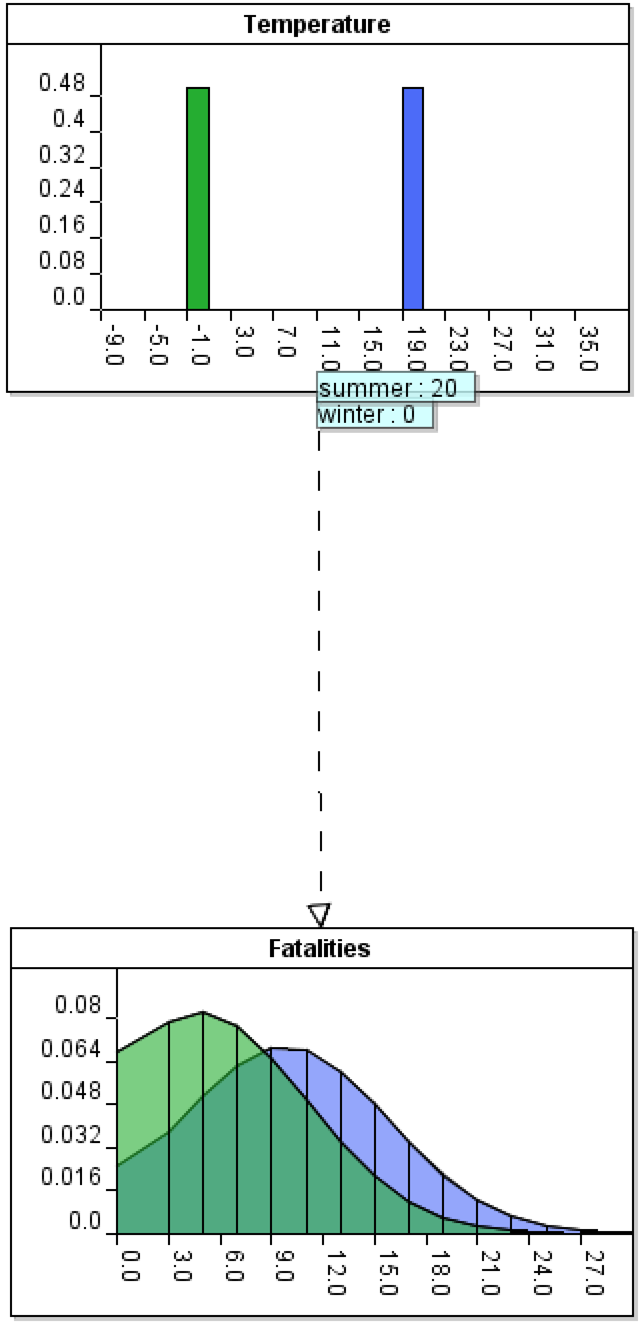
Combine diverse types of evidence

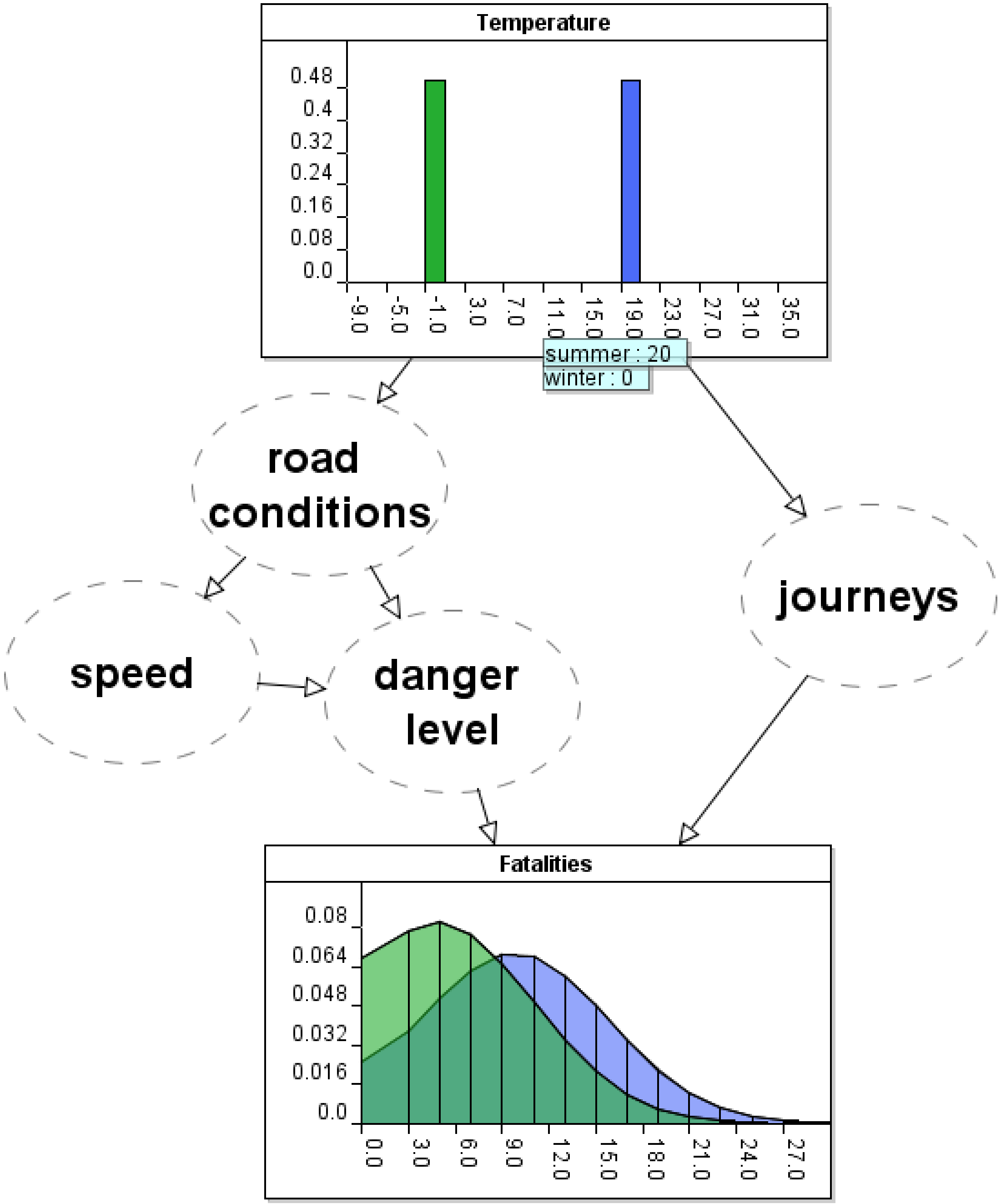
Visible auditable reasoning

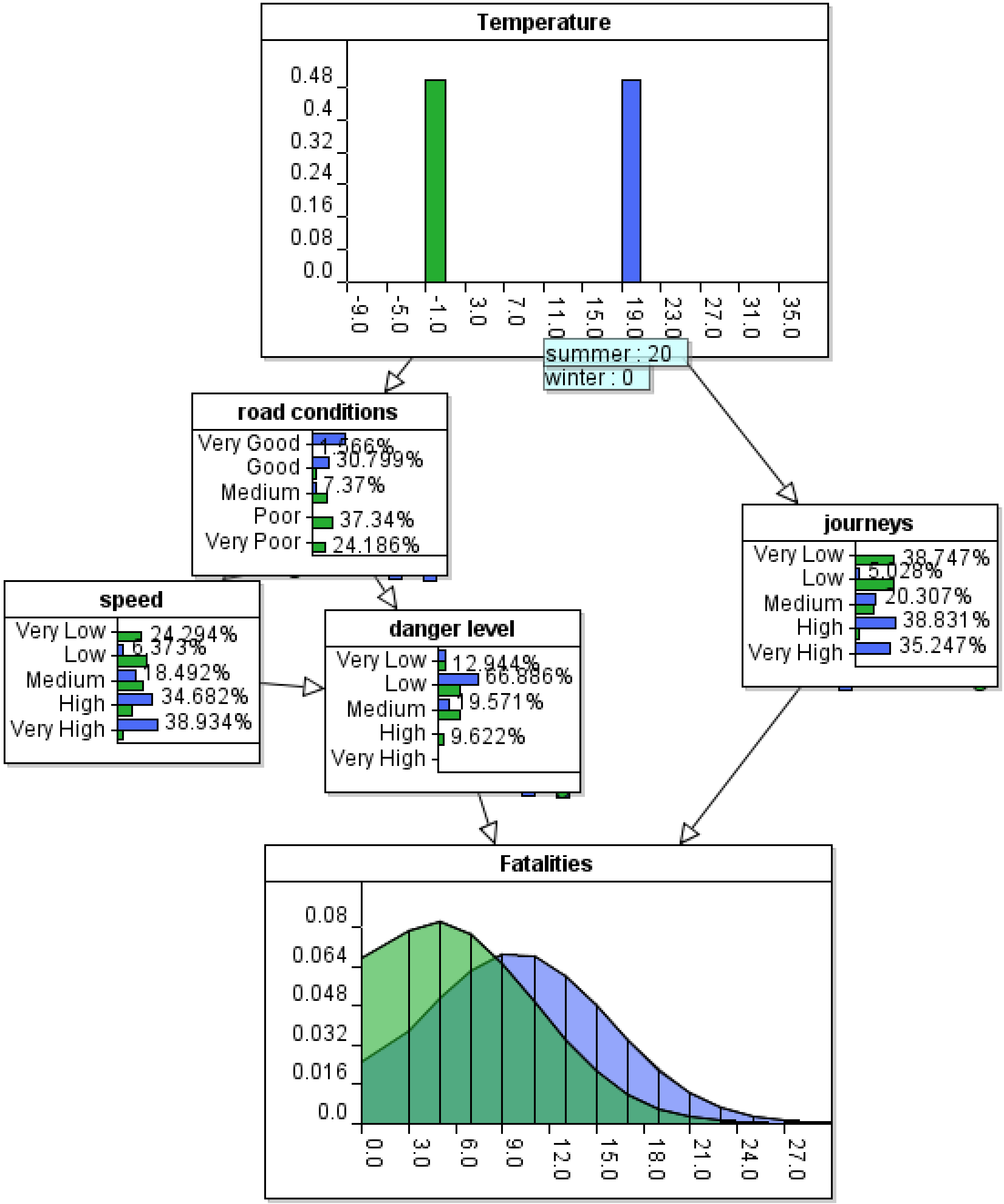
