28th International Joint Conference on Artificial Intelligence (IJCAI-19)

Macao, August 10-16, 2019

Tutorial:

Medical decision analysis with probabilistic graphical models

Francisco Javier Díez

Dept. Artificial Intelligence. UNED
Madrid, Spain

www.ia.uned.es/~fjdiez
www.cisiad.uned.es

OVERVIEW

1. Introduction: AI in medicine
2. Probabilistic diagnosis
3. Bayesian networks
4. Unicriterion decision analysis
5. Multicriteria decision analysis
6. Sensitivity analysis
7. Temporal models
8. Conclusion

1. Introduction: history of probabilistic AI in medicine

Probability in artificial intelligence

- A.I. was “born” in 1956, at the Dartmouth Conference
- In the first 25 or 30 years, many researchers questioned that probability could play a significant role in A.I.
- First reason (cf. [Sutton and Barto, 1998]):
  - Computers were already good at arithmetic operations
  - but could not perform “easy” tasks (easy for a little child): vision (image understanding), natural language, planning…
  - Those tasks could not be solved with arithmetic operations; they require conceptual reasoning (symbol manipulation $\rightarrow$ LISP).
  - Probabilistic “reasoning” consisted mainly in number crunching, not in conceptual reasoning.
- Second reason: limitations of probabilistic methods.

ARTIFICIAL INTELLIGENCE **11** (1978) 115-144

Categorical and Probabilistic Reasoning in Medical Diagnosis*

Peter Szolovits, Ph.D.

Clinical Decision Making Group, Laboratory for Computer Science, Massachusetts Institute of Technology, Cambridge, MA 02139, U.S.A.

Stephen G. Pauker, M.D. **

Department of Medicine, New England Medical Center Hospital, Tufts University School of Medicine, Boston, MA 02111, U.S.A.

Recommended by N. S. Sridharan

ABSTRACT

Medical decision making can be viewed along a spectrum, with categorical (or deterministic) reasoning at one extreme and probabilistic (or evidential) reasoning at the other. In this paper we examine the flowchart as the prototype of categorical reasoning and decision analysis as the prototype of probabilistic reasoning. Within this context we compare PIP, INTERNIST, CASNET, and MYCIN—four of the present programs which apply the techniques of artificial intelligence to medicine. Although these systems can exhibit impressive expert-like behavior, we believe that none of them is yet capable of truly expert reasoning. We suggest that a program which can demonstrate expertise in the area of medical consultation will have to use a judicious combination of categorical and probabilistic reasoning—the former to establish a sufficiently narrow context and the latter to make comparisons among hypotheses and eventually to recommend therapy.

Limitations of probability for AI in medicine

P. Szolovits. *Artificial Intelligence in Medicine*. Westview Press, 1982.

"The chief **disadvantages** of the decision theoretic approach are the difficulties of obtaining reasonable estimates of probabilities and utilities for a particular analysis. Although techniques such as sensitivity analysis help greatly to indicate which potential inaccuracies are unimportant, the lack of adequate data often forces artificial simplifications of the problem and lowers confidence in the outcome of the analysis. Attempts to extend these techniques to large medical domains in which multiple disorders may co-occur, temporal progressions of findings may offer important diagnostic clues, or partial effects of therapy can be used to guide further diagnostic reasoning, **have not been successful**. The typical language of probability and utility theory is not rich enough to discuss such issues, and its extension within the original spirit leads to untenably large decision problems. […]

A second difficulty for decision analysis is the relatively mysterious reasoning of a decision theoretic program—an explanation of the results is to he understood in terms of the numeric manipulations involved in expected value computations, which is not a natural way of thinking for most people."

Historic evolution of probabilistic AI

- 1960s and 1970s: naïve-Bayes diagnostic systems
  - able to diagnose better than physicians in restricted problems
- expert system Prospector (Hart and Duda, 1977)
  - used approximate Bayesian reasoning
  - found a molybdenum deposit valued in $1,000,000
  - was the first commercial success of A.I.
- Bayesian networks (Pearl, 1982, 1986, 1988)
  - overcame the limitations of the naïve Bayes
- Nowadays: probabilistic graphical models (PGMs) are used more and more in A.I.
  - tasks: diagnosis, planning, learning (incl. deep learning)…
  - fields: medicine, robotics, computer vision, e-commerce…

2. Probabilistic diagnosis

2.1. Basic concepts of probabilistic diagnosis

Probabilistic diagnosis with one finding

- Example:
  - Prevalence of a disease: 14%
  - Sensitivity of a test: 70%
  - Specificity of the test: 91%

- Questions:
  - What is the positive predictive value (PPV)?
    - If the test is positive, what is the probability that the patient has the disease?
  - What is the negative predictive value (NPV)?
    - If the test is negative, what is the probability that the patient does not have the disease?

Basic concepts for medical diagnosis

- Disease *E*, result of a test *T*
- Parameters of the model
  - Prevalence: $P(+e)$
  - Sensitivity: $P(+t|+e)$
  - Specificity: $P(\neg t|\neg e)$
- Predictive values:
  - Positive PV: $P(+e|+t)$
  - Negative PV: $P(\neg e|\neg t)$

2.2. Bayes theorem

Bayes theorem

- We knew that

$$P(x|y) = \frac{P(x,y)}{P(y)}$$ by the definition of $P(x|y)$

$$P(x,y) = P(x)\cdot P(y\,|\,x)$$ by the definition of $P(y|x)$

$$P(y) = \sum_x P(y|x)\cdot P(x)$$ by the theorem of total prob.

- Combining these results:

$$P(x|y) = \frac{P(x,y)}{P(y)} = \frac{P(x)\cdot P(y|x)}{P(y)} = \frac{P(x)\cdot P(y|x)}{\sum_{x'} P(x')\cdot P(y|x')}$$

- It means that knowing $P(x)$ and $P(y|x)$ we compute $P(x|y)$.

Predictive value of a finding

- Positive predictive value: $P(+e|+h)$

$$P(+e|+h) = \frac{P(+e)\cdot P(+h|+e)}{P(+e)\cdot P(+h|+e) + P(\neg e)\cdot P(+h|\neg e)}$$

$$\text{PPV} = \frac{\text{prev}\cdot\text{sens}}{\text{prev}\cdot\text{sens} + (1-\text{prev})\cdot(1-\text{spec})}$$

- Negative predictive value: $P(\neg e|+h)$

$$P(\neg e|\neg h) = \frac{P(\neg e)\cdot P(\neg h|\neg e)}{P(+e)\cdot P(\neg h|+e) + P(\neg e)\cdot P(\neg h|\neg e)}$$

$$\text{NPV} = \frac{(1-\text{prev})\cdot\text{spec}}{\text{prev}\cdot(1-\text{sens}) + (1-\text{prev})\cdot\text{spec}}$$

Probabilistic diagnosis with two findings

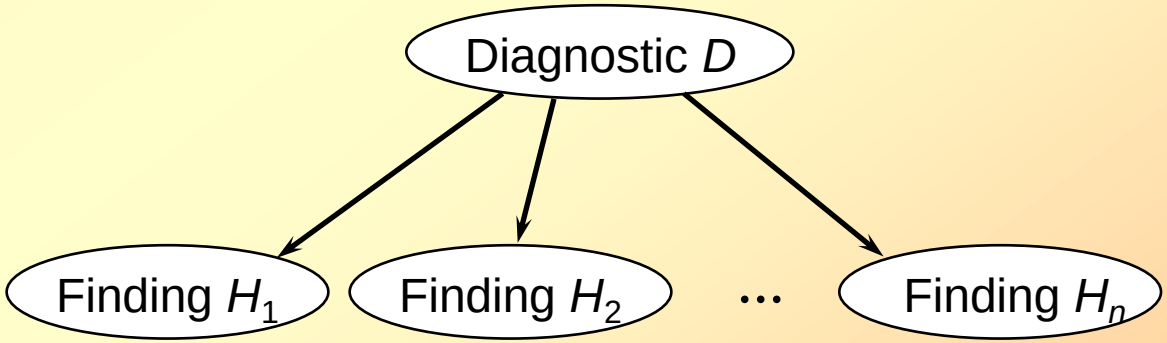
- Example:
  - Prevalence of the disease: 14%
  - Sensitivity of test C: 70%
  - Specificity of test C: 91%
  - Sensitivity of test E: 90%
  - Specificity of test E: 93%
- Questions:
  - What is the posterior probability for each combination of findings?

2.3. The naïve Bayes method

The naïve Bayes method

- Two hypotheses:
  - Diagnostics are mutually exclusive
    $\Rightarrow$ every patient has at most one disease
  - Findings are conditionally independent given the diagnostics

- Graphical representation:



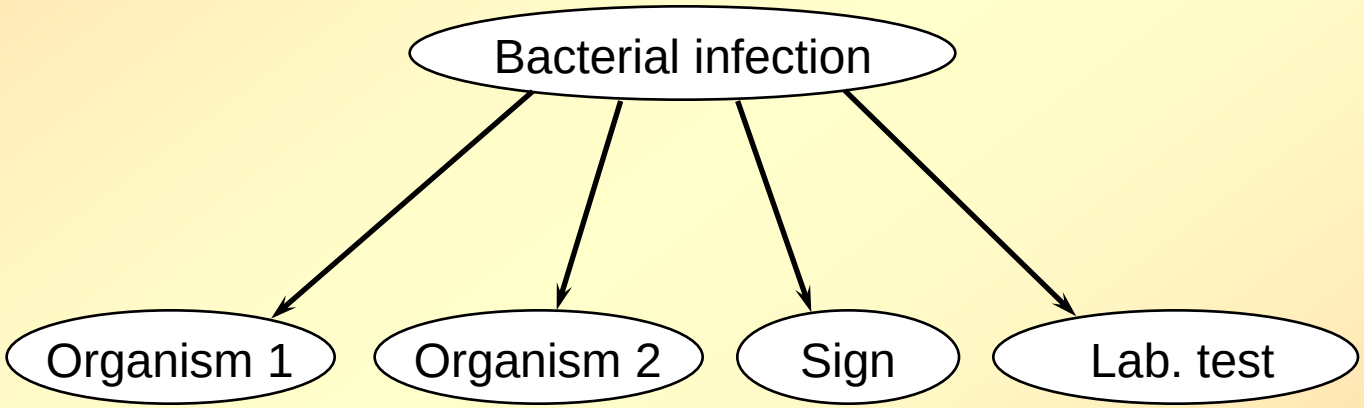
Succesfull applications of the naïve-Bayes

- Lodwick GS, Haun CL, Smith WE, et al. , “Computer diagnosis of primary bone tumors: A preliminary report” *Radiology* **80** (1963) 273-275.
- Overall JE, Williams CM, “Conditional probability program for diagnosis or thyroid function” *JAMA* **183** (1963) 307-313.
- Toronto AF, Veasy LG, Warner HR, “Evaluation of a computer program for diagnosis of congenital heart disease” *Progress in Cardiovascular Diseases* **5** (1963) 362-377.

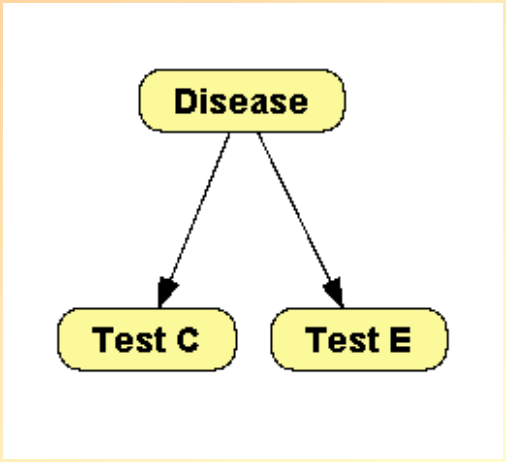
  Warner HR, Toronto AF, Veasy LG, “Experience with Bayes’ theorem for computer diagnosis of congenital heart disease” *Annals New York Acad. Sciences* **115** (1964) 558-567.
- de Dombal FT, Leaper JR, Staniland JR, et al., “Computer-aided diagnosis of acute abdominal pain” *BMJ* **2** (1972) 9-13.
- Gorry GA, Kassirer JP, Essig A, Schwartz WB, “Decision analysis as the basis for computer-aided management of acute renal failure” *Amer. J Med* **55** (1973) 473-484.
- Gorry GA, Silverman H, Pauker SG, “Capturing clinical expertise: A computer program that considers clinical responses to digitalis” *Amer. J. Med* **64** (1978) 452-460.

Limitations of the naïve Bayes

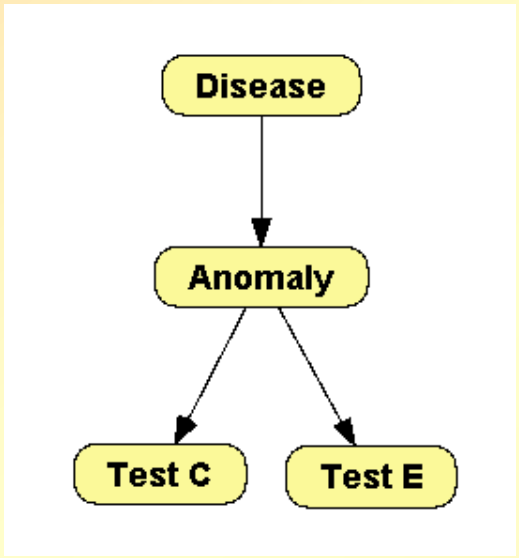
- In general the diagnostics are not mutually exclusive.
- In general, findings are not conditionally independent.



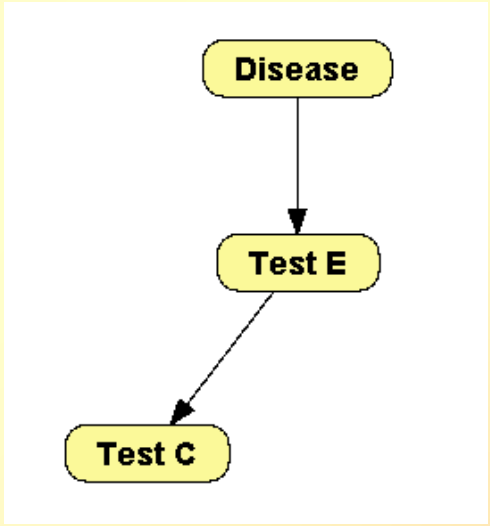
Three cases



Test results are conditionally independent given the disease



Correlation, even when the disease is present or absent



Test C is conditionally independent of the disease given test E

In the three cases the sensitivity and specificity of the tests (wrt the disease) are the same, but the posterior probabilities are different

Impact of correlation on the posterior prob.

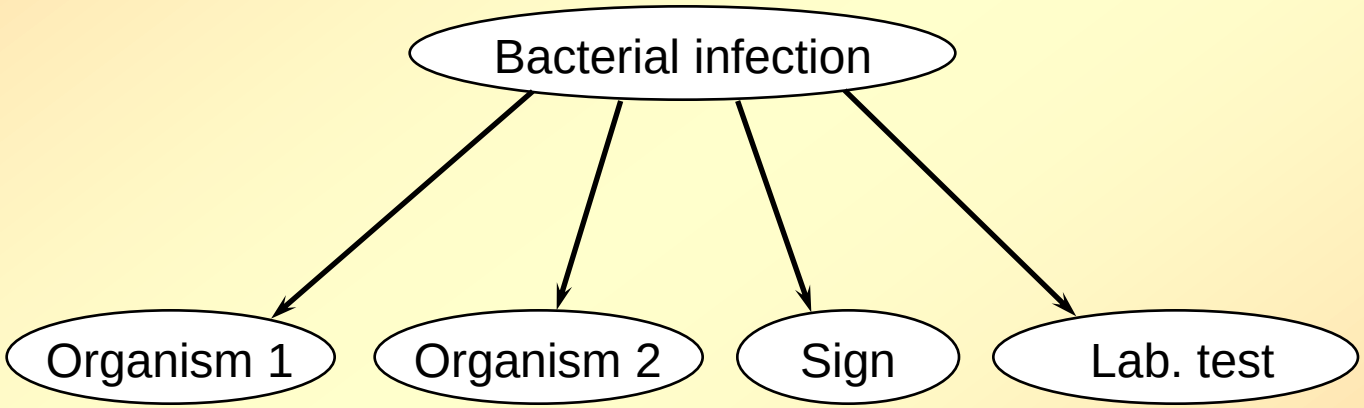
- Extreme case:
  test results are conditionally independent given the disease
  - $P(+d \mid +c, +e) = 0.9421$
  - maximum increase in the posterior probability
- Opposite extreme case:
  test C is conditionally independent of the disease given test E:
  - $P(+d \mid +c, +e) = 0.5587 = P(+d \mid +e)$
  - no increase in the posterior probability ≡ no new information
- Intermediate cases:
  correlation among findings
  - $0.5587 < P(+d \mid +c, +e) < 0.9421$
  - the bigger the correlation, the smaller the information contributed.

Prob. diagnosis with two findings (revisited)

- Example:
  - Prevalence of the disease: 14%
  - Sensitivity of test C: 70%
  - Specificity of test C: 91%
  - Sensitivity of test E: 90%
  - Specificity of test E: 93%
- Questions:
  - What is the posterior probability for each combination of findings?
- The problem is ill-specified
  - The solution depends on the correlation among findings

Limitations of the naïve Bayes

- In general the diagnostics are not mutually exclusive.
- In general, findings are not conditionally independent.



- These limitations are rarely discussed in the books of medical decision analysis and evidence-based medicine.
- This is the only method presented in those books.

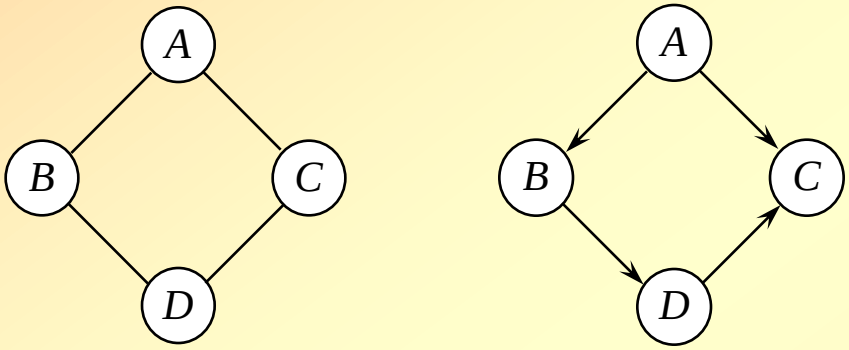
3. Bayesian networks

Probabilistic graphical models

- Elements of a PGM
  - Qualitative component (**structure**): a graph
    - Links usually represent causal relations
  - Quantitative components (**parameters**): potentials
    - A conditional probability for each chance node
    - A value function for each value node
- Relation between the graph and the prob. distribution
  - Every node in the graph represents a variable of the prob.
  - The graph represents the dependencies of the prob. distr.