Why Bayesian networks are needed for risk assessment

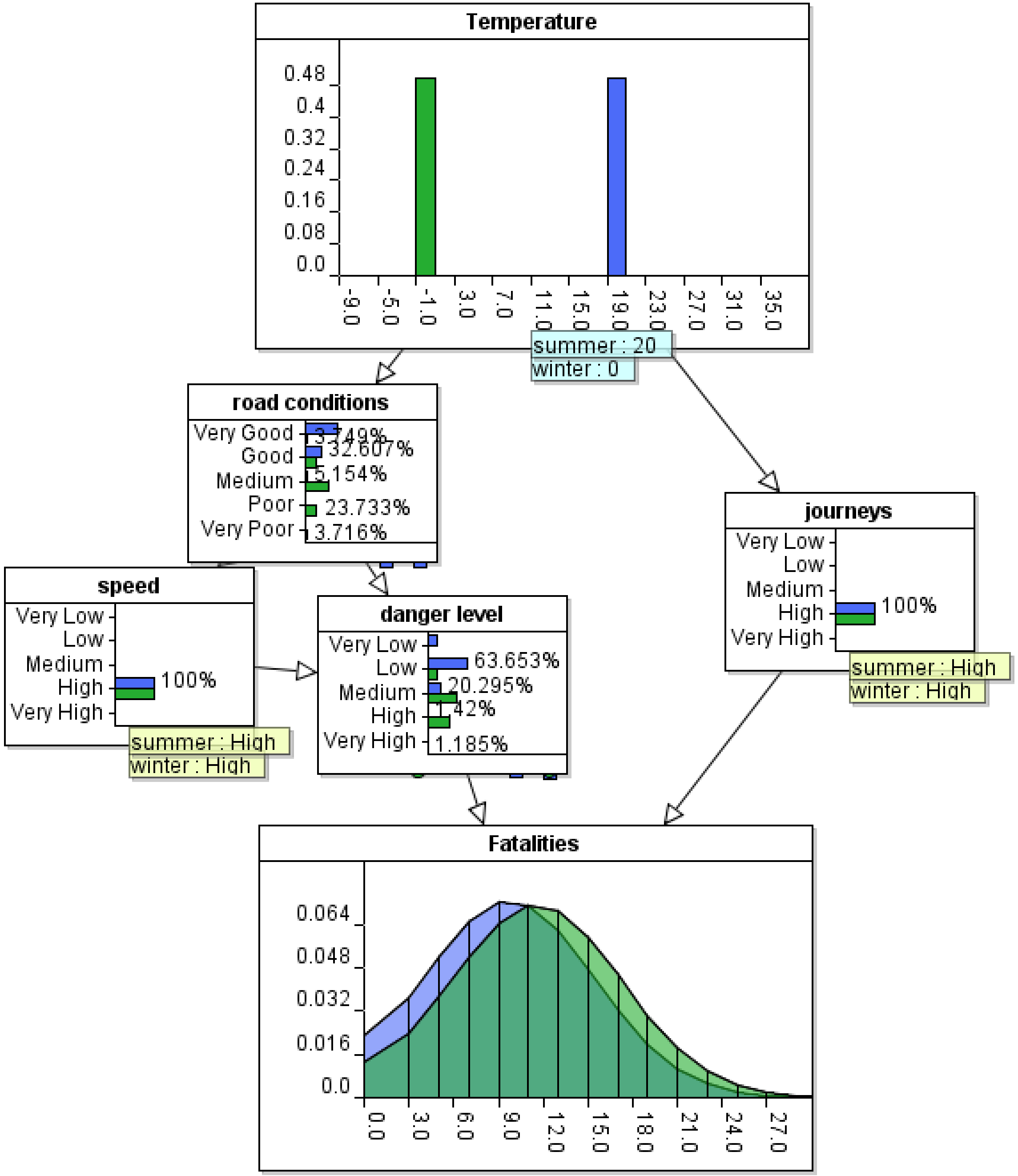
Assessing Risk of Road Fatalities: Classic Statistical Approach



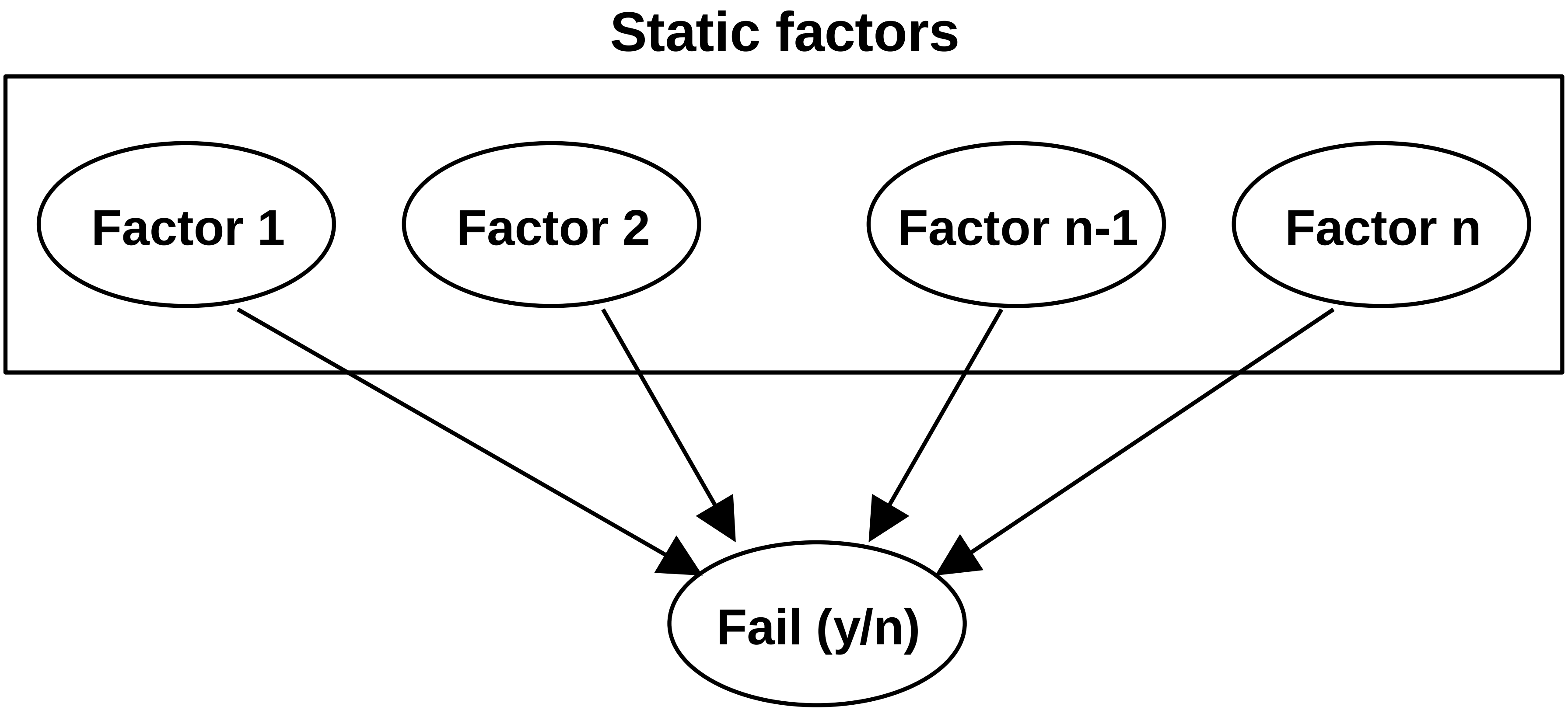

Temperature
0.48
0.4
0.32
0.24
0.16
0.08
0.0
-9.0
-5.0
-1.0
3.0
7.0
11.0
15.0
19.0
23.0
27.0
31.0
35.0
summer : 20
winter : 0
Fatalities
0.08
0.064
0.048
0.032
0.016
0.0
0.0
3.0
6.0
9.0
12.0
15.0
18.0
21.0
24.0
27.0


Temperature
0.48
0.4
0.32
0.24
0.16
0.08
0.0
-9.0
-5.0
-1.0
3.0
7.0
11.0
15.0
19.0
23.0
27.0
31.0
35.0
summer : 20
winter : 0
road conditions
journeys
speed
danger level
Fatalities
0.08
0.064
0.048
0.032
0.016
0.0
0.0
3.0
6.0
9.0
12.0
15.0
18.0
21.0
24.0
27.0

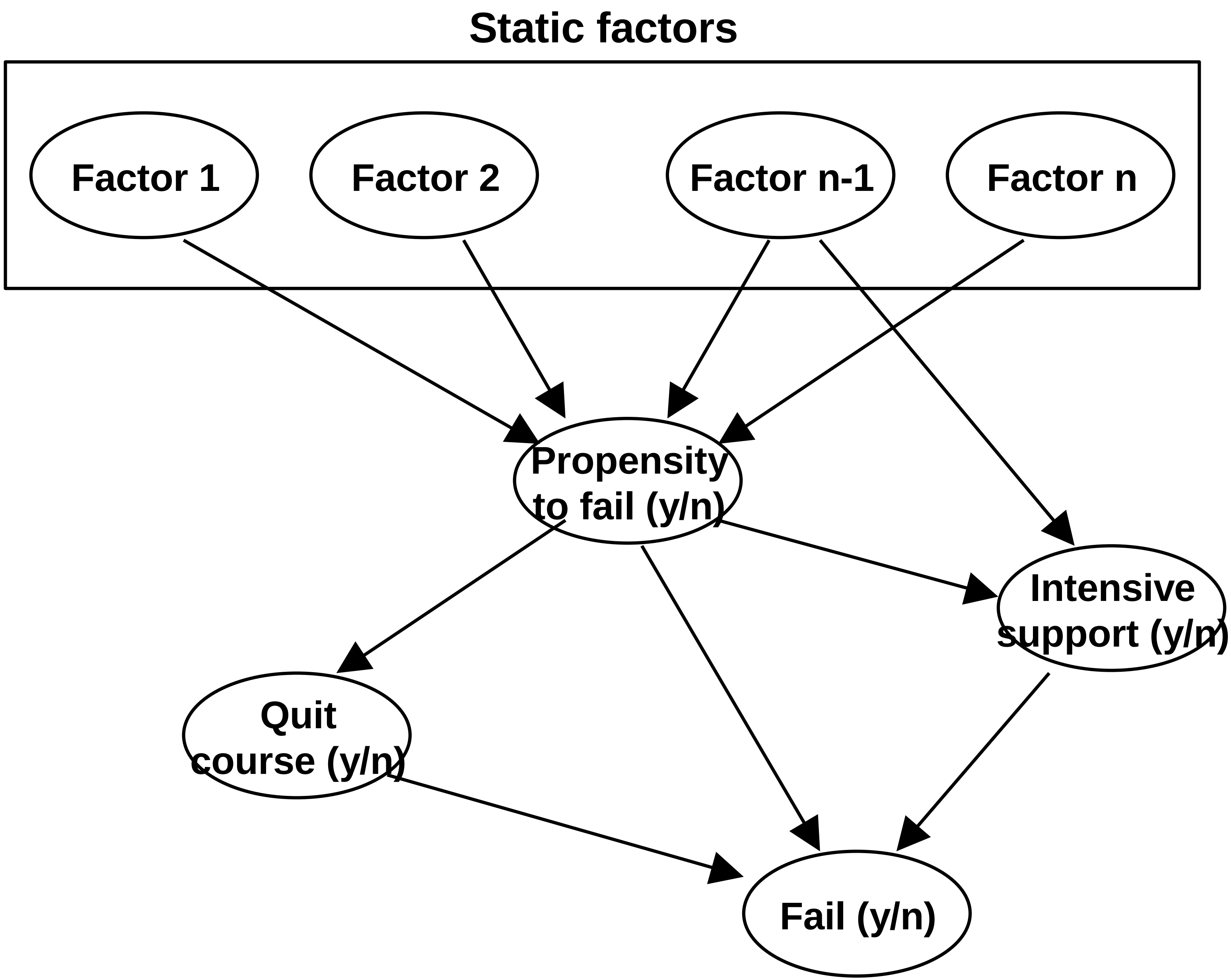

Temperature
0.48
0.4
0.32
0.24
0.16
0.08
0.0
-9.0
-5.0
-1.0
3.0
7.0
11.0
15.0
19.0
23.0
27.0
31.0
35.0
summer : 20
winter : 0
road conditions
Very Good
Good
Medium
Poor
Very Poor
30.799%
7.37%
37.34%
24.186%
journeys
Very Low
Low
Medium
High
Very High
38.747%
5.028%
20.307%
38.831%
35.247%
speed
Very Low
Low
Medium
High
Very High
24.294%
6.373%
18.492%
34.682%
38.934%
danger level
Very Low
Low
Medium
High
Very High
12.944%
66.886%
9.571%
9.622%
Fatalities
0.08
0.064
0.048
0.032
0.016
0.0
0.0
3.0
6.0
9.0
12.0
15.0
18.0
21.0
24.0
27.0


Temperature
0.48
0.4
0.32
0.24
0.16
0.08
0.0
-9.0
-5.0
-1.0
3.0
7.0
11.0
15.0
19.0
23.0
27.0
31.0
35.0
summer : 20
winter : 0
road conditions
Very Good
Good
Medium
Poor
Very Poor
13.749%
32.607%
15.154%
23.733%
13.716%
journeys
Very Low
Low
Medium
High
Very High
100%
summer : High
winter : High
speed
Very Low
Low
Medium
High
Very High
100%
summer : High
winter : High
danger level
Very Low
Low
Medium
High
Very High
63.653%
20.295%
1.42%
1.185%
Fatalities
0.064
0.048
0.032
0.016
0.0
0.0
3.0
6.0
9.0
12.0
15.0
18.0
21.0
24.0
27.0