3.1. Definition of BN

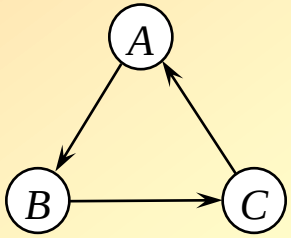
Notions about graphs

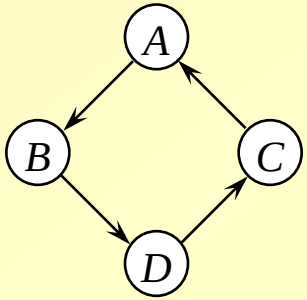


- Basic concepts
  - Definition: a set of nodes and links (vertices and edges)
  - Two types of links: directed / undirected
  - Open path (*A-B*, *A-B-C-D*), closed path (*A-B-C-D-A*),
- In directed graphs:
  - parent, child, ancestor, descendant.

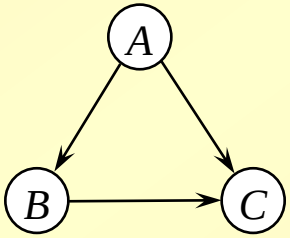
Directed graphs: cycles and loops

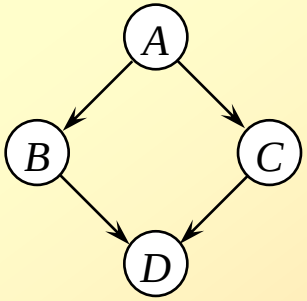
- Cycles

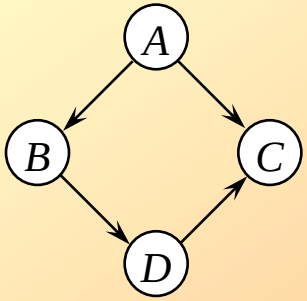




- Loops







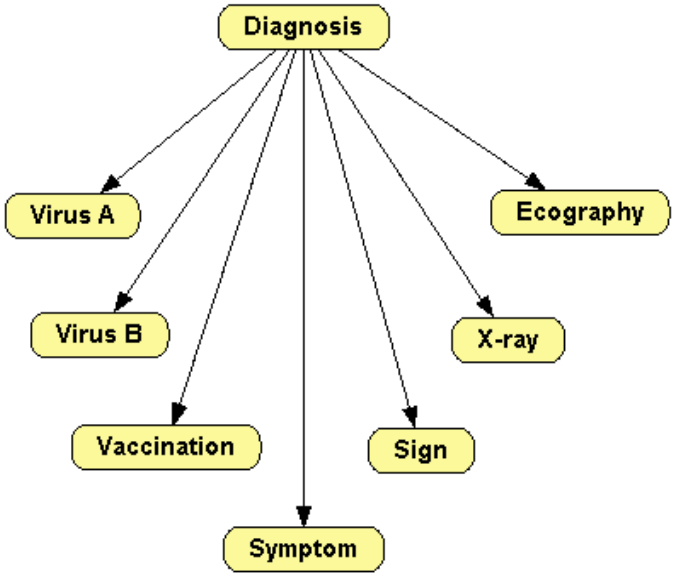
Definition of Bayesian network

- Elements:
  - a set of variables $\{X_i\}$
  - an acyclic directed graph
    - every node in the graph represents a variable X_i
  - a conditional probability distribution (usually a table) for each variable: $P(x_i \mid pa(x_i))$
    - for a node without parents: $P(x_i \mid pa(x_i)) = P(x_i)$

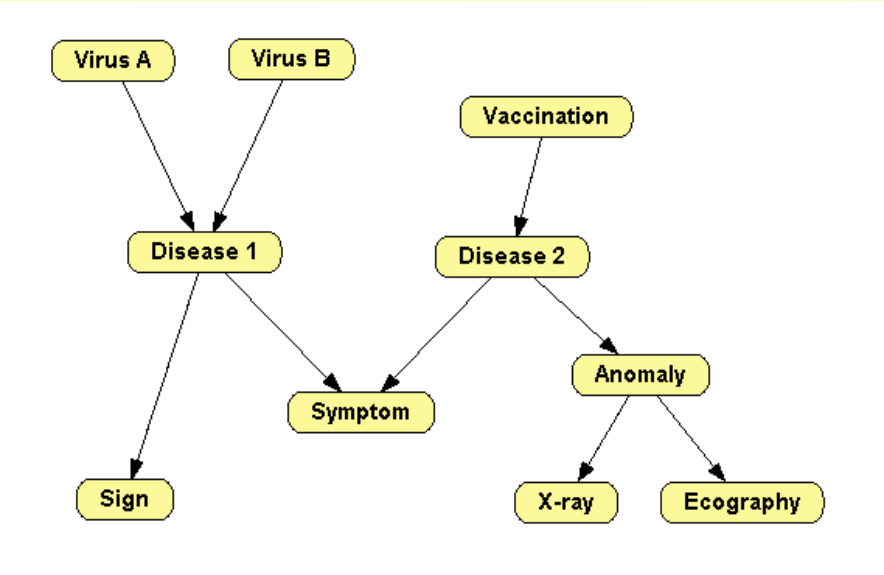
- Result: join probability for the network

$$P(x_1, \cdots, x_n) = \prod_{i=1}^{n} P(x_i \mid pa(x_i))$$

Naïve Bayes



Bayesian network

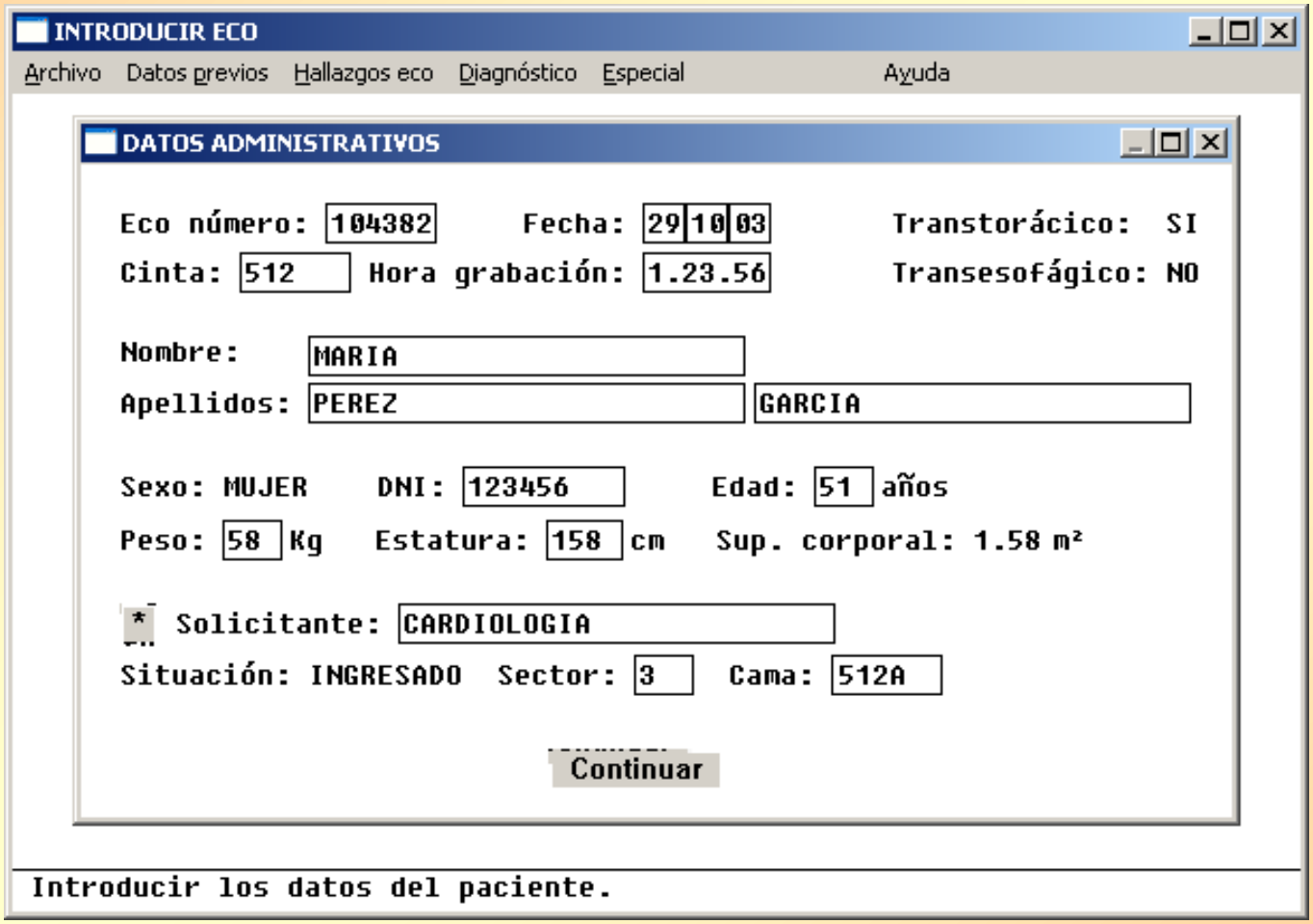


3.2. Examples of BNs

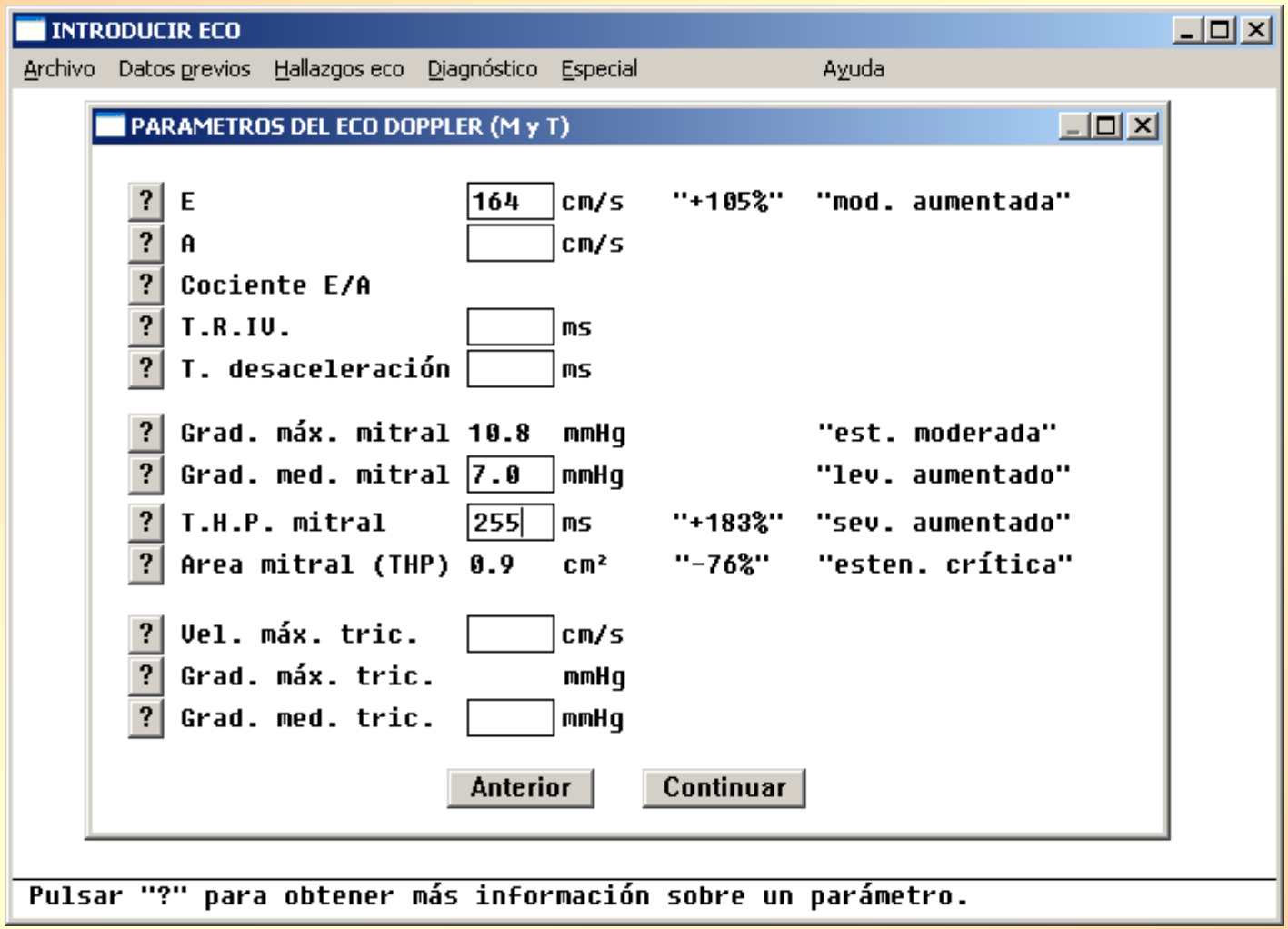
Examples of BNs

- Medical Bayesian networks we have built
  - DIAVAL: echocardiography (valvulopathies)
    F. J. Díez' thesis, 1994
  - Prostanet: urology (prostate cancer)
    Carmen Lacave's thesis, 2003
  - Nasonet: nasopharyngeal cancer spread
    Severino Galán's thesis, 2003
  - HEPAR II: liver diseases
    Agnieszka Onisko's thesis, 2003
  - Catarnet: Cataract surgery
    Nuria Alonso's thesis, 2009

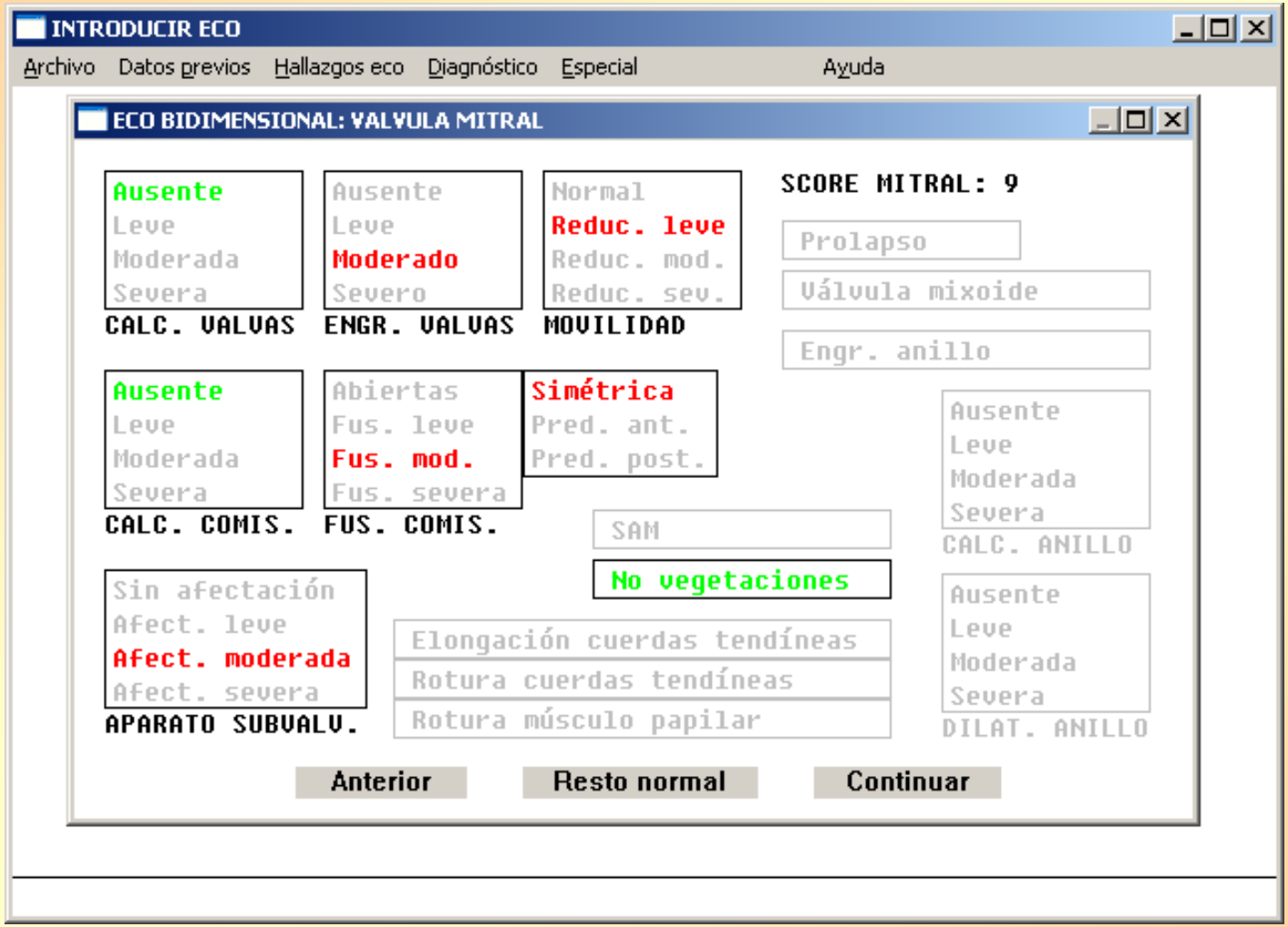
DIAVAL


INTRODUCIR ECO
Archivo
Datos previos
Hallazgos eco
Diagnóstico
Especial
Ayuda
DATOS ADMINISTRATIVOS
Eco número: 104382
Fecha: 29 10 03
Transtorácico: SI
Cinta: 512
Hora grabación: 1.23.56
Transesofágico: NO
Nombre: MARIA
Apellidos: PEREZ
GARCIA
Sexo: MUJER
DNI: 123456
Edad: 51 años
Peso: 58 Kg
Estatura: 158 cm
Sup. corporal: 1.58 m²
* Solicitante: CARDIOLOGIA
Situación: INGRESADO
Sector: 3
Cama: 512A
Continuar
Introducir los datos del paciente.

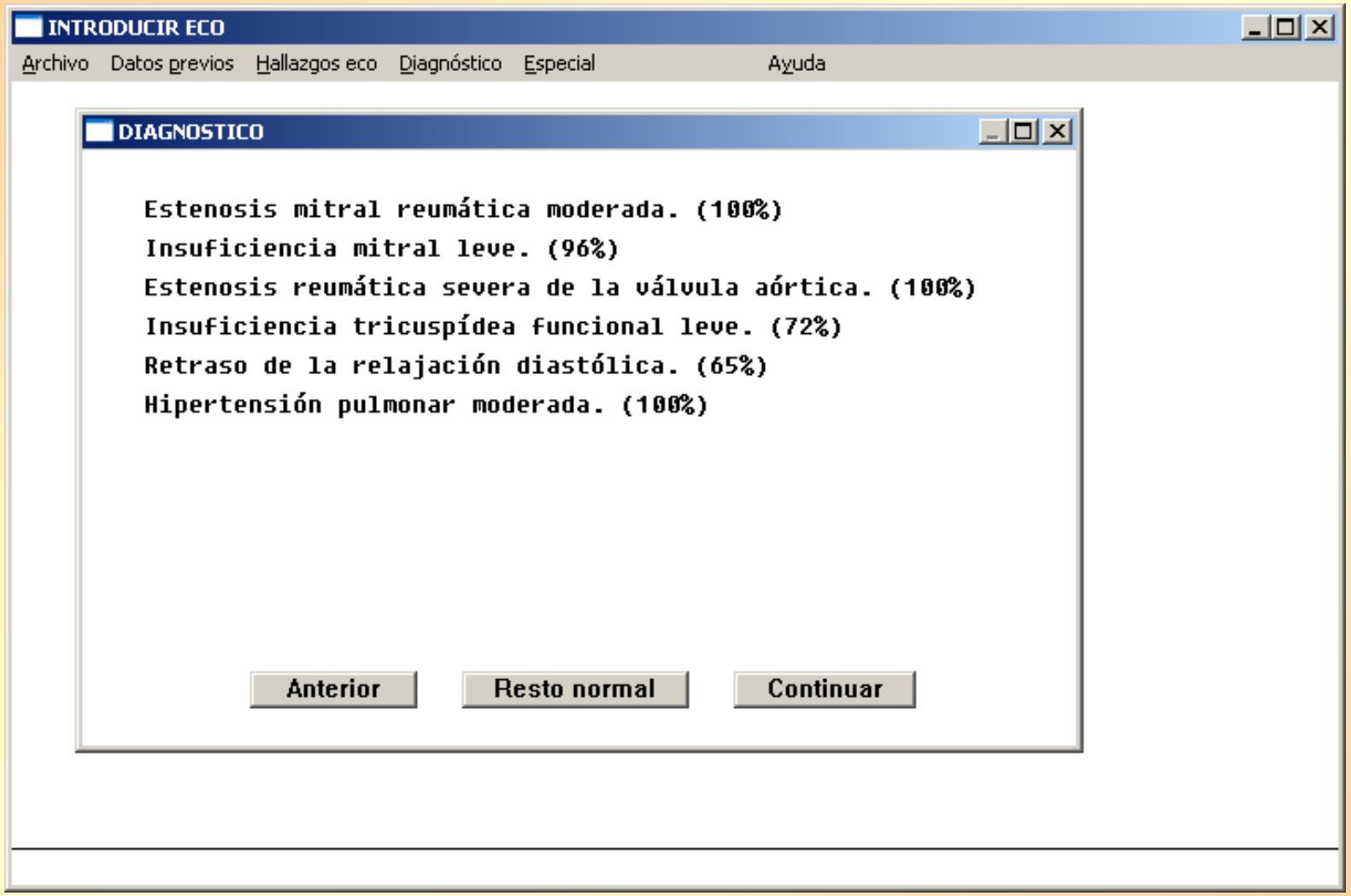
DIAVAL: numeric findings



DIAVAL: qualitative findings



DIAVAL: diagnostics



DIAVAL: final report

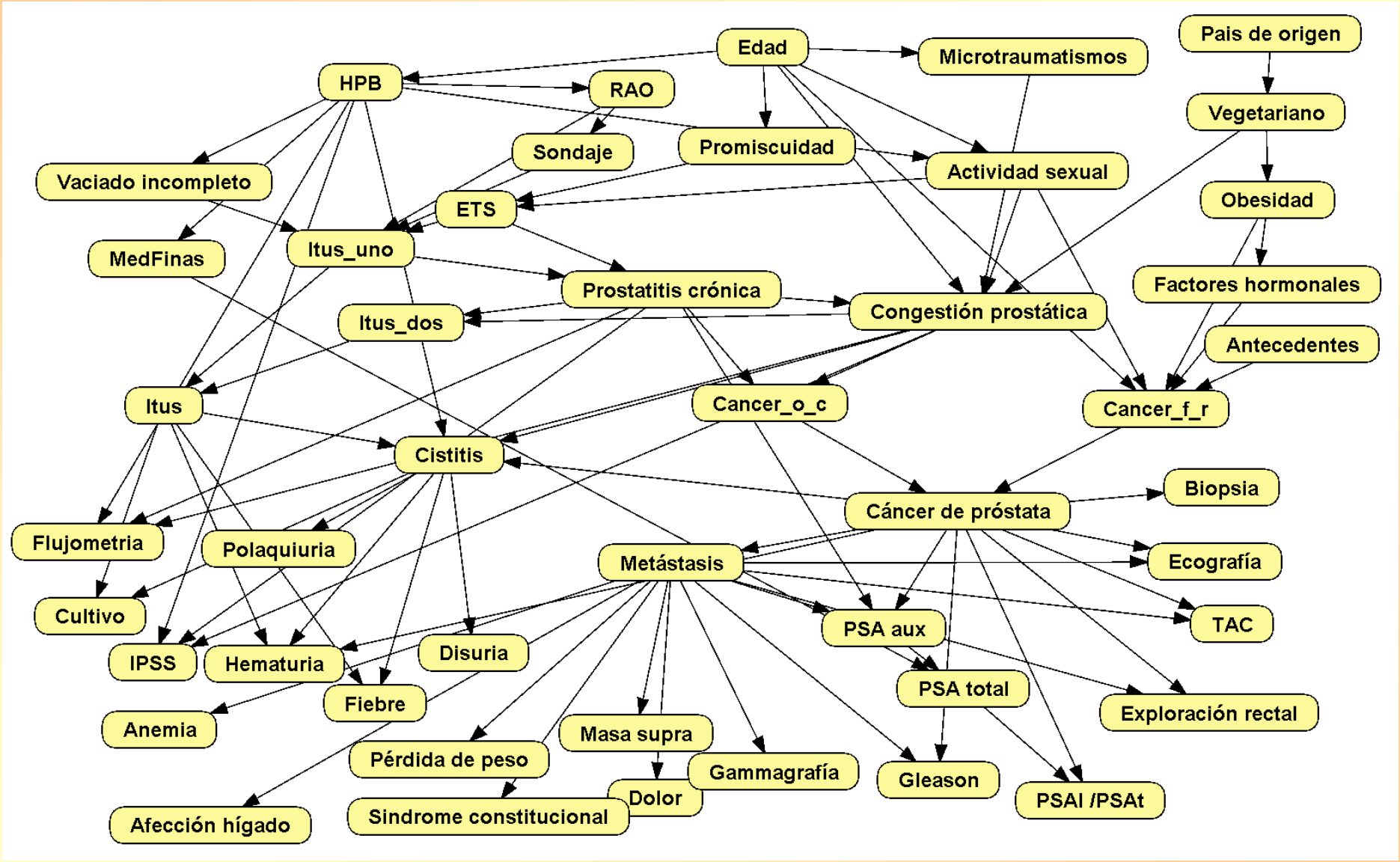
in a text editor

INFORME.TXT - Bloc de notas

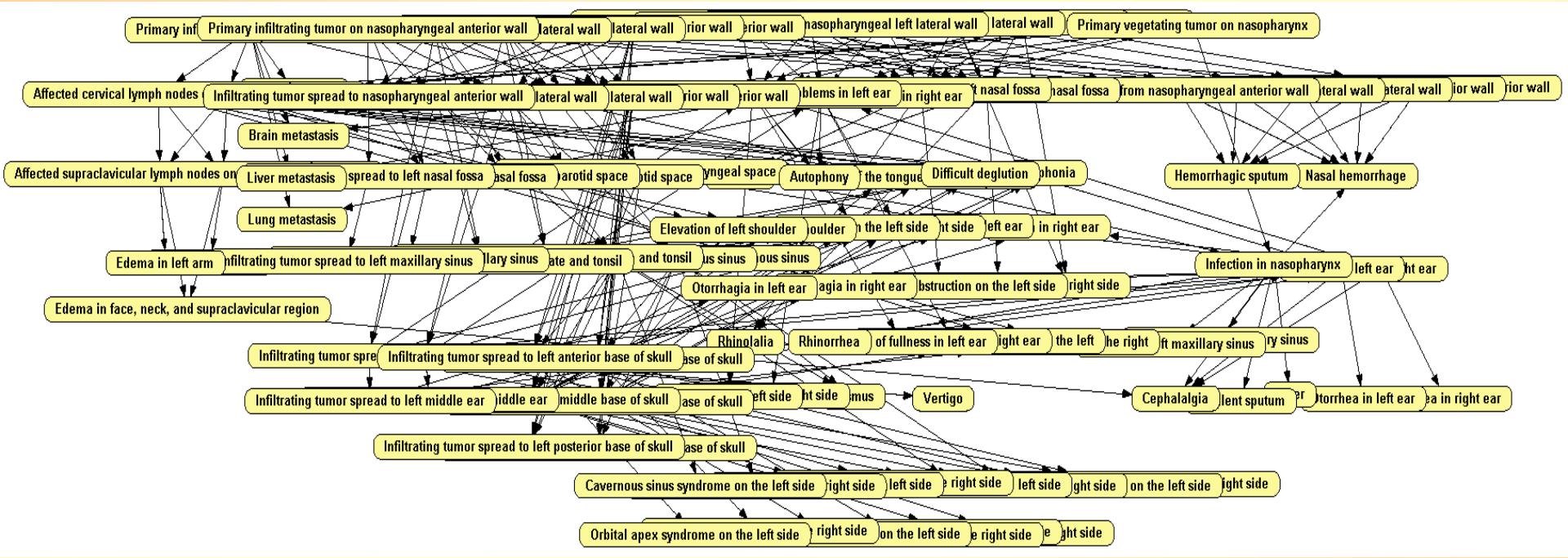
Archivo Edición Formato Ver Ayuda

```
Datos administrativos
=====================
  Nº eco: 104382.  Fecha: 29/10/3.  Cinta: 512.
  MARIA GARCIA PEREZ.  DNI: 1234567.
  Edad: 51 años.  Mujer.
  Peso: 58 Kg.  Estatura: 158 cm.  Sup. corporal: 1.58 m².
  Solicitante: CARDIOLOGIA.
  Ingresada, sector 3, cama 512A.

Síntomas
========
  Disnea de grado II.

Válvula mitral
==============
  Area (eco 2D): 1.2 cm².
  Velocidad onda E: 164 cm/s.
  Gradiente máximo: 10.8 mmHg.
  Gradiente medio: 7.0 mmHg.
  Tiempo de hemipresión: 255 ms.
  Area (THP): 0.9 cm².
  Engrosamiento moderado de las valvas.
  No calcificación de las valvas mitrales.
  Reducción leve de la movilidad.


[...]


Contractilidad segmentaria normal.

Pericardio normal.

DIAGNOSTICO
===========
  Estenosis mitral reumática moderada.
  Insuficiencia mitral leve.
  Estenosis reumática severa de la válvula aórtica.
  Insuficiencia tricuspídea funcional leve.
  Retraso de la relajación diastólica.
  Hipertensión pulmonar moderada.


                    Dra. Elena Iturralde  |
```

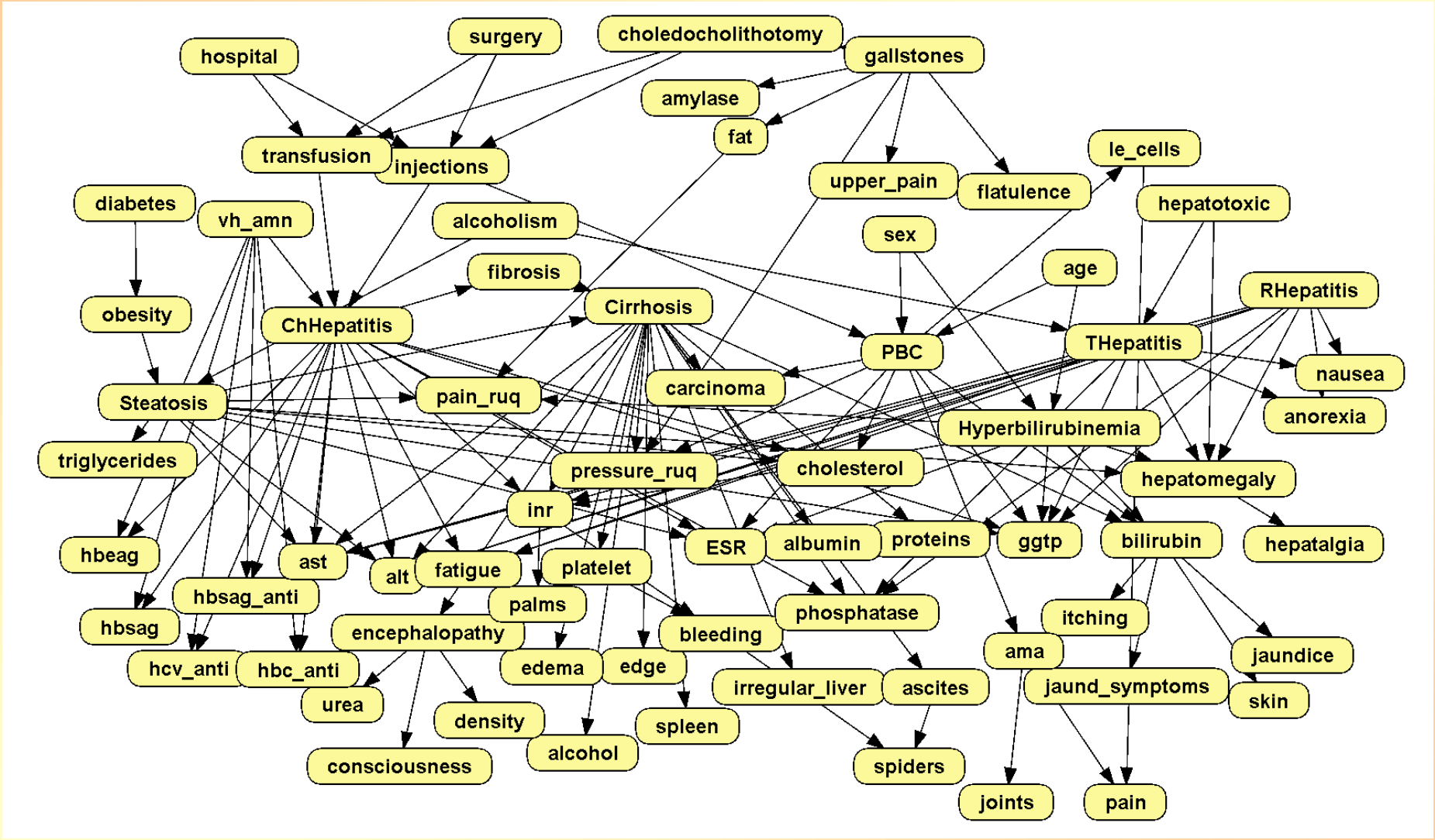
Prostanet (for prostate diseases)



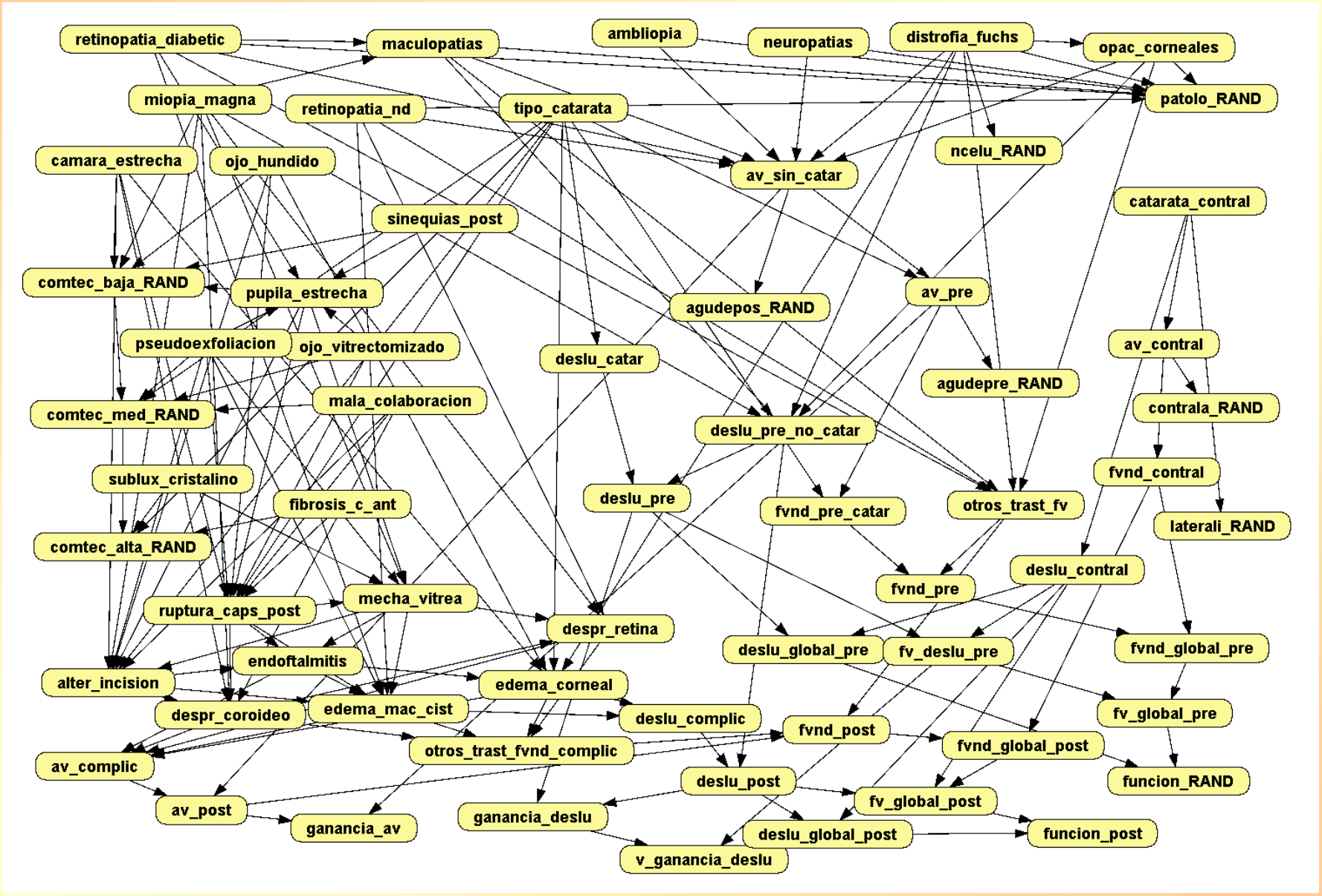
Nasonet (nasopharyngeal cancer spread)