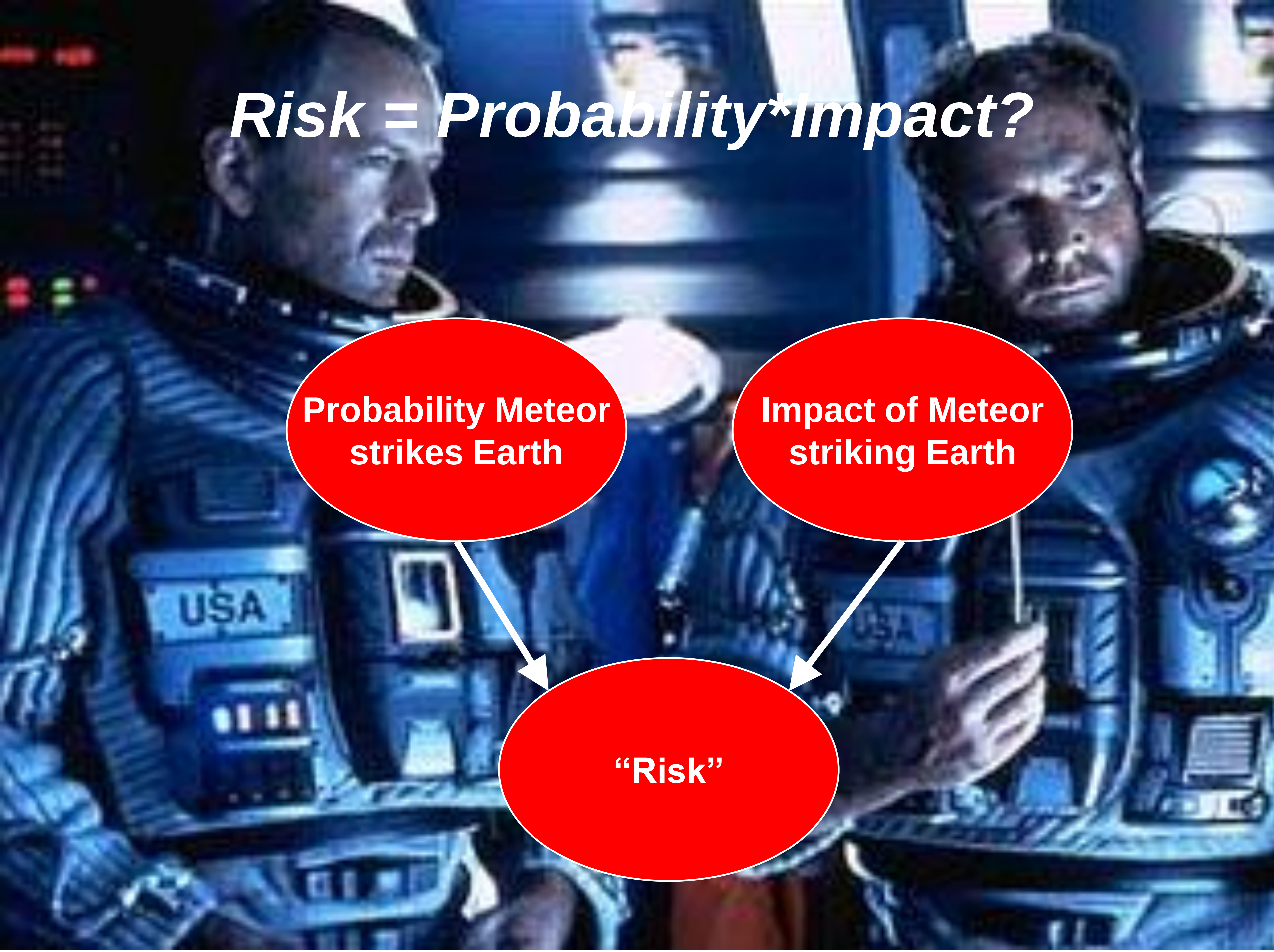
Classic (but wrong) approach to risk



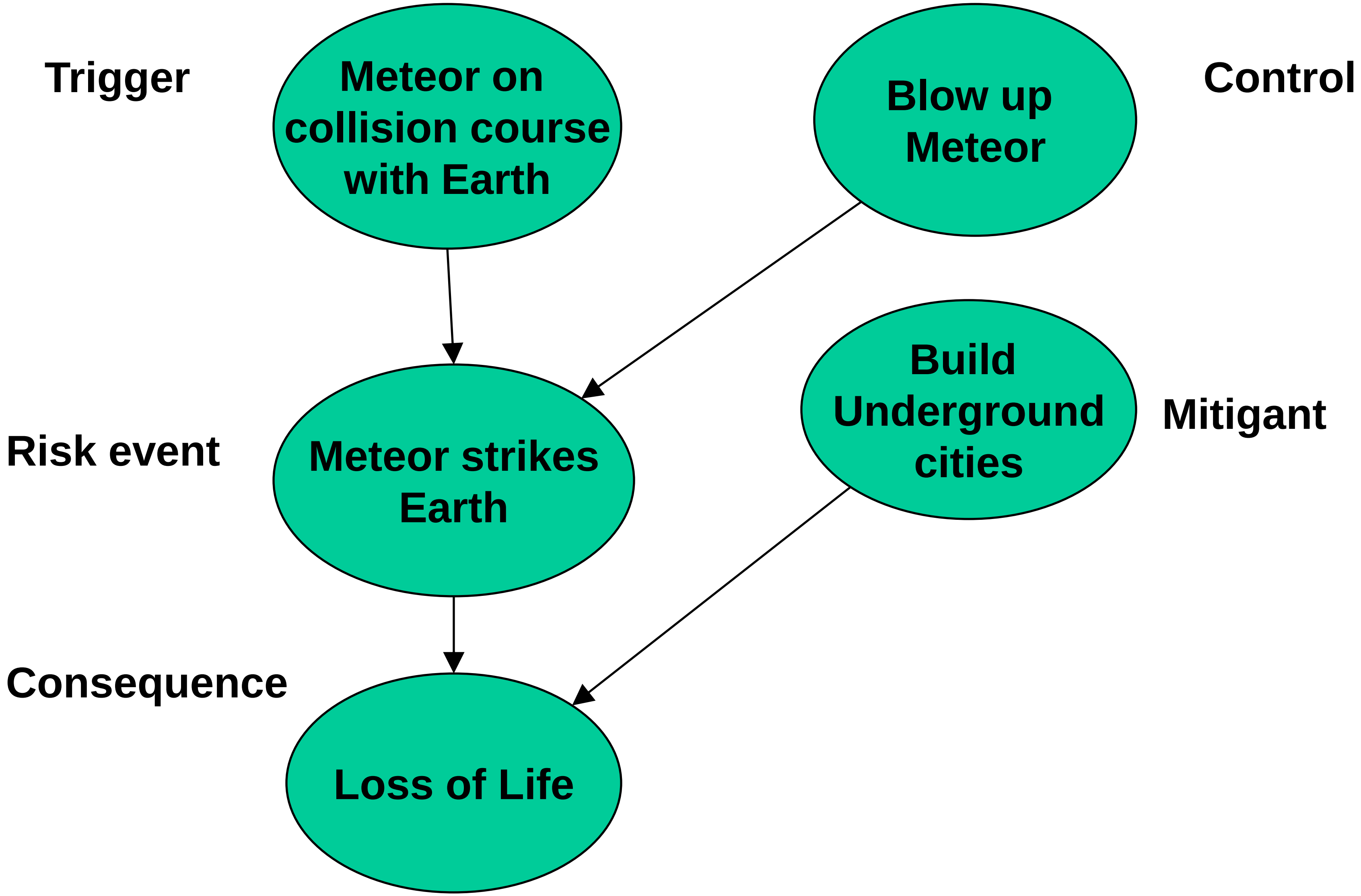
What we really need



*Risk = Probability*Impact?*



Bayesian Net with causal view of risk



The challenges

Bayesian Networks: Barriers and Challenges

Resistence to subjective probabilities

Building realistic models

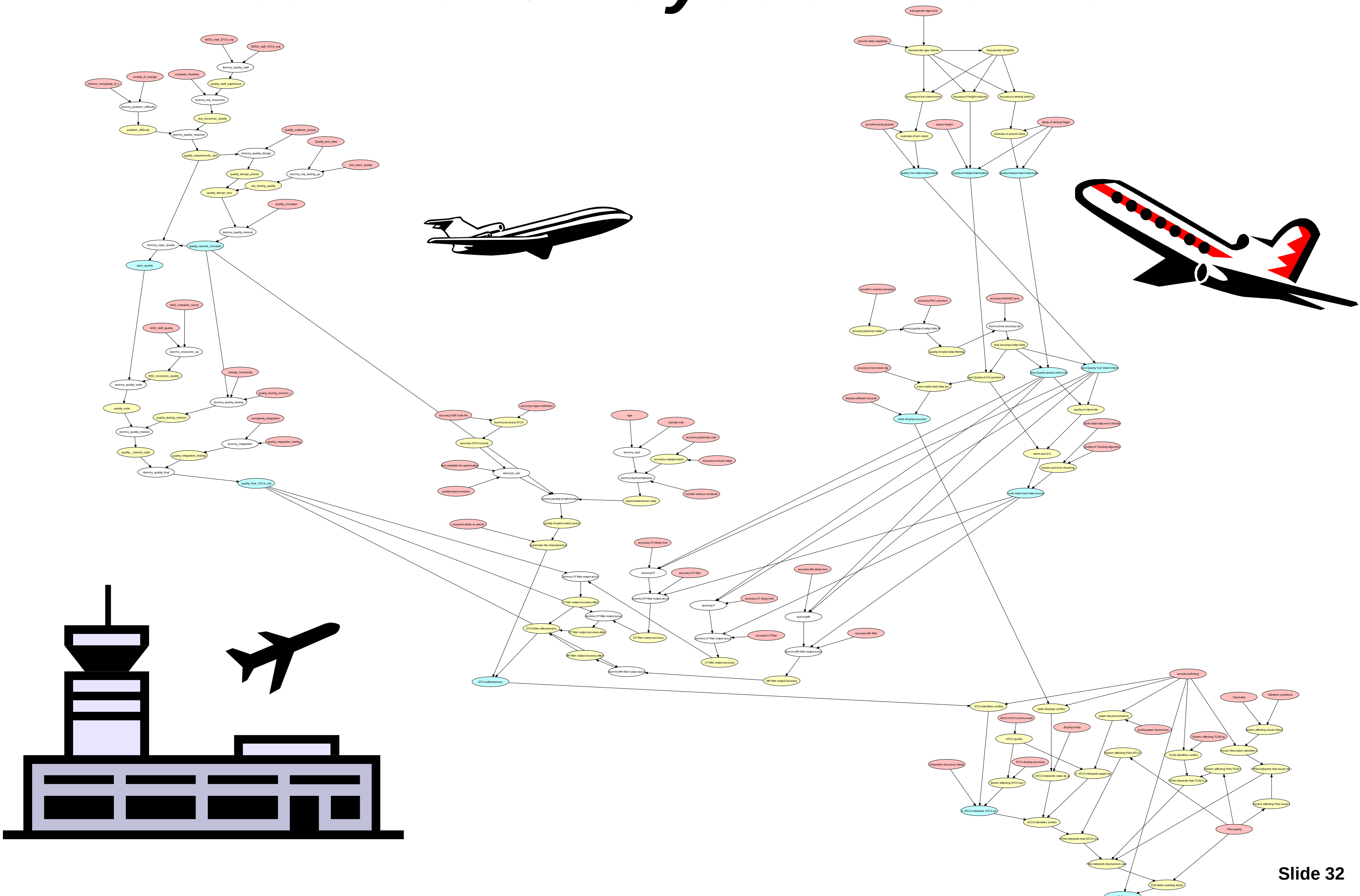
Handling continuous variables properly

Resistence to Bayes

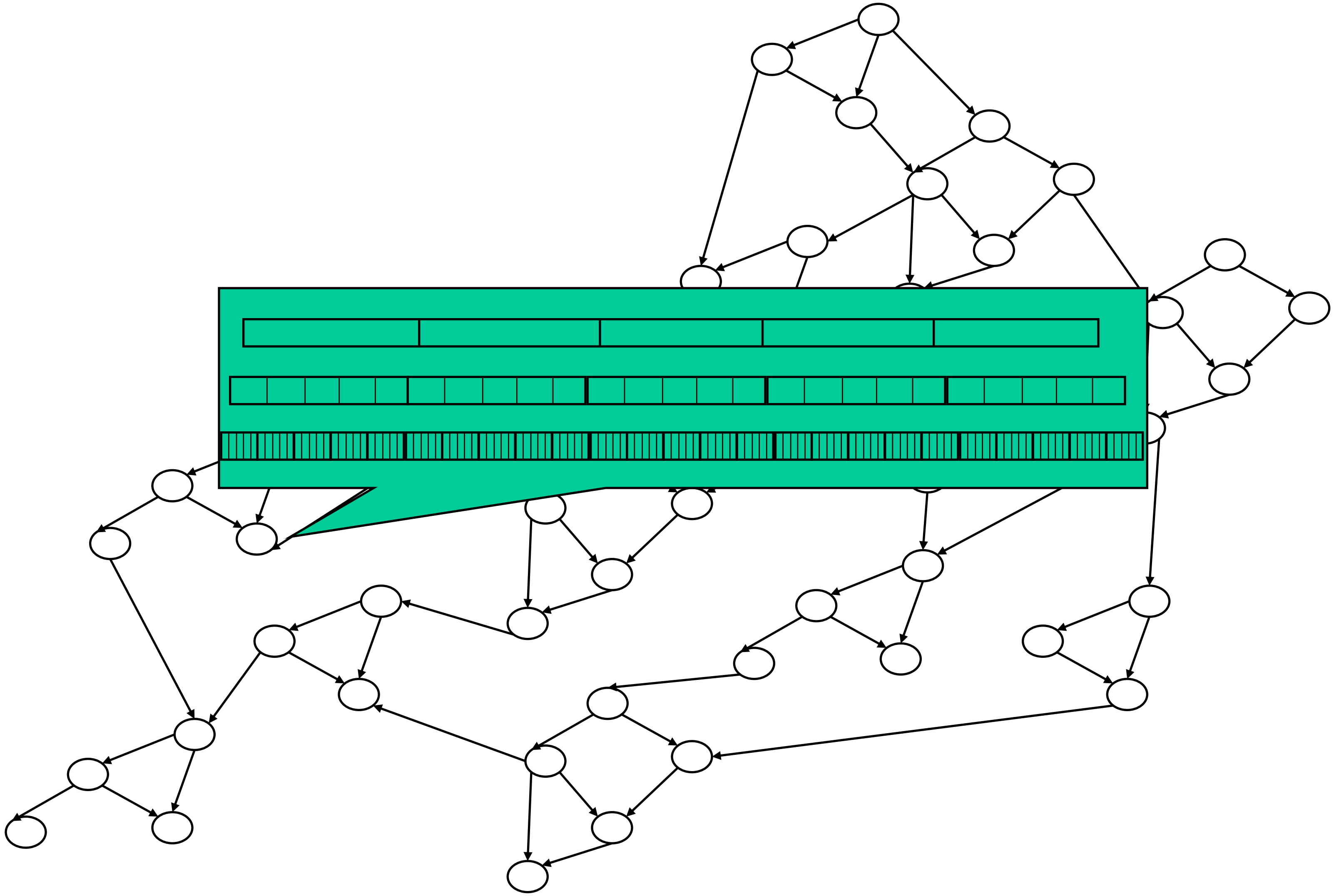