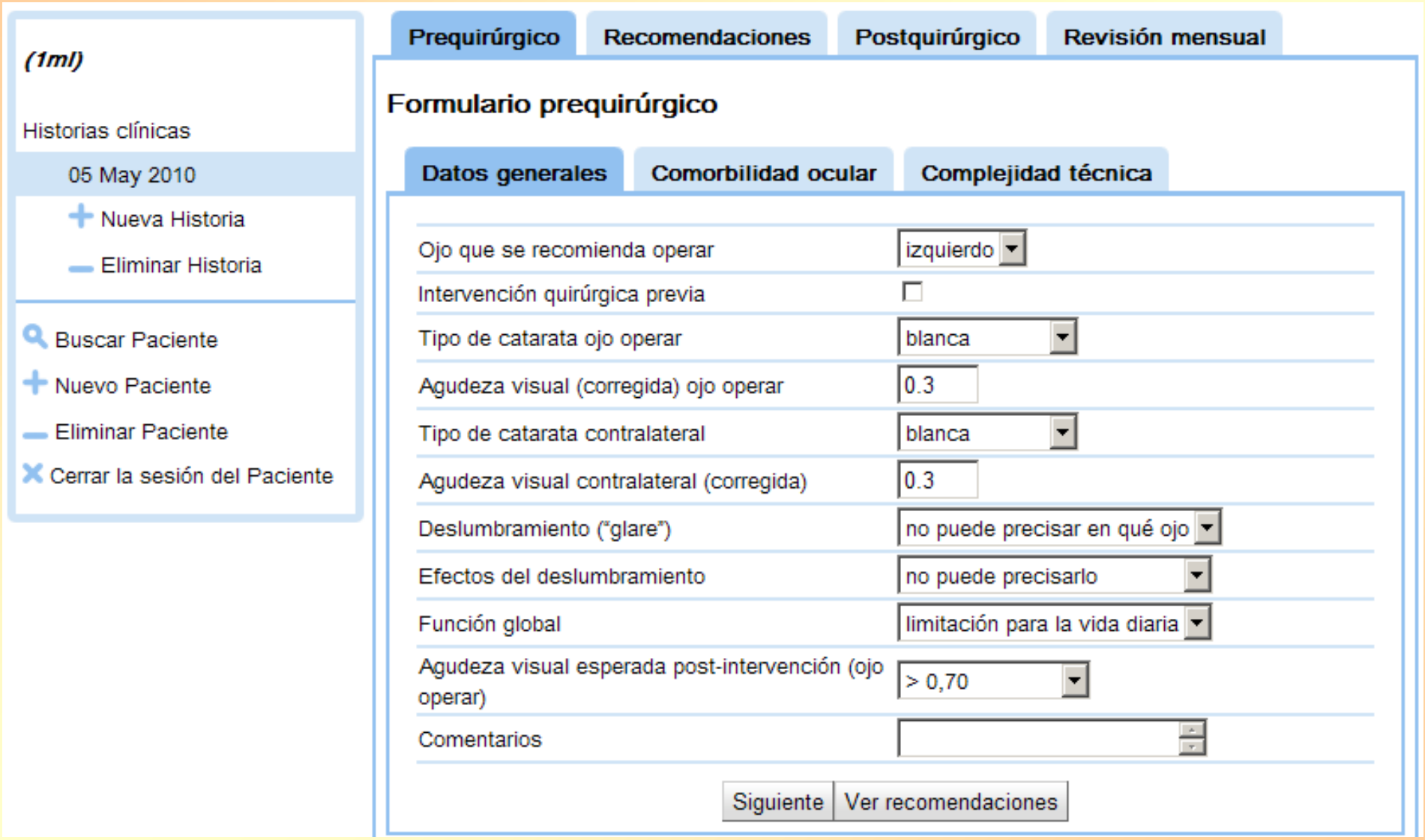
Hepar II (liver diseases)



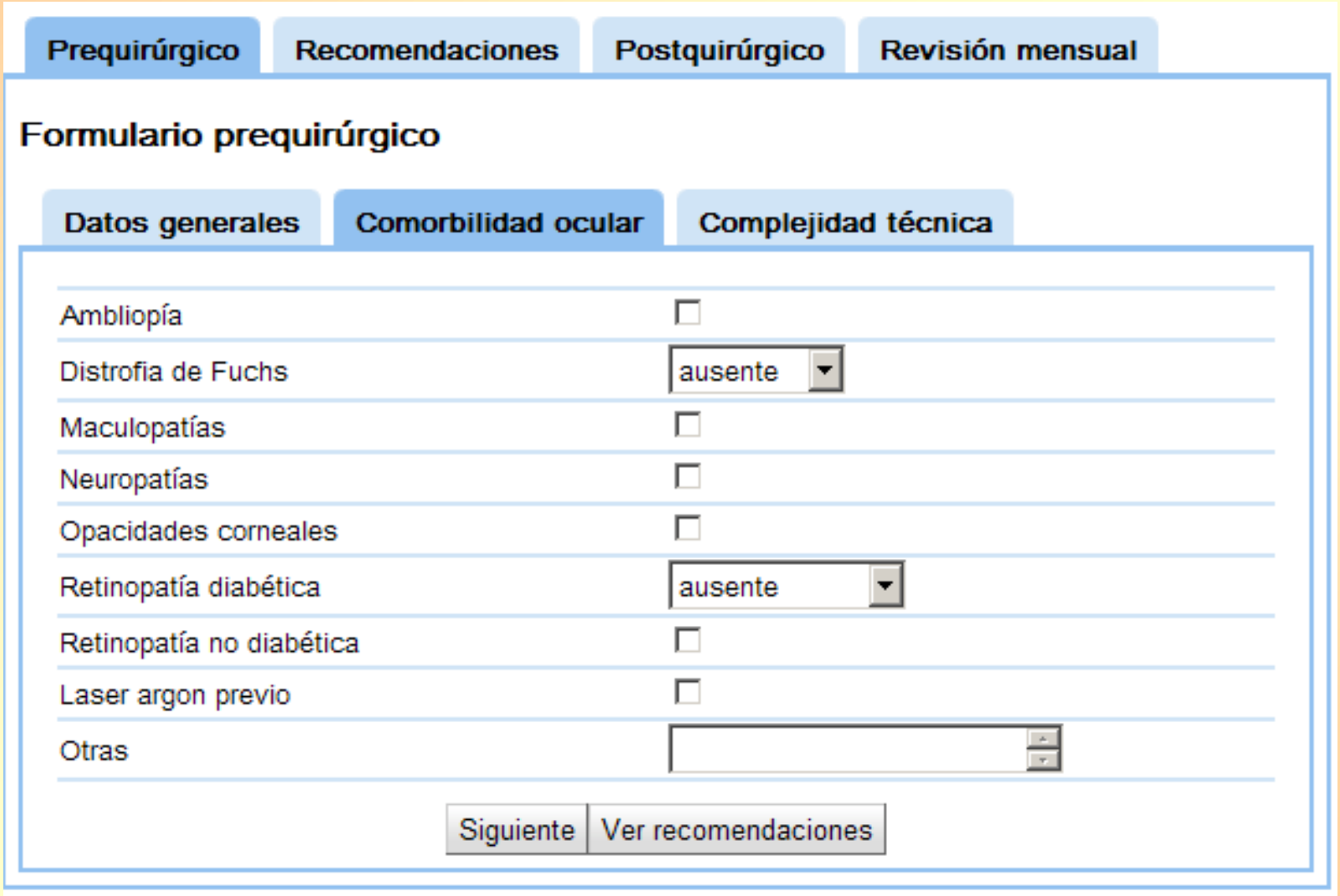
Catarnet (cataract surgery)



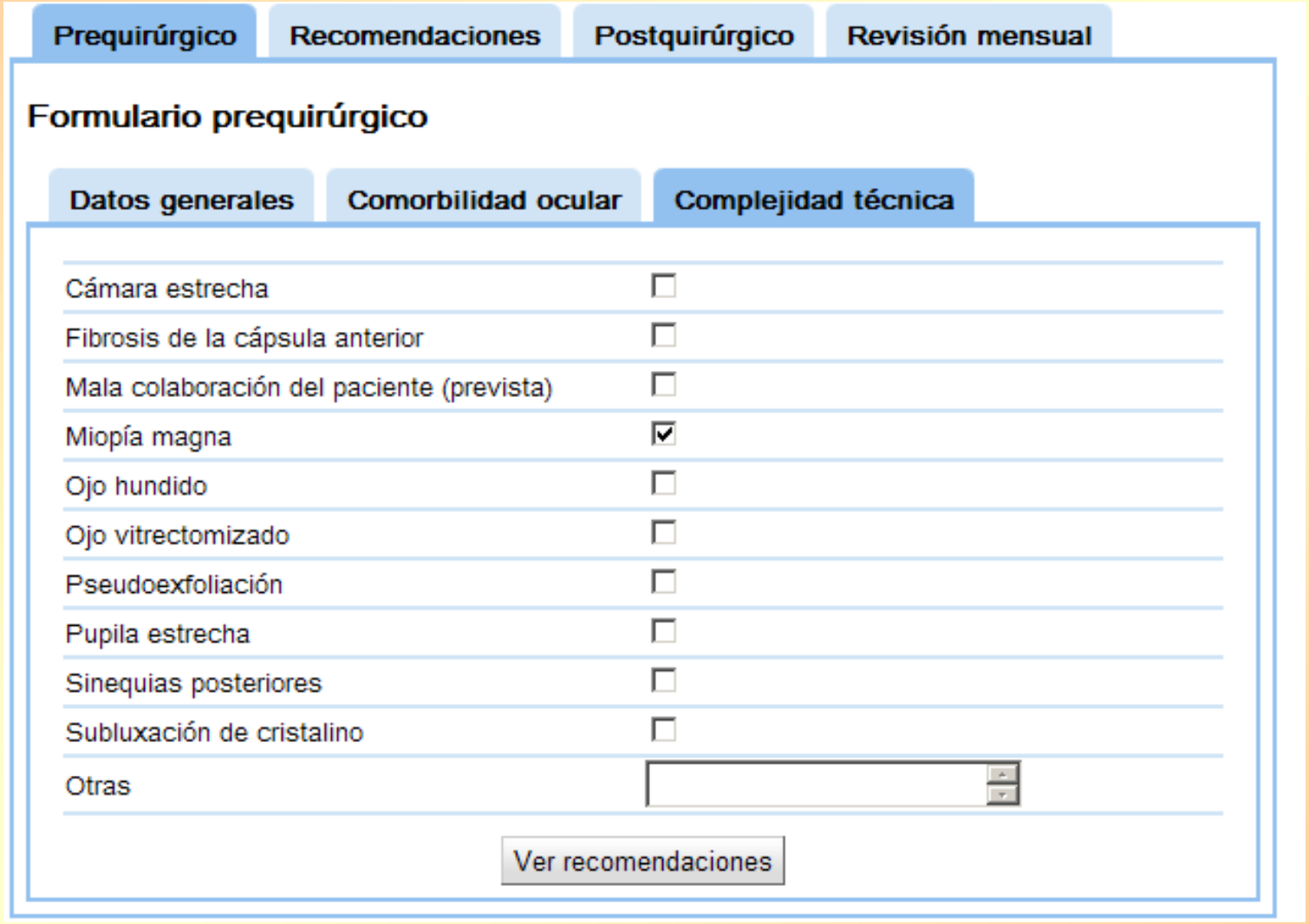
Input: 1. General data



Input: 2. Ocular comorbidity



Input: 3. Surgical complexity



Output: 1. Expert panel's recommendations

Prequirúrgico | **Recomendaciones** | Postquirúrgico | Revisión mensual

Recomendaciones de SAD-Catar

Panel de expertos

Recomendación:	**Facoemulsificación apropiada**
Mediana de las puntuaciones (1 a 9):	8,5
Grado de acuerdo:	Acuerdo

▼ Escenario

Variable	Valor
A.V. contralateral	≥ 0,2 y ≤ 0,4
A.V. previa en el ojo a operar	≥ 0,2 y ≤ 0,4
Patología asociada a la catarata	Catarata simple
Lateralidad de la catarata	Bilateral
Complejidad técnica	Moderada por presencia de: ▪ miopía magna (leve) ▪ catarata blanca (moderada)
Función visual	Dificultades en las actividades de la vida diaria

Explicación

Output: 2. BN recommendation

Red bayesiana CatarNet

Recomendación:	**9 (Totalmente recomendada)**
Mejoría en A.V. (máx. 6):	5,2
Mejoría en deslumbramiento (máx. 5):	1,7

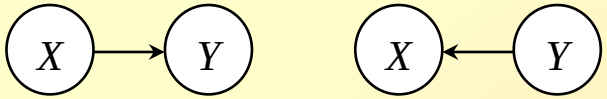
▼ Probabilidades

Función visual post-intervención	**Probabilidad**
Sin problemas	0,057
Dificultades para el ocio	0,830
Dificultades para la vida diaria	0,113
AV post-intervención	**Probabilidad**
≤ 0,15	0,029
> 0,15 y ≤ 0,4	0,088
> 0,4 y ≤ 0,7	0,047
> 0,7	0,836
Deslumbramiento post-intervención	**Probabilidad**
Deslumbramiento	0,544
Complicaciones	**Probabilidad**
Desprendimiento de coroides	0,001
Desprendimiento de retina	0,080
Edema corneal	0,042
Edema macular cistoide	0,020
Endoftalmitis	0,002

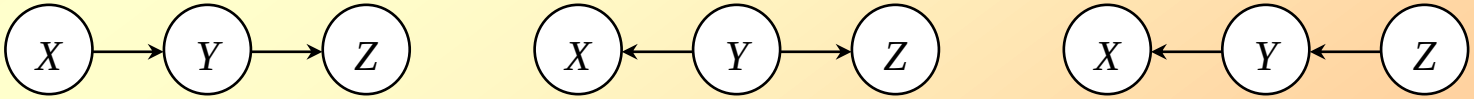
3.3. BNs and causality

Two interpretations of BNs

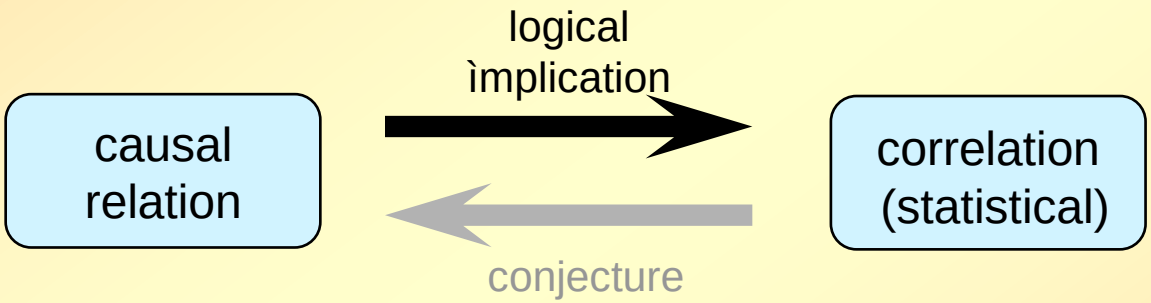
- Semantics of a Bayesian network:
  - As a mathematical model: probabilistic independencies
  - As a model of the real world: they usually represent causality
- Two models are mathematically equivalent when they represent the same set of independencies.
- But two different BNs can never have the same causal meaning.
- Example 1