- **OK – but even if I accept the calculations are ‘correct’ I don’t accept subjective priors**

There is no such thing as a truly objective frequentist approach

A Real World Bayesian Network



How to build big BNs?



Options for Building BNs

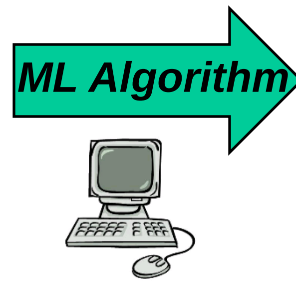
Structure and probability tables all learnt from data only (‘machine learning’)

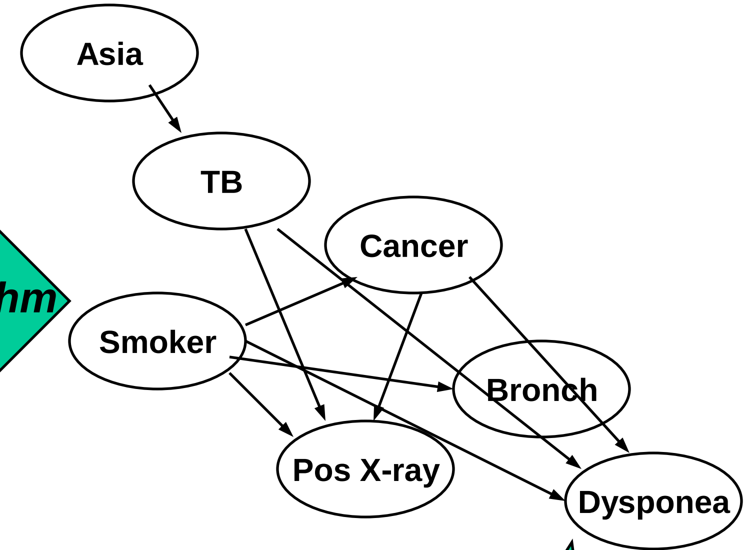
Structure informed by experts, probability tables learnt from data

Structure and tables built by experts

Machine Learning Option

VisitAsia	Tuberculosis	Smoking	Cancer	TbOrCa	XRay	Bronchitis	Dyspnea
No_Visit	Absent	Smoker	Absent	False	Normal	Present	Present
No_Visit	Absent	Smoker	Absent	False	Normal	Absent	Absent
No_Visit	Absent	Smoker	Absent	False	Normal	Absent	Absent
No_Visit	Absent	NonSmoker	Absent	False	Normal	Absent	Present
No_Visit	Absent	Smoker	Absent	False	Normal	Absent	Absent
No_Visit	Absent	Smoker	Absent	False	Normal	Absent	Absent
No_Visit	Absent	NonSmoker	Absent	False	Normal	Present	Present
No_Visit	Absent	Smoker	Present	True	Abnormal	Present	Present
No_Visit	Absent	NonSmoker	Absent	False	Normal	Absent	Absent
No_Visit	Absent	NonSmoker	Absent	False	Normal	Absent	Absent
No_Visit	Absent	Smoker	Absent	False	Normal	Absent	Present
No_Visit	Absent	Smoker	Absent	False	Normal	Present	Present
No_Visit	Absent	NonSmoker	Absent	False	Normal	Present	Present
No_Visit	Absent	Smoker	Absent	False	Normal	Present	Absent
No_Visit	Absent	NonSmoker	Absent	False	Normal	Absent	Absent
No_Visit	Absent	Smoker	Absent	False	Normal	Absent	Absent
No_Visit	Absent	Smoker	Absent	False	Normal	Present	Present
No_Visit	Absent	NonSmoker	Absent	False	Normal	Present	Present
No_Visit	Absent	NonSmoker	Absent	False	Normal	Present	Present
No_Visit	Absent	NonSmoker	Absent	False	Normal	Absent	Absent
No_Visit	Absent	NonSmoker	Absent	False	Normal	Absent	Absent
No_Visit	Absent	NonSmoker	Absent	False	Normal	Absent	Present
No_Visit	Absent	Smoker	Absent	False	Normal	Present	Present
No_Visit	Absent	NonSmoker	Absent	False	Normal	Absent	Present
No_Visit	Absent	Smoker	Present	True	Abnormal	Present	Present
No_Visit	Absent	NonSmoker	Absent	False	Normal	Absent	Present
No_Visit	Absent	NonSmoker	Absent	False	Normal	Present	Present
No_Visit	Absent	Smoker	Absent	False	Normal	Absent	Absent
No_Visit	Absent	NonSmoker	Absent	False	Normal	Absent	Absent
No_Visit	Absent	Smoker	Absent	False	Normal	Absent	Absent
No_Visit	Absent	NonSmoker	Absent	False	Normal	Absent	Absent
No_Visit	Absent	NonSmoker	Absent	False	Normal	Present	Absent
No_Visit	Absent	Smoker	Absent	False	Normal	Present	Present
No_Visit	Absent	NonSmoker	Absent	False	Normal	Absent	Absent
No_Visit	Absent	Smoker	Absent	False	Normal	Absent	Absent
No_Visit	Absent	Smoker	Absent	False	Normal	Absent	Absent
No_Visit	Absent	Smoker	Absent	False	Normal	Absent	Absent
No_Visit	Absent	NonSmoker	Absent	False	Normal	Absent	Absent
No_Visit	Absent	Smoker	Absent	False	Normal	Present	Present
No_Visit	Absent	NonSmoker	Absent	False	Normal	Absent	Absent
No_Visit	Absent	NonSmoker	Absent	False	Normal	Absent	Absent
No_Visit	Absent	NonSmoker	Absent	False	Normal	Absent	Absent
No_Visit	Absent	NonSmoker	Absent	False	Normal	Absent	Absent
No_Visit	Absent	Smoker	Absent	False	Normal	Present	Present
No_Visit	Absent	Smoker	Absent	False	Abnormal	Absent	Absent
No_Visit	Absent	Smoker	Absent	False	Normal	Present	Absent
No_Visit	Absent	NonSmoker	Absent	False	Normal	Present	Present

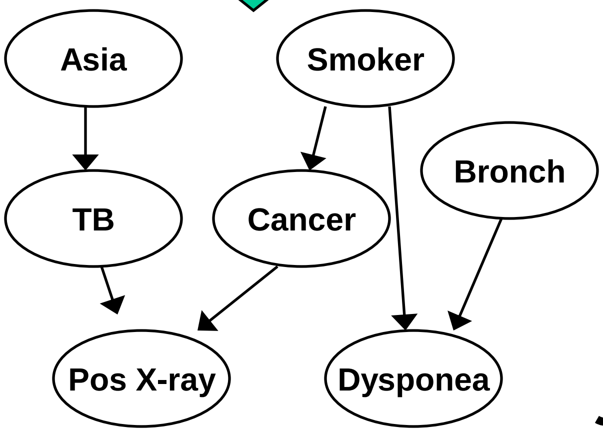


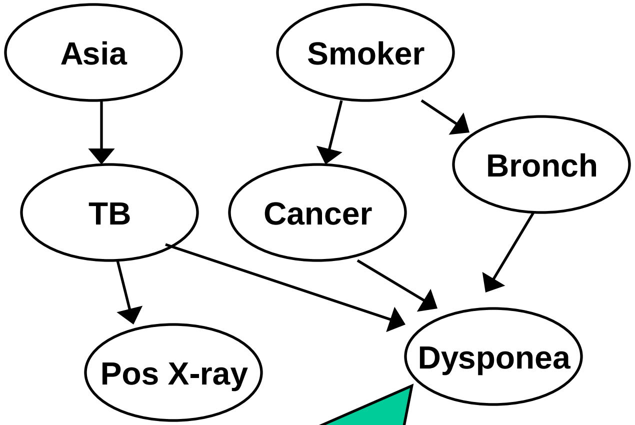


Has bronchitis	yes				no			
Has lung cancer	yes		no		yes		no	
Has tuberculosis	yes	no	yes	no	yes	no	yes	no
yes	0.25373134	0.56179774	0.5235602	0.73964494	0.45833334	0.6666667	0.7518797	0.6430868
no	0.74626863	0.43820223	0.4764398	0.26035503	0.5416667	0.33333334	0.24812031	0.35691318