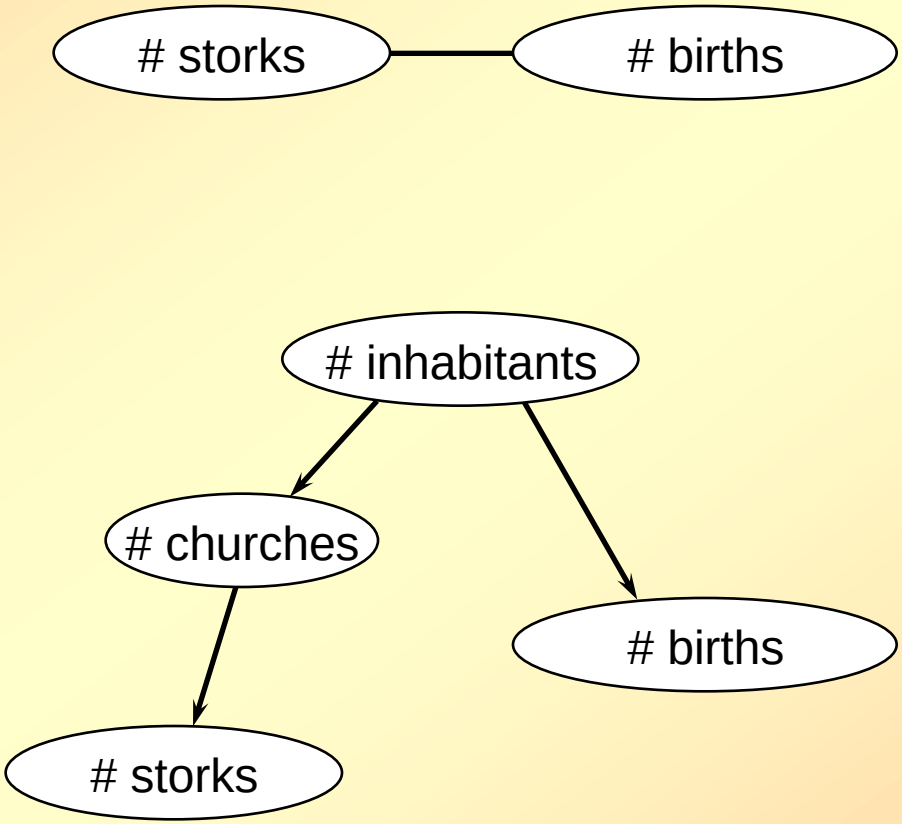
- Example 2



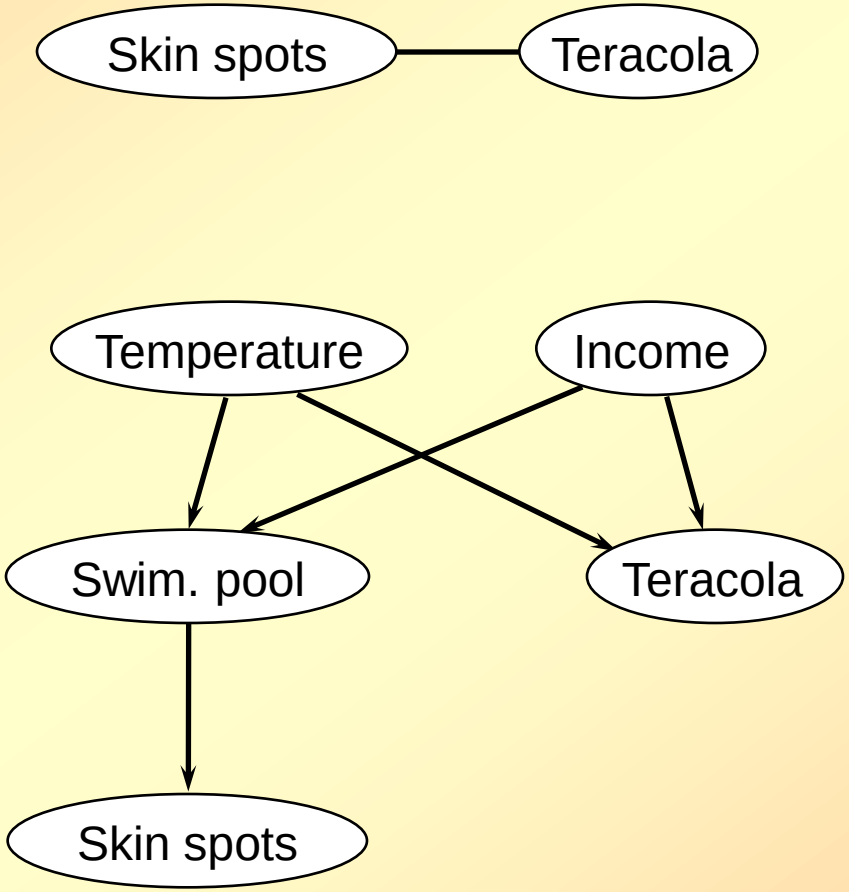
Correlation does not imply causality



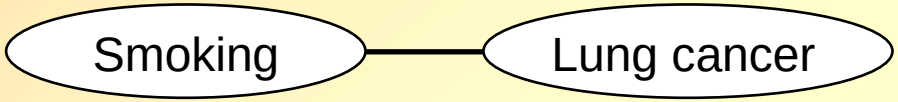
Correlation does not imply causality (example 1)

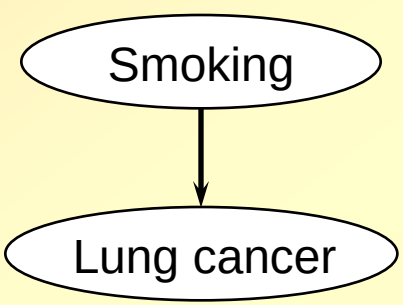


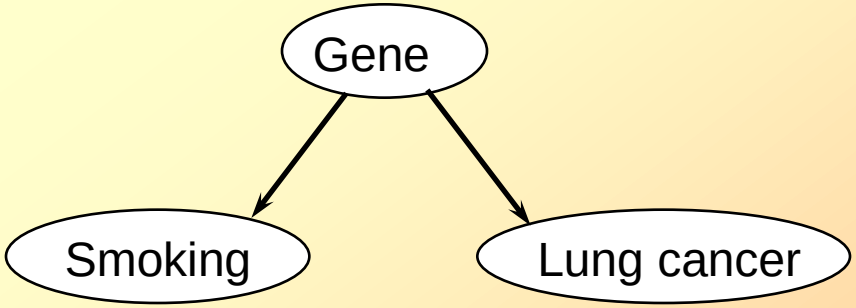
Correlation does not imply causality (example 2)



Correlation does not imply causality (example 3)

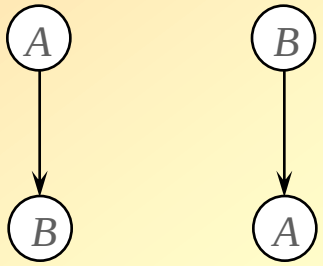




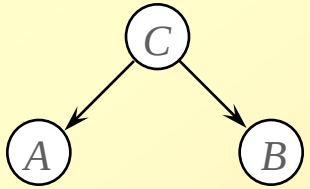


Several types of correlation

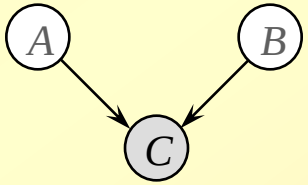
- Direct cause



- Common cause



- Selection bias

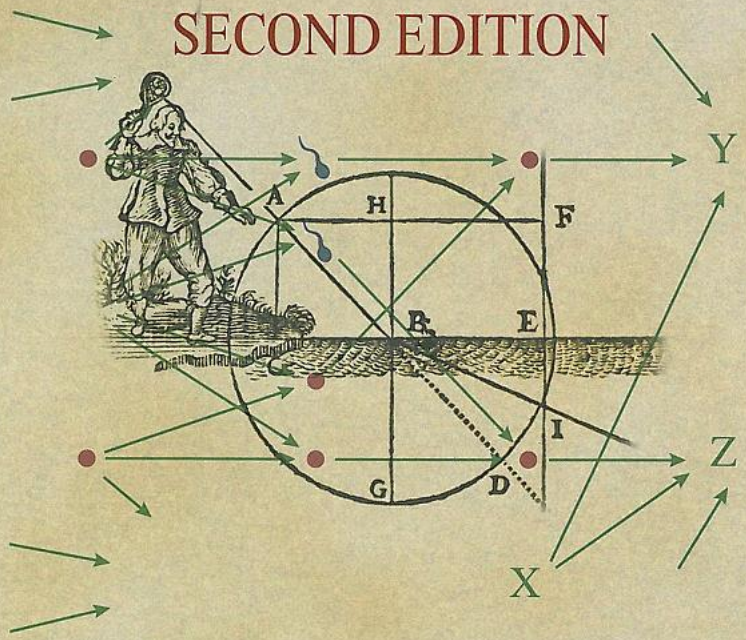


(example: Berkson bias)

Correlation without direct causality

CAUSALITY

SECOND EDITION



MODELS, REASONING, AND INFERENCE

JUDEA PEARL

Miguel Hernan

MIGUEL HERNAN

Search this section

Home
Teaching
Research
Causal Inference Book
Editorial Posts and Commentaries
Scientific Meetings
HIV-CAUSAL Collaboration
Positions Available

Causal Inference Book

My colleague Jamie Robins and I are working on a book that provides a cohesive presentation of concepts of, and methods for, causal inference. Much of this material is currently scattered across journals in several disciplines or confined to technical articles. We expect that the book will be of interest to anyone interested in causal inference, e.g., epidemiologists, statisticians, psychologists, economists, sociologists, political scientists, computer scientists... The book is divided in 3 parts of increasing difficulty: causal inference without models, causal inference with models, and causal inference from complex longitudinal data.

We are making drafts of selected book sections available on this website. The idea is that interested readers can submit suggestions or criticisms before the book is published. To share any comments, please email me or visit @causalinference on Facebook. To cite the book, please use "Hernán MA, Robins JM (2018). Causal Inference. Boca Raton: Chapman & Hall/CRC, forthcoming."

Follow the links below to access different parts of the book:

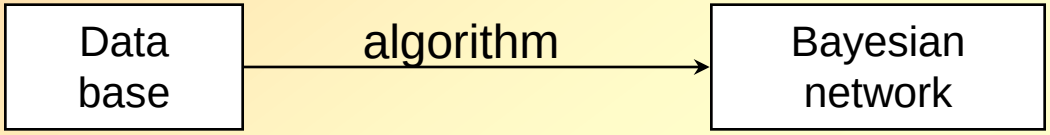
- Part I, Chapters 1-10 (updated 4 October 2017)
- Part II, Chapters 11-17 (updated 5 March 2017)

www.hsph.harvard.edu/miguel-hernan/causal-inference-book

3.4. Building BNs

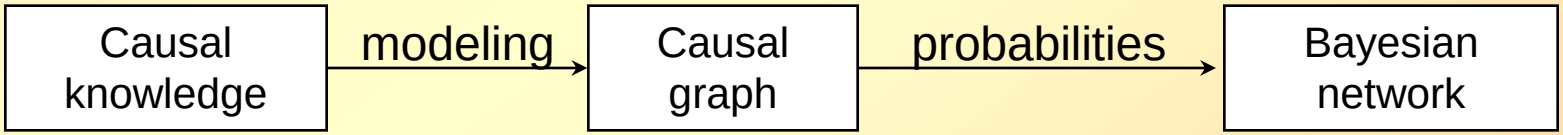
How to build a Bayesian network

- From a database



  - There are many algorithms, several new algorithms every year
  - Similar to statistical methods (logistic regression, neural nets...)

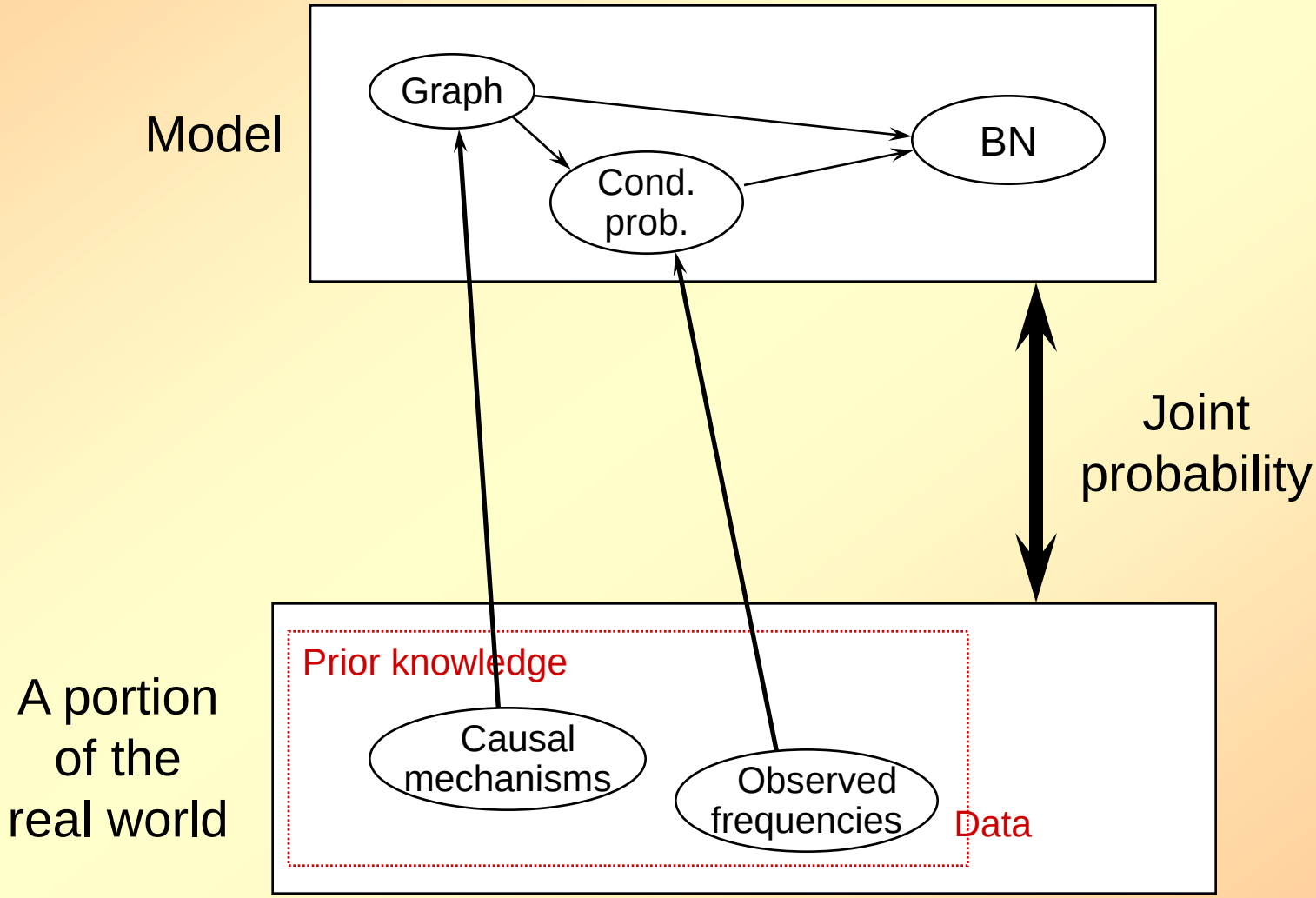
- With a human expert's help



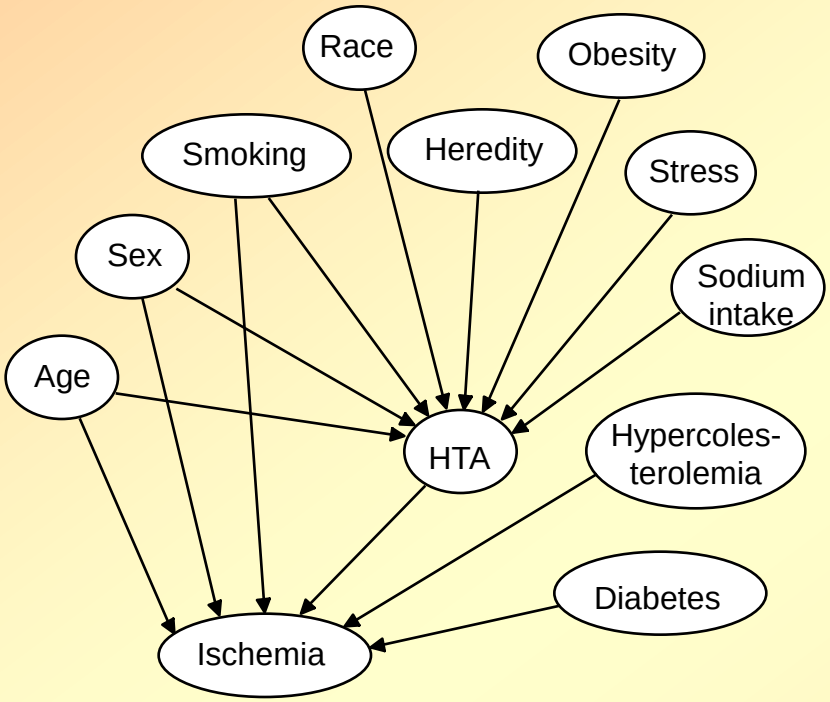
- Hybrid methods:
  - experts → structure; database → probabilities
  - experts → initial model; new cases → refine the probabilities

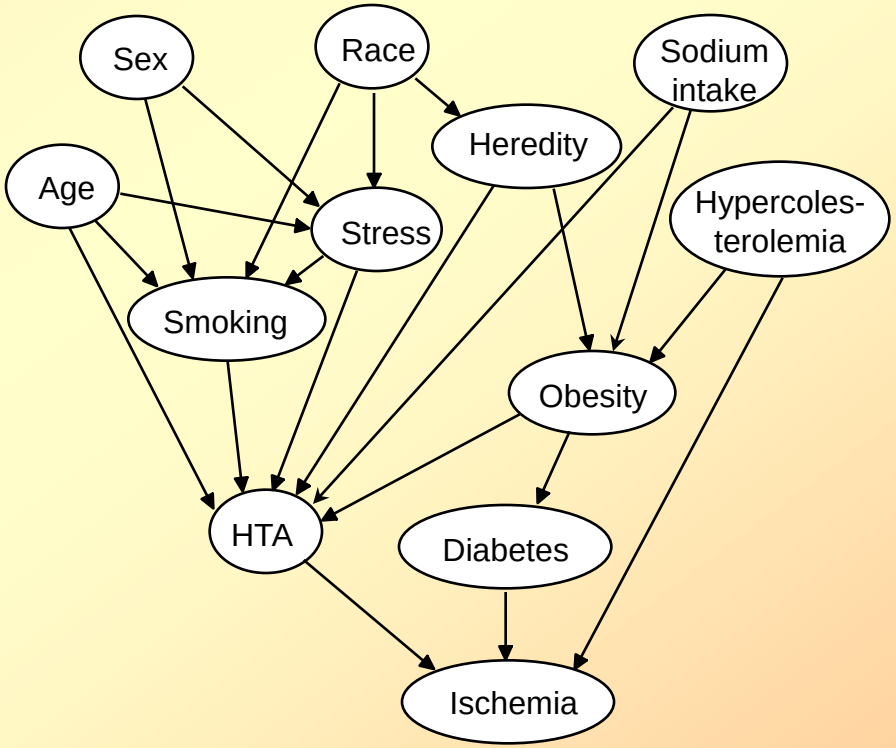
3.4.1. Building BNs with causal knowledge

Building BNs with causal knowledge



PROBLEMS DUE TO LACK OF CAUSAL KNOWLEDGE (1)





Where do the probabilities come from?