Structure informed by experts

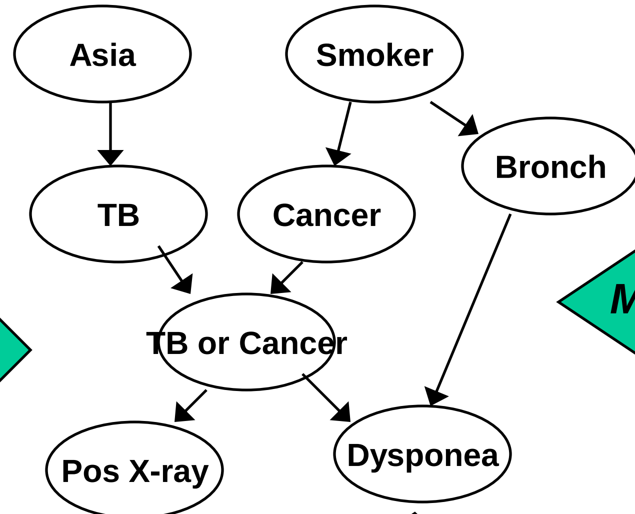
VisitAsia	Tuberculosis	Smoking	Cancer	TbOrCa	XRay	Bronchitis	Dyspnea
No_Visit	Absent	Smoker	Absent	False	Normal	Present	Present
No_Visit	Absent	Smoker	Absent	False	Normal	Absent	Absent
No_Visit	Absent	Smoker	Absent	False	Normal	Absent	Absent
No_Visit	Absent	NonSmoker	Absent	False	Normal	Absent	Present
No_Visit	Absent	Smoker	Absent	False	Normal	Absent	Absent
No_Visit	Absent	Smoker	Absent	False	Normal	Absent	Absent
No_Visit	Absent	NonSmoker	Absent	False	Normal	Present	Present
No_Visit	Absent	Smoker	Present	True	Abnormal	Present	Present
No_Visit	Absent	NonSmoker	Absent	False	Normal	Absent	Absent
No_Visit	Absent	NonSmoker	Absent	False	Normal	Absent	Absent
No_Visit	Absent	Smoker	Absent	False	Normal	Absent	Present
No_Visit	Absent	Smoker	Absent	False	Normal	Present	Present
No_Visit	Absent	NonSmoker	Absent	False	Normal	Present	Present
No_Visit	Absent	Smoker	Absent	False	Normal	Present	Absent
No_Visit	Absent	NonSmoker	Absent	False	Normal	Absent	Absent
No_Visit	Absent	Smoker	Absent	False	Normal	Absent	Absent
No_Visit	Absent	Smoker	Absent	False	Normal	Present	Present
No_Visit	Absent	NonSmoker	Absent	False	Normal	Present	Present
No_Visit	Absent	NonSmoker	Absent	False	Normal	Present	Present
No_Visit	Absent	NonSmoker	Absent	False	Normal	Absent	Absent
No_Visit	Absent	NonSmoker	Absent	False	Normal	Absent	Absent
No_Visit	Absent	NonSmoker	Absent	False	Normal	Absent	Present
No_Visit	Absent	Smoker	Absent	False	Normal	Present	Present
No_Visit	Absent	NonSmoker	Absent	False	Normal	Absent	Present
No_Visit	Absent	Smoker	Present	True	Abnormal	Present	Present
No_Visit	Absent	NonSmoker	Absent	False	Normal	Absent	Present
No_Visit	Absent	Smoker	Absent	False	Normal	Present	Present
No_Visit	Absent	Smoker	Absent	False	Normal	Absent	Absent
No_Visit	Absent	NonSmoker	Absent	False	Normal	Absent	Absent
No_Visit	Absent	Smoker	Absent	False	Normal	Absent	Absent
No_Visit	Absent	NonSmoker	Absent	False	Normal	Absent	Absent
No_Visit	Absent	NonSmoker	Absent	False	Normal	Present	Absent
No_Visit	Absent	Smoker	Absent	False	Normal	Present	Present
No_Visit	Absent	NonSmoker	Absent	False	Normal	Absent	Absent
No_Visit	Absent	Smoker	Absent	False	Normal	Absent	Absent
No_Visit	Absent	Smoker	Absent	False	Normal	Absent	Absent
No_Visit	Absent	Smoker	Absent	False	Normal	Absent	Absent
No_Visit	Absent	NonSmoker	Absent	False	Normal	Absent	Absent
No_Visit	Absent	Smoker	Absent	False	Normal	Present	Present
No_Visit	Absent	NonSmoker	Absent	False	Normal	Absent	Absent
No_Visit	Absent	NonSmoker	Absent	False	Normal	Absent	Absent
No_Visit	Absent	NonSmoker	Absent	False	Normal	Absent	Absent
No_Visit	Absent	NonSmoker	Absent	False	Normal	Absent	Absent
No_Visit	Absent	Smoker	Absent	False	Normal	Present	Present
No_Visit	Absent	Smoker	Absent	False	Abnormal	Absent	Absent
No_Visit	Absent	Smoker	Absent	False	Normal	Present	Absent
No_Visit	Absent	NonSmoker	Absent	False	Normal	Present	Present

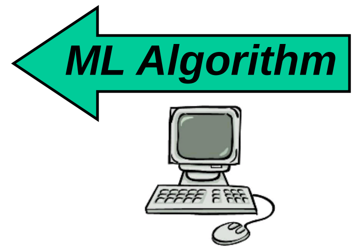




Has bronchitis	yes				no			
Has lung cancer	yes		no		yes		no	
Has tuberculosis	yes	no	yes	no	yes	no	yes	no
yes	0.25373134	0.56179774	0.5235602	0.73964494	0.45833334	0.6666667	0.7518797	0.6430868
no	0.74626863	0.43820223	0.4764398	0.26035503	0.5416667	0.33333334	0.24812031	0.35691318

Structure and tables by experts





VisitAsia	Tuberculosis	Smoking	Cancer	TbOrCa	XRay	Bronchitis	Dyspnea
No_Visit	Absent	Smoker	Absent	False	Normal	Present	Present
No_Visit	Absent	Smoker	Absent	False	Normal	Absent	Absent
No_Visit	Absent	Smoker	Absent	False	Normal	Absent	Absent
No_Visit	Absent	NonSmoker	Absent	False	Normal	Absent	Present
No_Visit	Absent	Smoker	Absent	False	Normal	Absent	Absent
No_Visit	Absent	Smoker	Absent	False	Normal	Absent	Absent
No_Visit	Absent	NonSmoker	Absent	False	Normal	Present	Present
No_Visit	Absent	Smoker	Present	True	Abnormal	Present	Present
No_Visit	Absent	NonSmoker	Absent	False	Normal	Absent	Absent
No_Visit	Absent	NonSmoker	Absent	False	Normal	Absent	Absent
No_Visit	Absent	Smoker	Absent	False	Normal	Absent	Present
No_Visit	Absent	Smoker	Absent	False	Normal	Present	Present
No_Visit	Absent	NonSmoker	Absent	False	Normal	Present	Present
No_Visit	Absent	Smoker	Absent	False	Normal	Present	Absent
No_Visit	Absent	NonSmoker	Absent	False	Normal	Absent	Absent
No_Visit	Absent	Smoker	Absent	False	Normal	Absent	Absent
No_Visit	Absent	Smoker	Absent	False	Normal	Present	Present
No_Visit	Absent	NonSmoker	Absent	False	Normal	Present	Present
No_Visit	Absent	NonSmoker	Absent	False	Normal	Present	Present
No_Visit	Absent	NonSmoker	Absent	False	Normal	Absent	Absent
No_Visit	Absent	NonSmoker	Absent	False	Normal	Absent	Absent
No_Visit	Absent	NonSmoker	Absent	False	Normal	Absent	Present
No_Visit	Absent	Smoker	Absent	False	Normal	Present	Present
No_Visit	Absent	NonSmoker	Absent	False	Normal	Absent	Present
No_Visit	Absent	Smoker	Present	True	Abnormal	Present	Present
No_Visit	Absent	NonSmoker	Absent	False	Normal	Absent	Present
No_Visit	Absent	Smoker	Absent	False	Normal	Present	Present
No_Visit	Absent	Smoker	Absent	False	Normal	Absent	Absent
No_Visit	Absent	NonSmoker	Absent	False	Normal	Absent	Absent
No_Visit	Absent	Smoker	Absent	False	Normal	Absent	Absent
No_Visit	Absent	NonSmoker	Absent	False	Normal	Absent	Absent
No_Visit	Absent	NonSmoker	Absent	False	Normal	Present	Absent
No_Visit	Absent	Smoker	Absent	False	Normal	Present	Present
No_Visit	Absent	NonSmoker	Absent	False	Normal	Absent	Absent
No_Visit	Absent	Smoker	Absent	False	Normal	Absent	Absent
No_Visit	Absent	Smoker	Absent	False	Normal	Absent	Absent
No_Visit	Absent	Smoker	Absent	False	Normal	Absent	Absent
No_Visit	Absent	NonSmoker	Absent	False	Normal	Absent	Absent
No_Visit	Absent	Smoker	Absent	False	Normal	Present	Present
No_Visit	Absent	NonSmoker	Absent	False	Normal	Absent	Absent
No_Visit	Absent	NonSmoker	Absent	False	Normal	Absent	Absent
No_Visit	Absent	NonSmoker	Absent	False	Normal	Absent	Absent
No_Visit	Absent	NonSmoker	Absent	False	Normal	Absent	Absent
No_Visit	Absent	Smoker	Absent	False	Normal	Present	Present
No_Visit	Absent	Smoker	Absent	False	Abnormal	Absent	Absent
No_Visit	Absent	Smoker	Absent	False	Normal	Present	Absent
No_Visit	Absent	NonSmoker	Absent	False	Normal	Present	Present

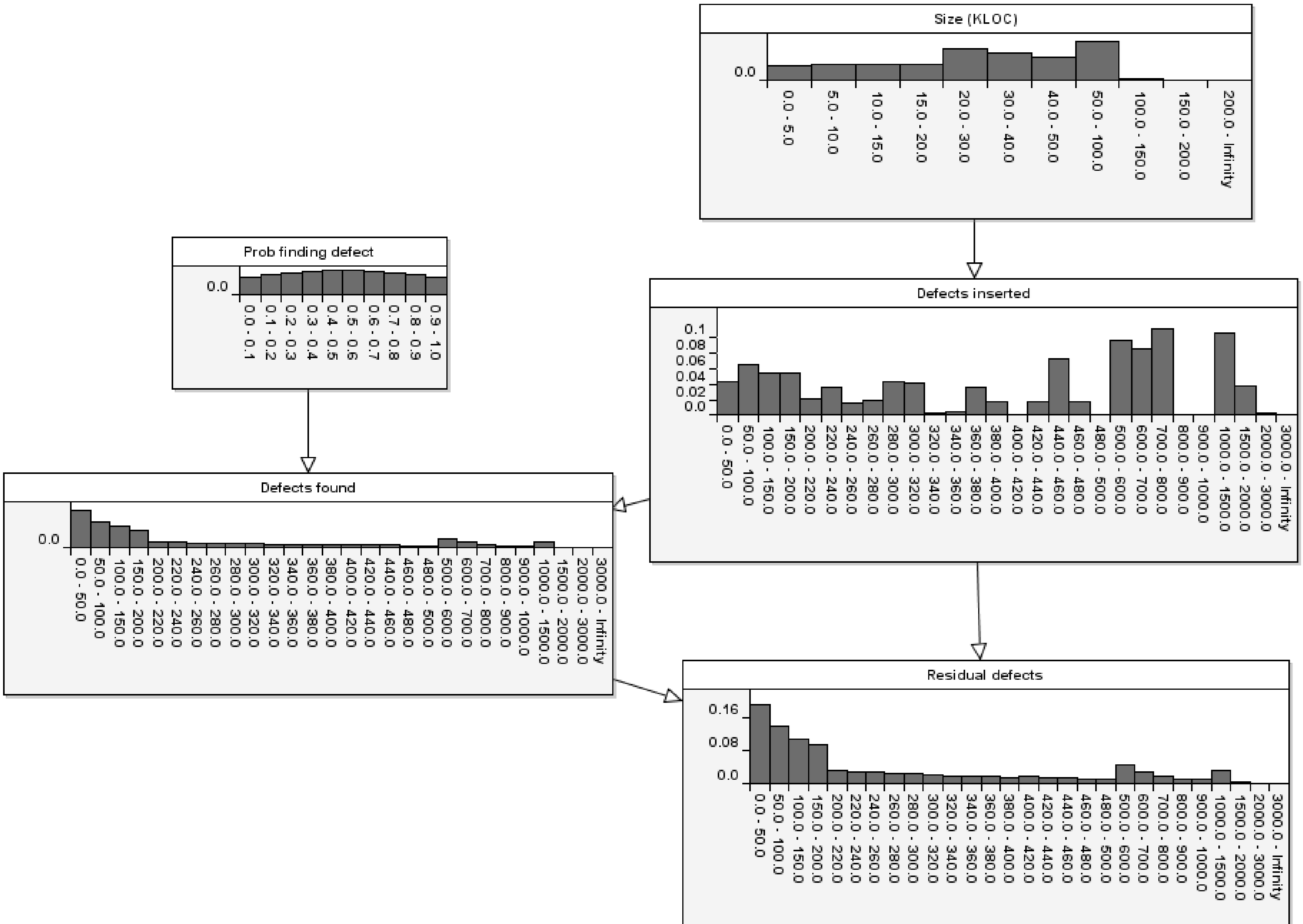
Has bronchitis	yes		no	
Tuberculosi...	yes	no	yes	no
yes	0.875	0.7619048	0.68085104	0.1122449
no	0.125	0.23809524	0.31914896	0.8877551

Handling continuous nodes

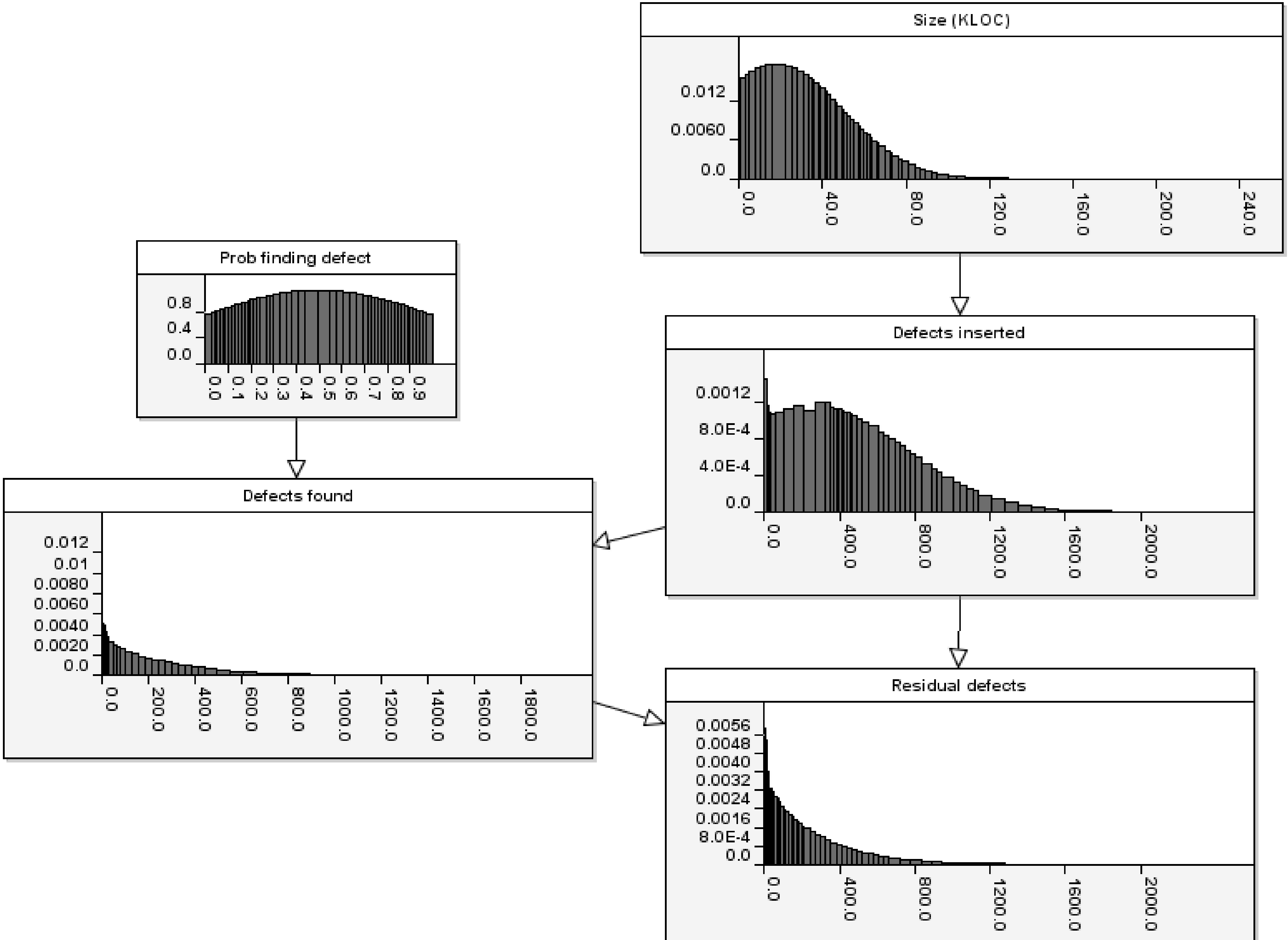
Static discretisation: inefficient and devastatingly inaccurate

Developments in dynamic discretisation will have revolutionary effect

Static discretization



Dynamic discretization



Typical Applications

Predicting reliability of critical systems

Typical Applications

Software defect prediction

Typical Applications

Aircraft accident traffic risk

Civil Aviation Authority

NATS

Typical Applications

Warranty return rates of electronic parts

Typical Applications

Operational risk in financial institutions

Typical Applications

Hazards in petrochemical industry

Typical Applications

R vs Levi Bellfield

Probabilistic and risk based legal arguments

Conclusions

Genuine risk assessment requires causal Bayesian networks

Bayesian networks have been used effectively in a range of real world problems.

Major remaining barrier to widespread use is conceptual/presentational

www.BayesianRisk.com

www.AgenaRisk.com