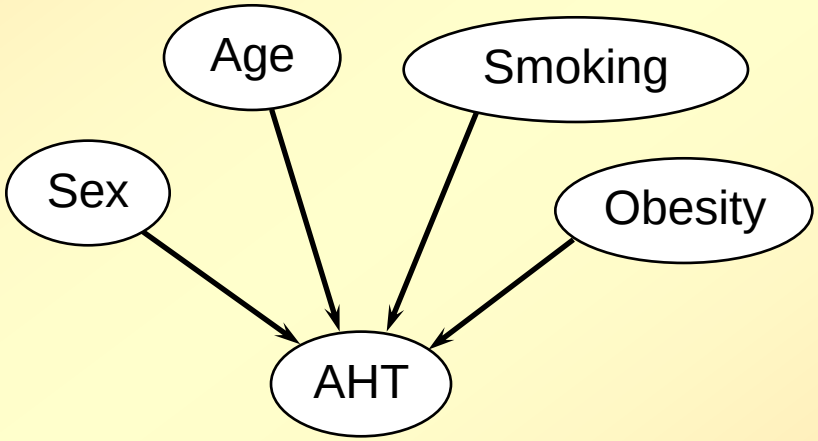
- Epidemiological studies
  - advantage: we obtain directly the parameters we need
  - disadvantage: time and cost; biases
- Medical literature
  - advantage: reliable, inexpensive
  - disadvantage: few qualitative data, few direct probabilities, different criteria, population-dependent, biases
- Databases
  - advantage: fast, inexpensive
  - disadvantage: small databases, selection biases
- Subjective estimates
  - advantage: relatively inexpensive
  - disadvantage: unavailability of experts, psychological biases

3.4.1.1. Canonical models

Canonical models

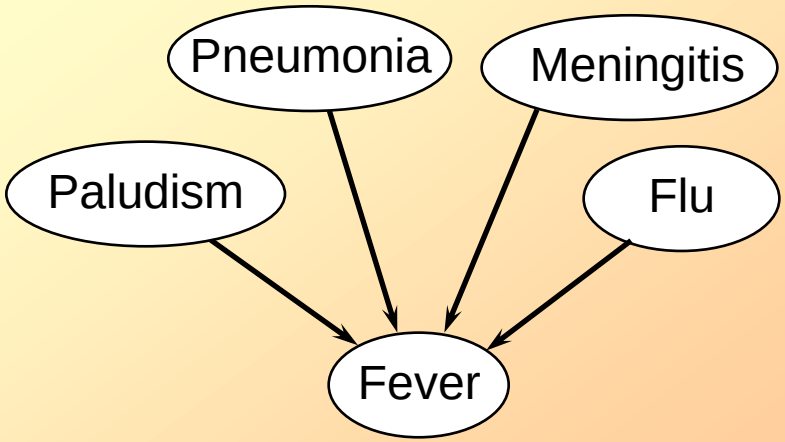
General model

- Probability table:
  $P(y \mid x_1, \ldots, x_n)$
- Factors that influence the prob. of X



Noisy OR

- Efficiency of each link:
  c_i
- Causes that can produce X



The noisy OR (hypotheses)

1. Each cause, by itself, is able to produce the effect, with a certain probability
   - which is less than 100% when there are *inhibitors*.
2. The effect is absent when no cause has produced it
   - i.e., when every cause is either absent or inhibited
3. If a cause has produced the effect, then the effect is present (regardless of the other causes)
4. Independence of causal influences

   ⇒ there is no interaction between the causes (or its inhibitors) when producing the effect

   ⇒ the probability of the effect is the probability that the first cause has produced it, plus the probability that the second cause produces it when the first has not, plus…

Application of the noisy OR when building BNs

- Advantages of the noisy OR
  - Easier to build, because it requires fewer parameters
    - from a database: more cases to estimate each parameter
    - from a human expert: fewer parameters and more intuitive
  - The computation of probability is more efficient (faster)
  - Possibility of explaining the reasoning:
    differencial diagnosis (explaining away)
- Two ways to establish the noisy OR
  - From a statistical study
  - Knowing the causal mechanisms

Canonical Probabilistic Models for Knowledge Engineering

Francisco J. Díez FJDIEZ@DIA.UNED.ES
Dept. Inteligencia Artificial, UNED
Juan del Rosal, 16, 28040 Madrid, Spain

Marek J. Druzdzel MAREK@SIS.PITT.EDU
Decision Systems Laboratory, School of Information Sciences and Intelligent Systems Program
University of Pittsburgh, Pittsburgh, PA 15260, USA

Abstract

The hardest task in knowledge engineering for probabilistic graphical models, such as Bayesian networks and influence diagrams, is obtaining their numerical parameters. Models based on acyclic directed graphs and composed of discrete variables, currently most common in practice, require for every variable a number of parameters that is exponential in the number of its parents in the graph, which makes elicitation from experts or learning from databases a daunting task. In this paper, we review the so called *canonical models*, whose main advantage is that they require much fewer parameters. We propose a general framework for them, based on three categories: deterministic models, ICI models, and simple canonical models. ICI models rely on the concept of *independence of causal influence* and can be subdivided into noisy and leaky. We then analyze the most common families of canonical models (the OR/MAX, the AND/MIN, and the noisy XOR), generalizing them and offering criteria for applying them in practice. We also briefly review temporal canonical models.

Contents

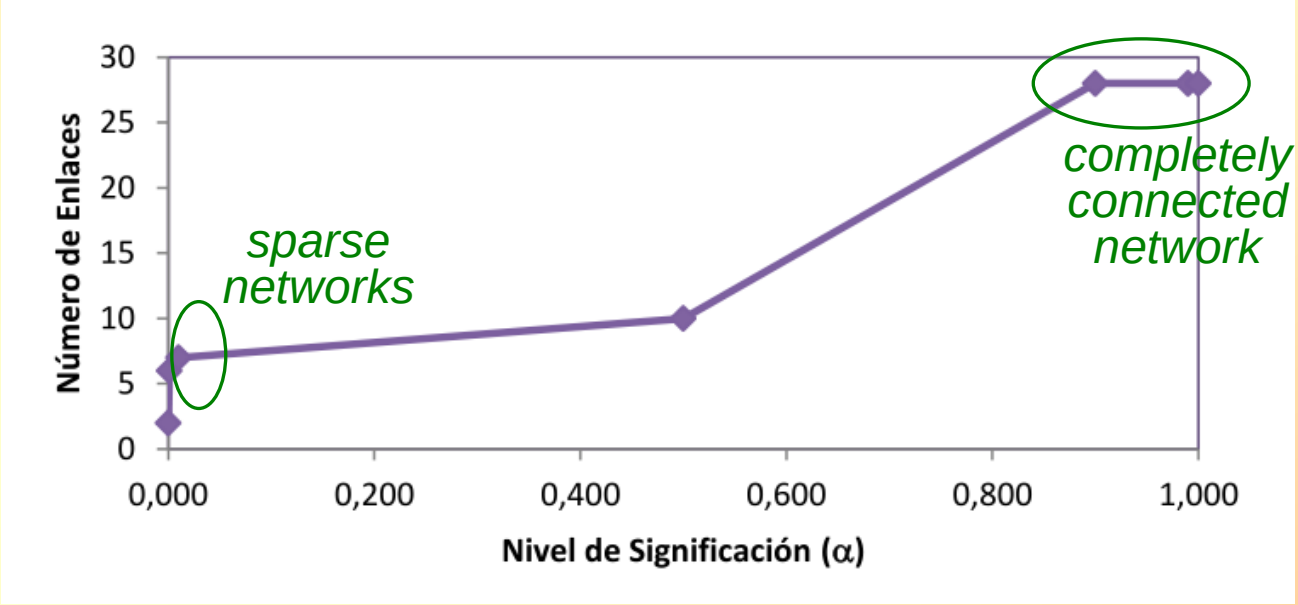
3.4.2. Learning BNs from data

Learning BNs from data

- Two possibilities of learning
  - automatic, interactive
- Two main algorithms:
  - Search-and-score
    - search
      - depart from a network with no links
      - add/remove/invert a link in each iteration
    - score
      - use a metric (there are several metrics available) to quantify how well the model matches the data
  - PC
    - depart from a fully-connected undirected graph
    - when two variables are independent, remove the link
      - more precisely, when the correlation is not statistically significant ($p < \alpha$)
    - when two variables are conditionally indep., remove the link
    - orient the remaining links to obtain a directed graph

The role of significance in the PC algorithm

- We set the value of the significance, α
- For each link, when $p > \alpha$ we assume that the correlation in the database is spurious (i.e., due randomness) and remove the link
- Low value of $\alpha \Rightarrow$ removing many links $\Rightarrow$ sparse network
- High value of $\alpha \Rightarrow$ keeping many links $\Rightarrow$ dense network



Advantages of interactive learning

- The system proposes, the user decides
  - Very useful for tuition
  - Useful for combining data with expert knowledge
  - Useful for debugging new algorithms (workbench)
- See www.openmarkov.org/docs/tutorial.

A comparison of both methods for building BNs

- Automatic learning from databases
  - Advantage: faster (graph + probabilities)
  - Limitation: medical databases are usually incomplete
    - Missing values $\rightarrow$ problem of imputation (rarely missing at random)
    - Missing variables $\rightarrow$ spurious correlations
  - Blackbox algorithm that returns non-causal models

    $\Rightarrow$ Human experts are reluctant to accept their advice
- With expert knowledge (“manual” method)
  - Only method possible when there is not a good-enough database
  - Difficulty in practice: getting the collaboration of experts
  - Building the structure of the causal is sometimes difficult
  - Obtaining the probabilities is even more difficult.

Summary: BNs vs. the naïve Bayes

- BNs can diagnose several diseases simultaneously.
- BNs do *not* assume conditional independence of findings.
- BNs are usually causal models
  - closer to doctors' reasoning: explanation of reasoning
  - probabilities are in general easier to obtain
- Three types of reasoning: abductive, deductive, inter-causal.
- They can combine data (from databases), epidemiological studies (scientific literature) and expert knowledge (doctors).

In spite of these advantages,
BNs are almost unknown in medicine.
No book for medical doctors mentions them!

4. Unicriterion decision analysis

4.1. Introductory examples

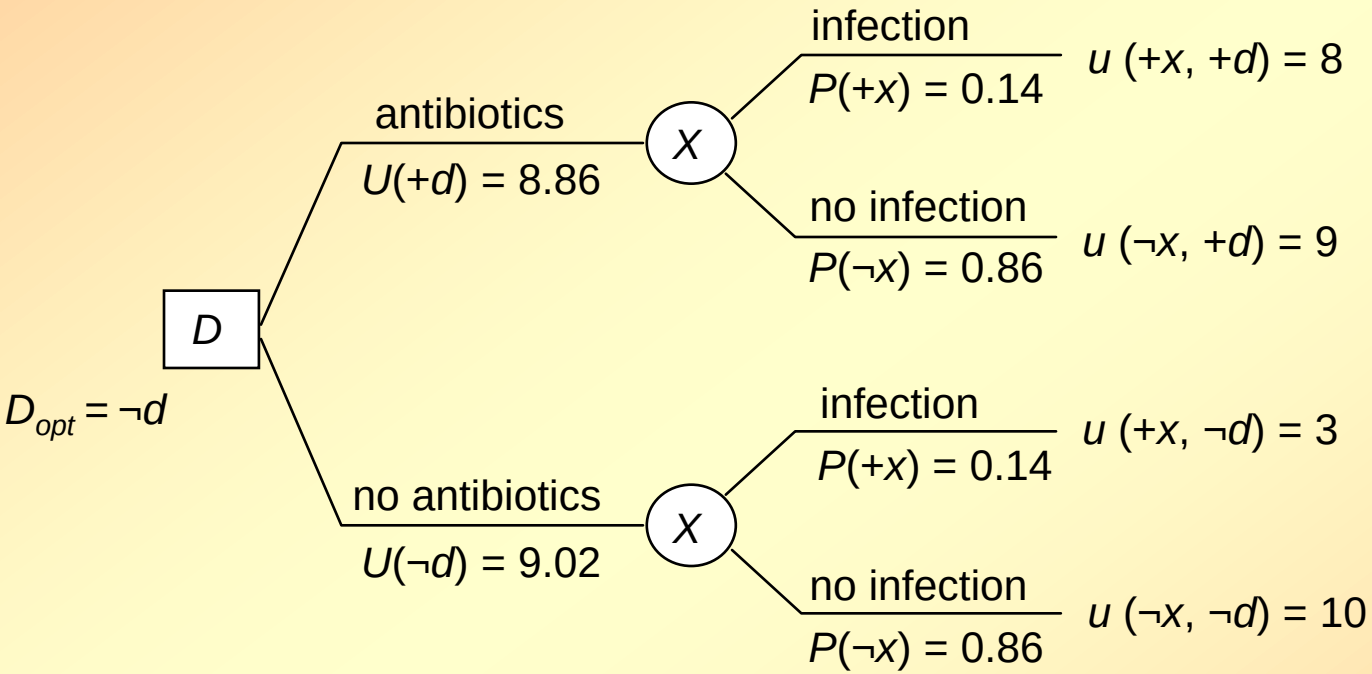
Medical example (1)

- Three variables
  - Chance variable: $X \rightarrow$ bacterial infection; $P(+x) = 0.14$
  - Decision: $D \rightarrow$ give antibiotics
  - Utility (value): $U \rightarrow$ effectiveness

$u\ (x, d)$	$+x$	$\neg x$
$+d$	8	9
$\neg d$	3	10

- When making the decision we do not know whether the patient is infected with the bacteria.
- Question: Should we give antibiotics?

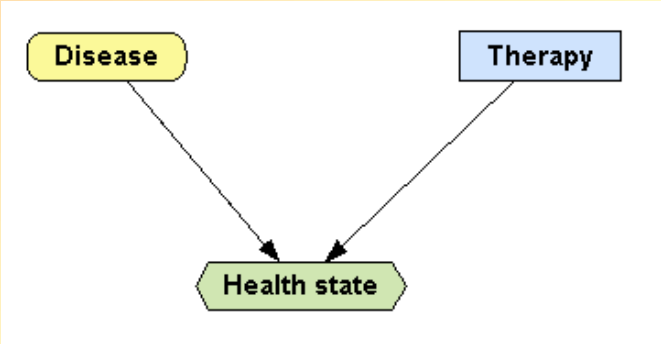
Decision tree (1)



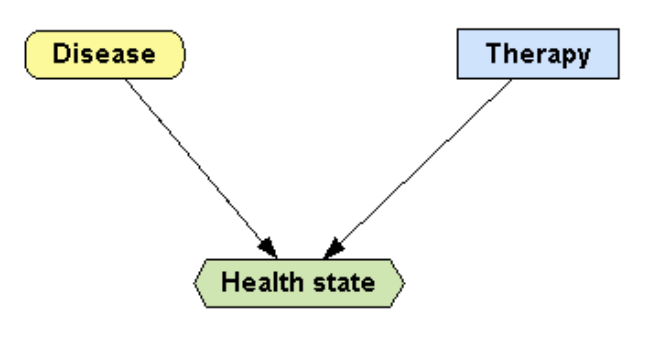
Optimal decision: $D_{opt} = \neg d \Rightarrow$ do not give antibiotics

Prognosis: $U = \max(U(+d), U(\neg d)) = \max(8.86, 9.02) = 9.02$

Influence diagram



DAN
(decision analysis network)



- Both models are identical.
- They generate the same decision tree.

Utility as a function of prevalence

$$U(d) = \sum_x u(x,d) \cdot P(x)$$

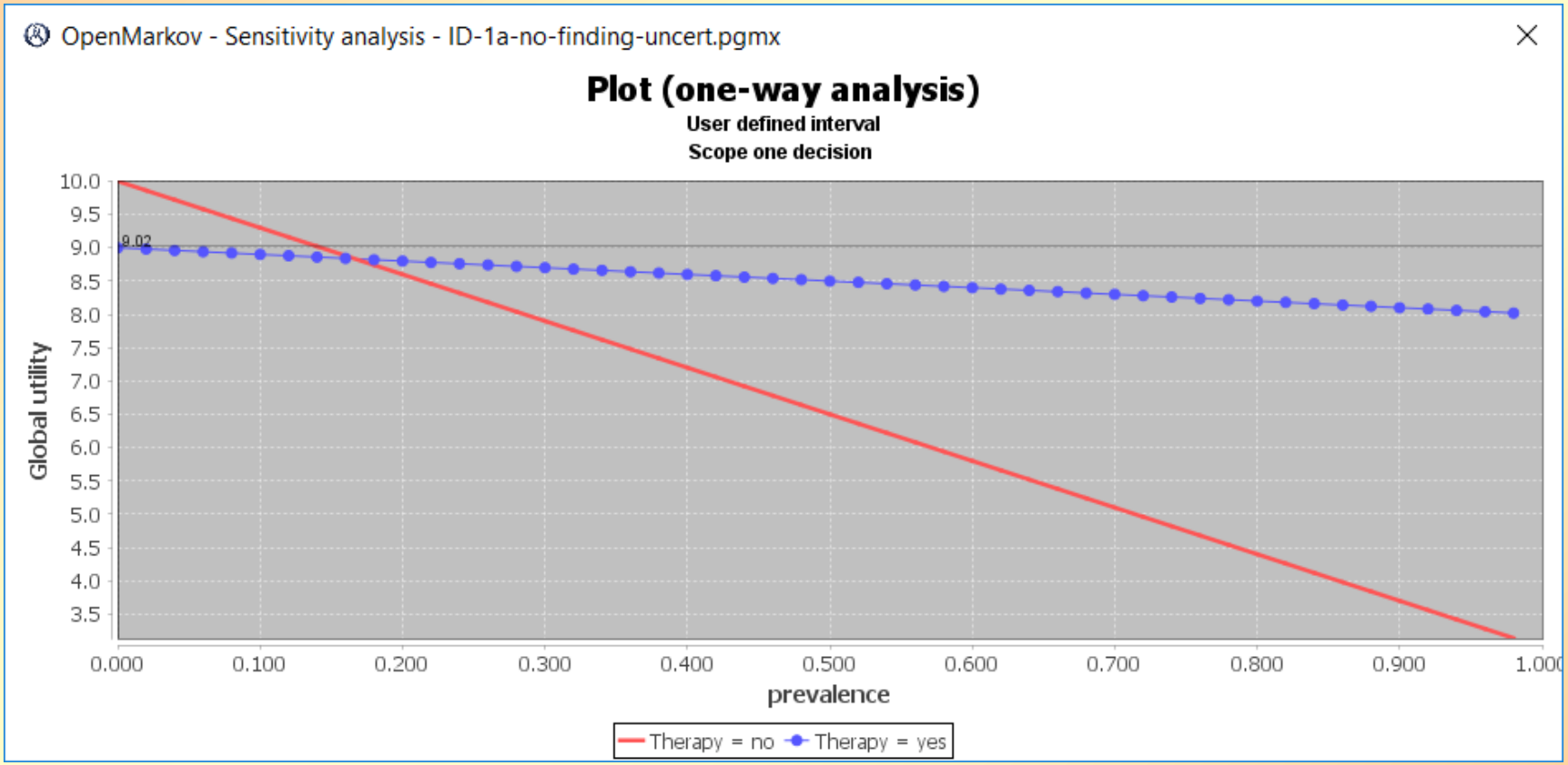
$$D_{opt} = \arg\max(U(+d), U(\neg d))$$

$$U = \max(U(+d), U(\neg d))$$

$P(+x)$	$U(+d)$	$U(\neg d)$	D_{opt}	U
0’00	9’00	10’00	$\neg d$	10’00
0’05	8’95	9’65	$\neg d$	9’79
0’14	8’86	9’02	$\neg d$	9’02
0’17	8’83	8’81	$+d$	8’83
0’40	8’60	7’20	$+d$	8’60
0’75	8’25	4’75	$+d$	8’25
1’00	8’00	3’00	$+d$	8’00

decision threshold

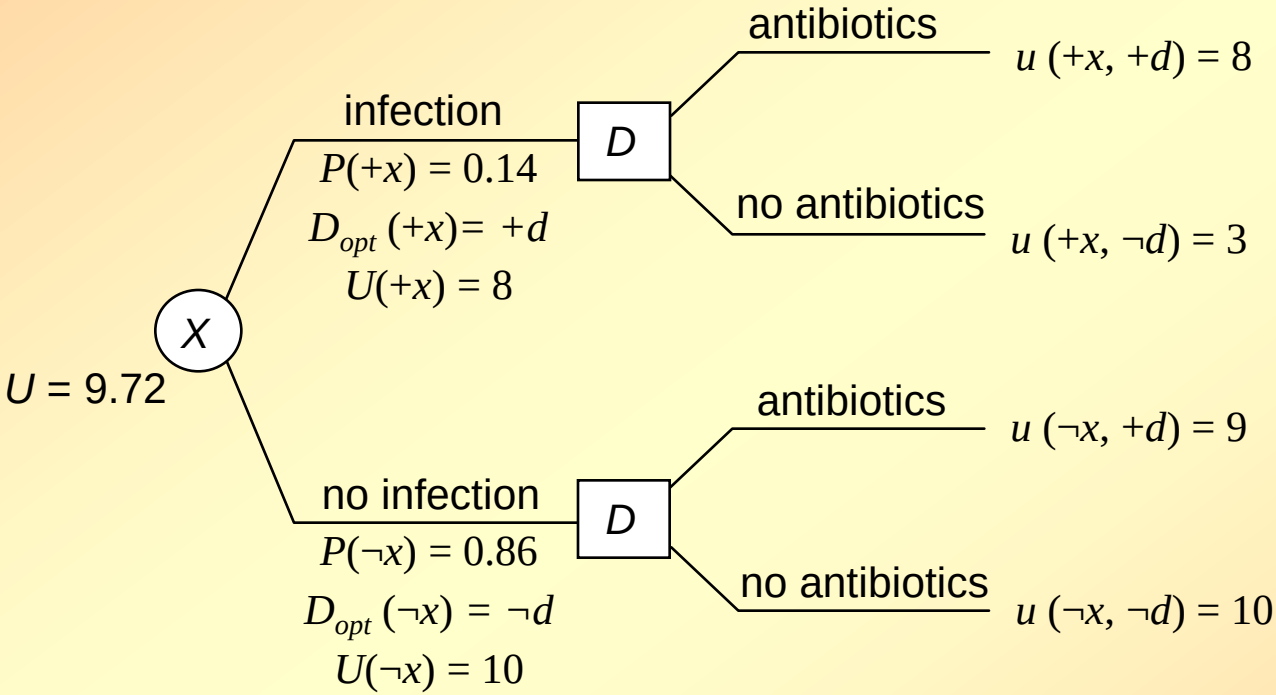
Utility as a function of prevalence



Medical example (2)

- In the previous scenario, what should we do if we knew with certainty whether the patient has the disease?
  - Question 1: What to do when infection is present?
  - Question 2: What to do when infection is absent?
- What is the average utility in this sub-population?

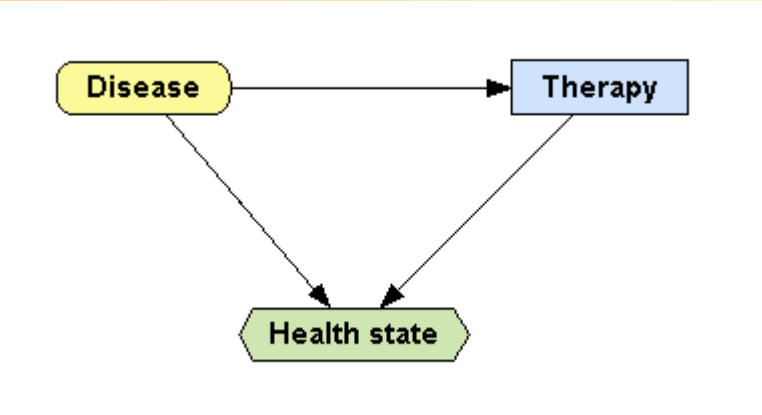
Decision tree (2)



Optimal decision: infection (+*x*) $\Rightarrow$ give antibiotics (+*d*)
no infection (¬*x*) $\Rightarrow$ do not give antibiotics (¬*d*)

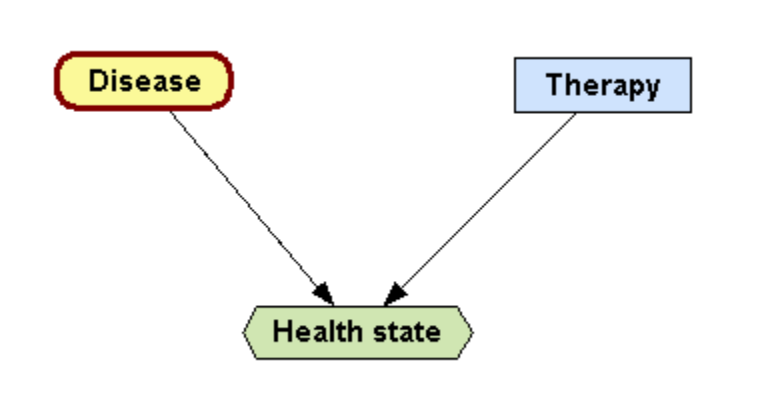
Expected utility: $U = 8 \times 0.14 + 10 \times 0.86 = 9.72$

Influence diagram



We have added an **information link**.

DAN



We have marked *Disease* as **always-observed**.

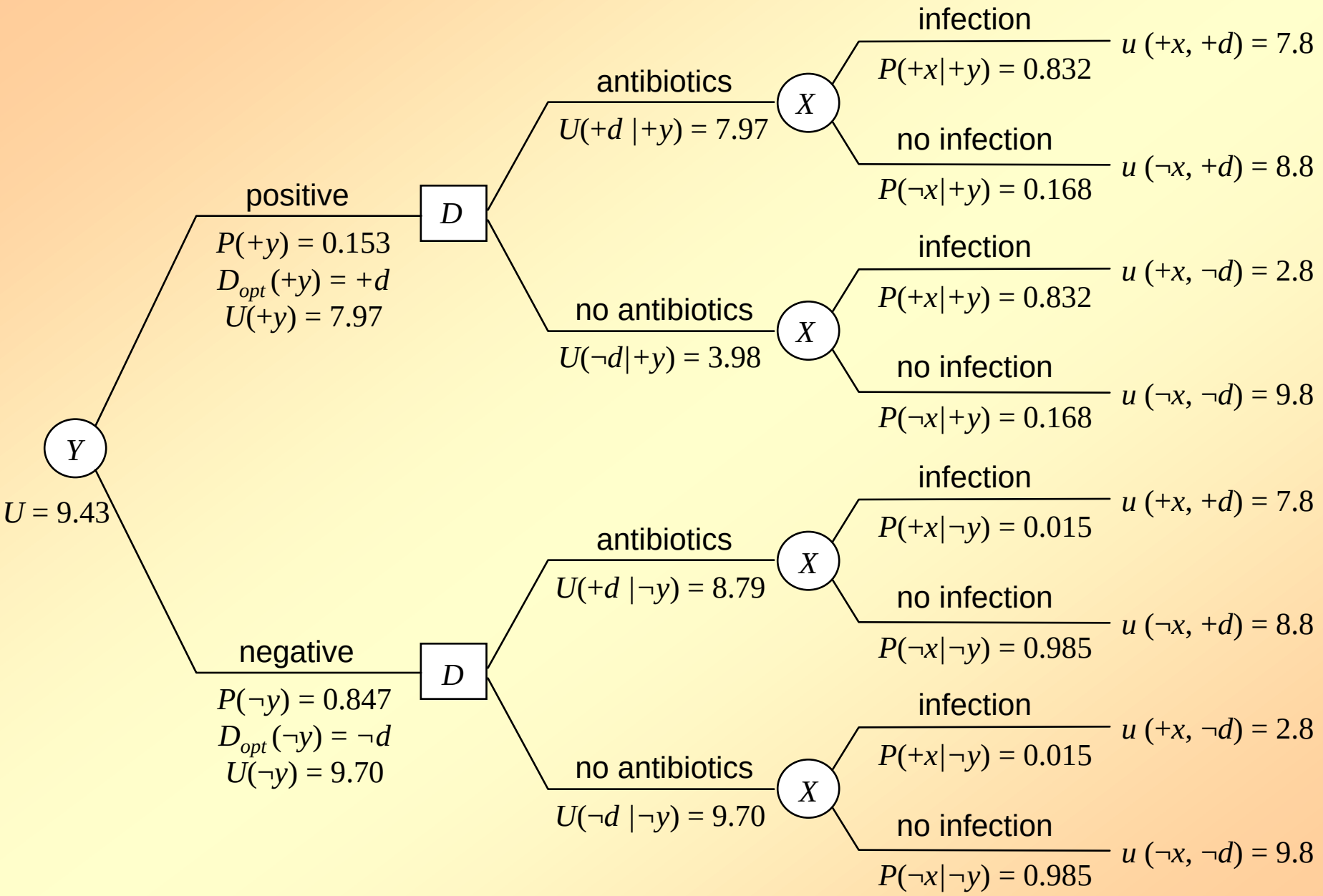
- Two different ways of saying that the value of *Disease* is known when making the decision *Therapy*.
- Both models are equivalent: they generate the same decision tree.

Medical example (3): The value of information

- Test Y for detecting X
  - sensitivity: $P(+y|+x) = 0.91$
  - specificity: $P(\neg y|\neg x) = 0.97$
  - cost: $u_{\text{test}}(x, d) = u_{\text{no test}}(x, d) - 0.2$

$u\ (x, d)$	$+x$	$\neg x$
$+d$	7'8	8'8
$\neg d$	2'8	9'8

- When making the decision we do **know** the result of the test.
- Question: Should we give antibiotics?


Y
U = 9.43
positive
P(+y) = 0.153
D_opt(+y) = +d
U(+y) = 7.97
negative
P(¬y) = 0.847
D_opt(¬y) = ¬d
U(¬y) = 9.70
D
D
antibiotics
U(+d |+y) = 7.97
no antibiotics
U(¬d|+y) = 3.98
antibiotics
U(+d |¬y) = 8.79
no antibiotics
U(¬d |¬y) = 9.70
X
X
X
X
infection
P(+x|+y) = 0.832
u (+x, +d) = 7.8
no infection
P(¬x|+y) = 0.168
u (¬x, +d) = 8.8
infection
P(+x|+y) = 0.832
u (+x, ¬d) = 2.8
no infection
P(¬x|+y) = 0.168
u (¬x, ¬d) = 9.8
infection
P(+x|¬y) = 0.015
u (+x, +d) = 7.8
no infection
P(¬x|¬y) = 0.985
u (¬x, +d) = 8.8
infection
P(+x|¬y) = 0.015
u (+x, ¬d) = 2.8
no infection
P(¬x|¬y) = 0.985
u (¬x, ¬d) = 9.8

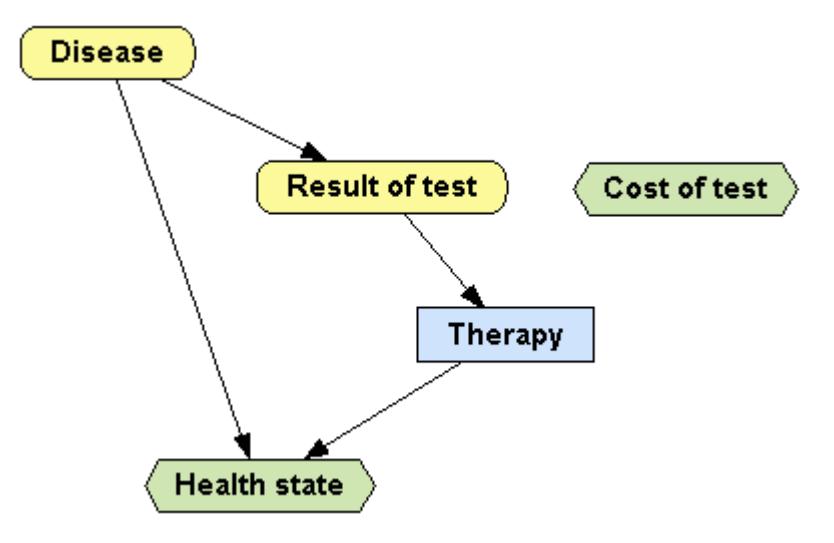
Policy and prognosis

- Policy:
  - When Y is positive: give antibiotics
  - When Y is negative: do not give antibiotics

- Prognosis
  - When Y is positive: $U(+y) = 7.97$
  - When Y is negative: $U(\neg y) = 9.70$
  - Global prognosis (average utility)

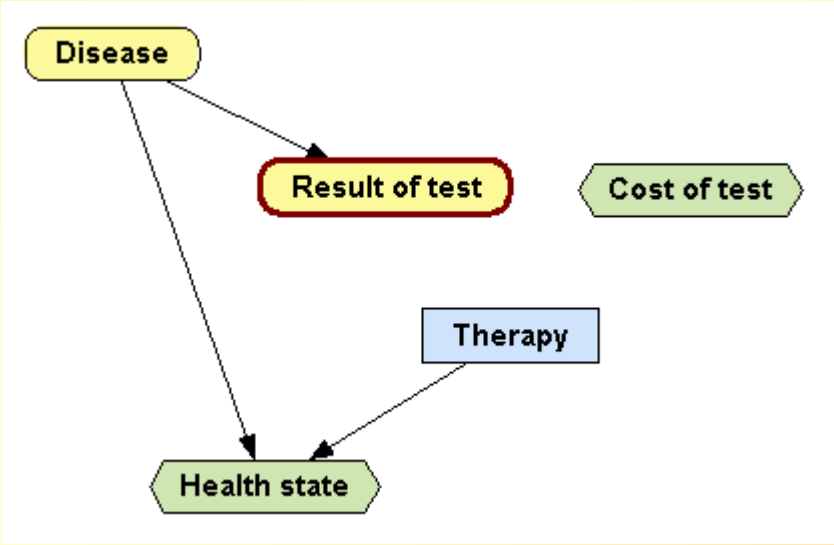
$$\begin{aligned} U_{\text{with test}} &= U(+y) \times P(+y) + U(\neg y) \times P(\neg y) \\ &= 7.97 \times 0.153 + 9.69 \times 0.847 \\ &= 9.43 \end{aligned}$$

Influence diagram



An **information link**
from *Result of test* to *Therapy*

DAN

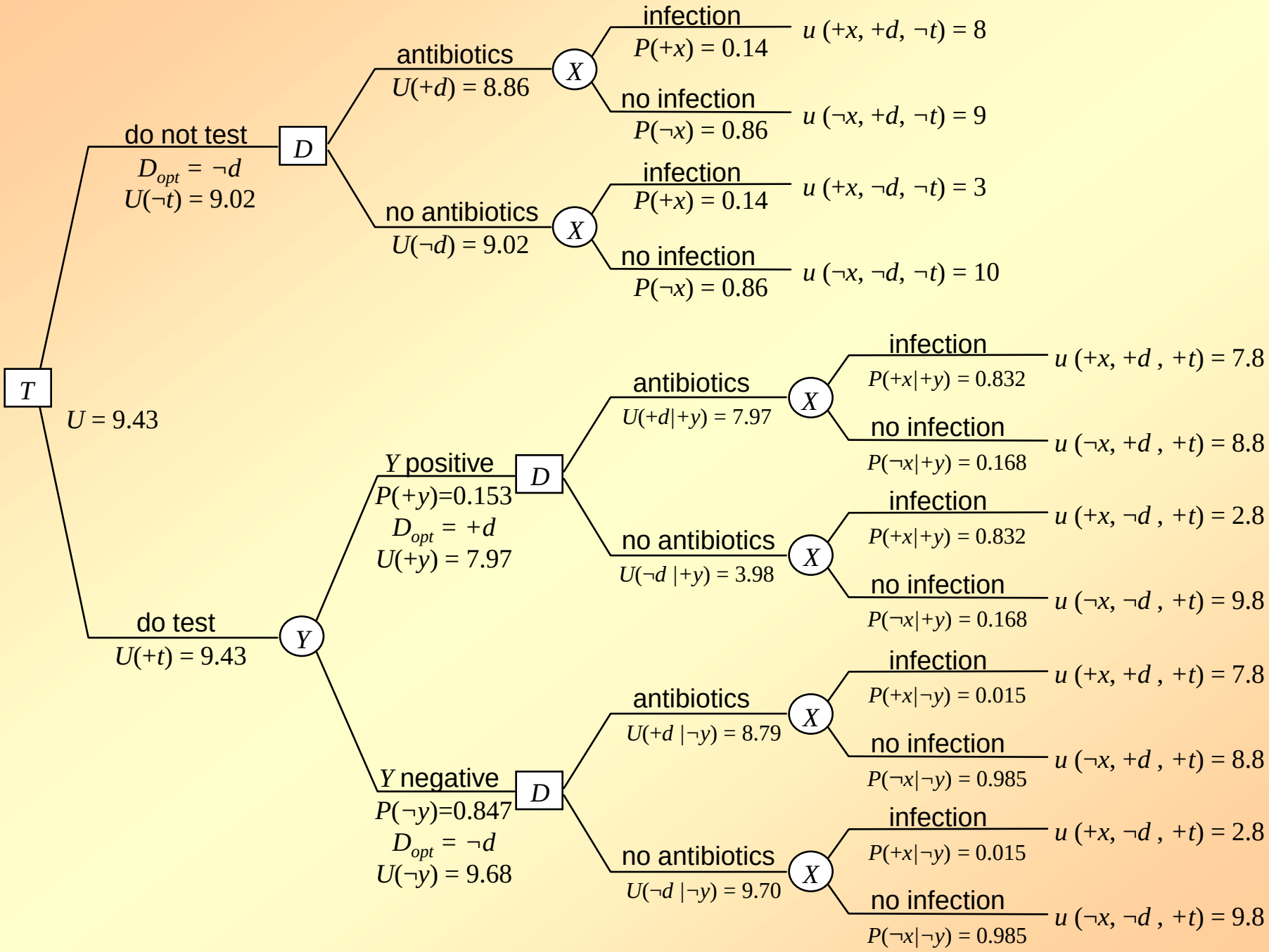


Result of test is marked as
always-observed.

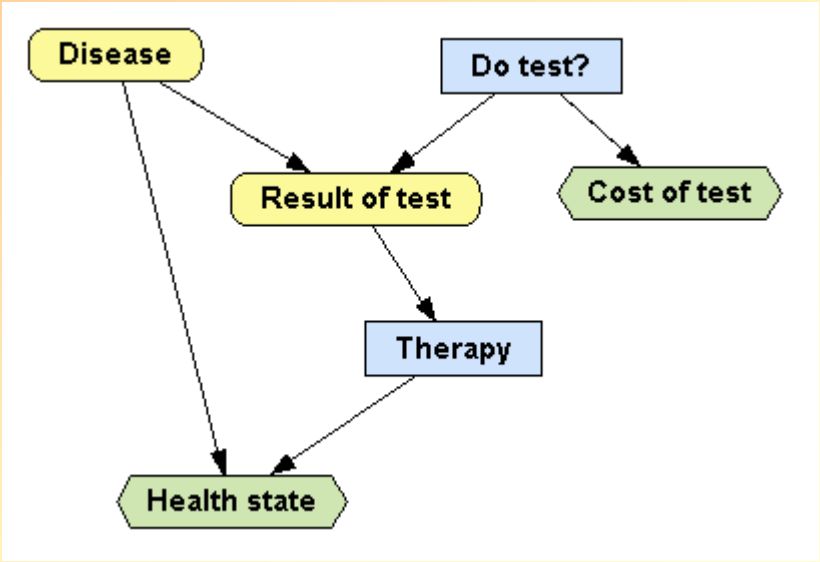
- Different ways of indicating the flow of information.
- Both models generate the same decision tree.

Medical example (4): deciding about a test

- Test Y
  - Advantage: gives information
  - Disadvantage: has a cost
- Is it worth doing the test?
- Three possible policies:
  1. Give the therapy to all patients, preventively
  2. Never apply the therapy
  3. Do test Y; apply the therapy only when it is positive

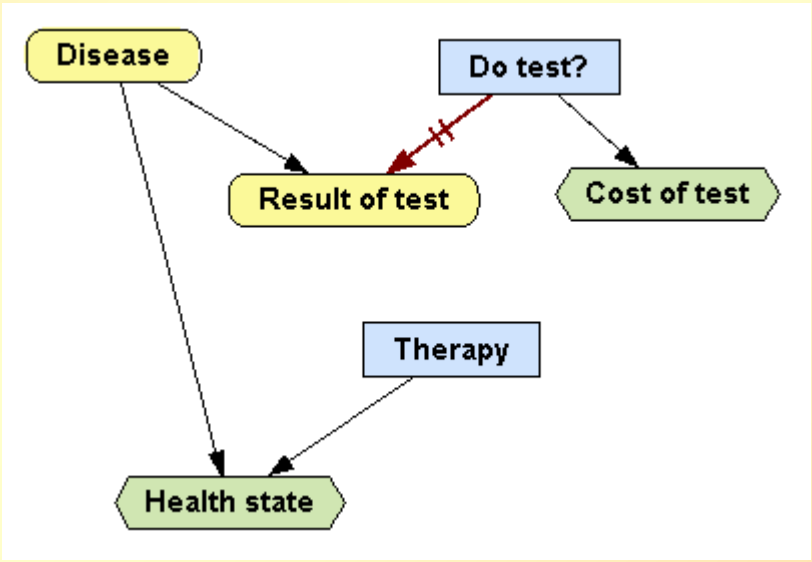

T
U = 9.43
do not test
D_opt = ¬d
U(¬t) = 9.02
D
antibiotics
U(+d) = 8.86
X
infection
P(+x) = 0.14
u (+x, +d, ¬t) = 8
no infection
P(¬x) = 0.86
u (¬x, +d, ¬t) = 9
no antibiotics
U(¬d) = 9.02
X
infection
P(+x) = 0.14
u (+x, ¬d, ¬t) = 3
no infection
P(¬x) = 0.86
u (¬x, ¬d, ¬t) = 10
do test
U(+t) = 9.43
Y
Y positive
P(+y)=0.153
D_opt = +d
U(+y) = 7.97
D
antibiotics
U(+d|+y) = 7.97
X
infection
P(+x|+y) = 0.832
u (+x, +d , +t) = 7.8
no infection
P(¬x|+y) = 0.168
u (¬x, +d , +t) = 8.8
no antibiotics
U(¬d |+y) = 3.98
X
infection
P(+x|+y) = 0.832
u (+x, ¬d , +t) = 2.8
no infection
P(¬x|+y) = 0.168
u (¬x, ¬d , +t) = 9.8
Y negative
P(¬y)=0.847
D_opt = ¬d
U(¬y) = 9.68
D
antibiotics
U(+d |¬y) = 8.79
X
infection
P(+x|¬y) = 0.015
u (+x, +d , +t) = 7.8
no infection
P(¬x|¬y) = 0.985
u (¬x, +d , +t) = 8.8
no antibiotics
U(¬d |¬y) = 9.70
X
infection
P(+x|¬y) = 0.015
u (+x, ¬d , +t) = 2.8
no infection
P(¬x|¬y) = 0.985
u (¬x, ¬d , +t) = 9.8

Influence diagram



An **information link**.
Total ordering of the decisions

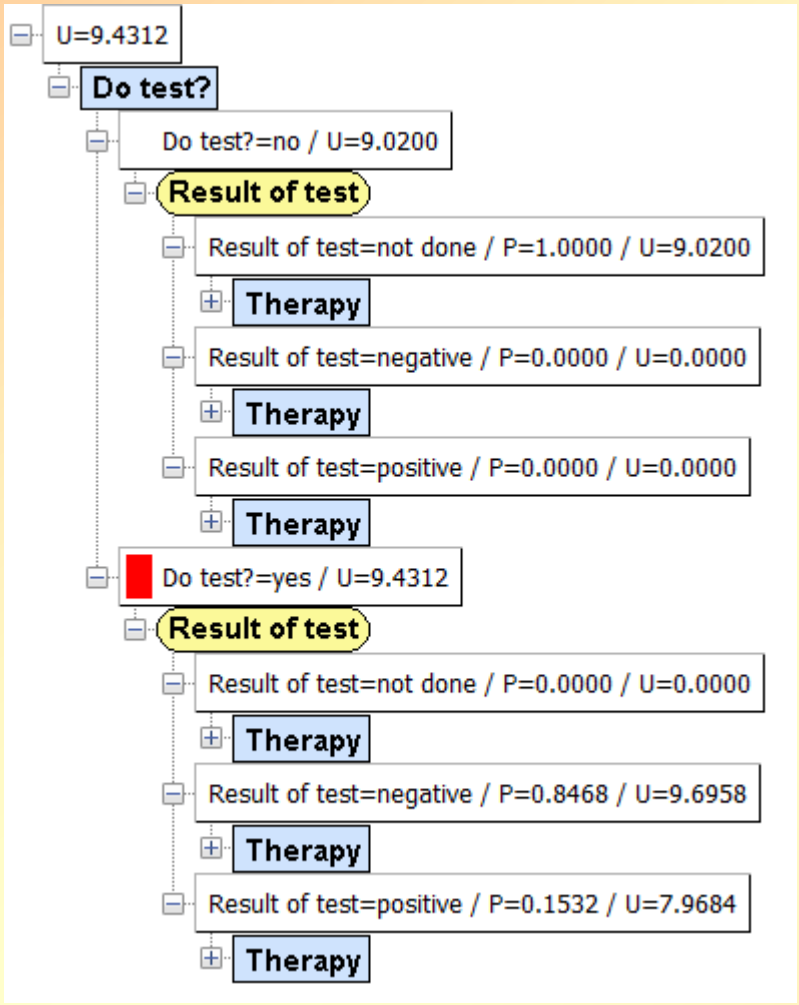
DAN



Restrictions. **Revelation link**.
The decisions are not ordered.

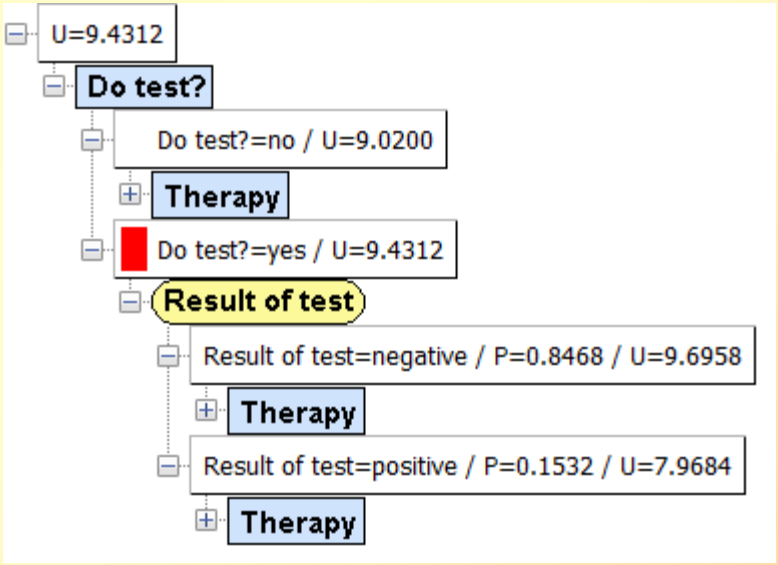
- Different ways of indicating the flow of information.
- The decision trees are different but <u>equivalent</u>:
  the same probabilities, utilities, and policies.

Decision tree generated by the ID



symmetric

Decision tree generated by the DAN



asymmetric

Conditional prob. for *Result of test*

in the ID

Node Potential: Result of test

Relation Type: Table Reor

Do test?	no	no	yes	yes
Disease	absent	present	absent	present
positive	0	0	0.03	0.91
negative	0	0	0.97	0.09
not done	1	1	0	0

- dummy value:
  test not done

in the DAN

Node Potential: Result of test

Relation Type: Table Reor

Do test?	no	no	yes	yes
Disease	absent	present	absent	present
positive	0	0	0.03	0.91
negative	0	0	0.97	0.09

- restrictions
- no dummy value

Hands-on exercise 3

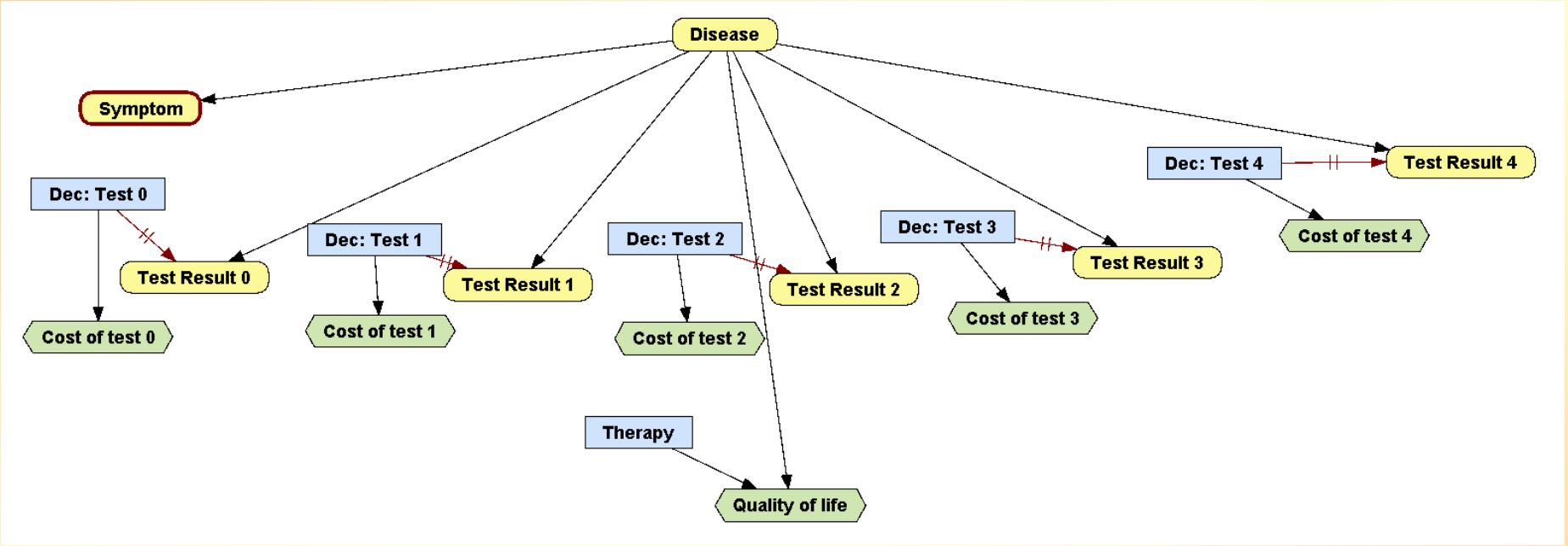
Exercise: Optimal stratety for two tests

Test	sensitivity	specificity	discomfort
A	0.60	0.92	0.0003 QALY
B	0.80	0.91	0.0001 QALY

Disease →	absent	present
therapy	38 QALY	30 QALY
no therapy	40 QALY	20 QALY

- QALY is a unit of effectiveness
- Question: What is the most effective strategy?

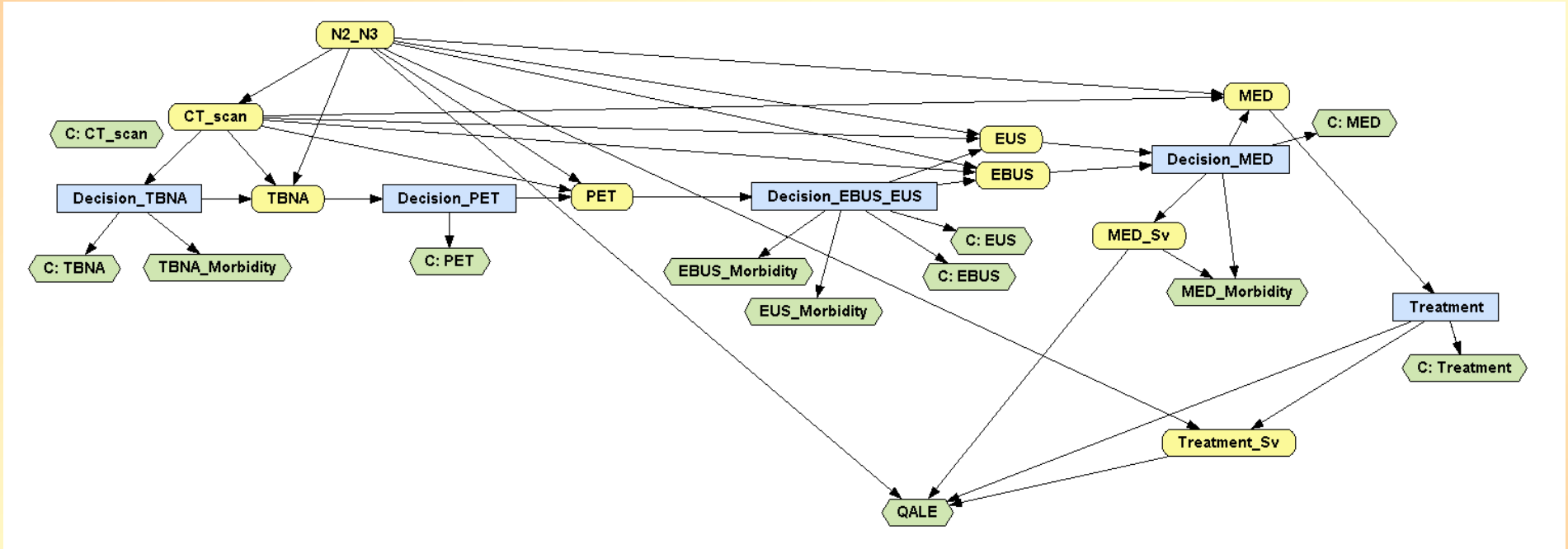
The n-test problem



- Computationally hard: $n!$ possible orderings of the tests.
- We have developed an any-space algorithm for this problem
- and a fast algorithm (9 minutes for the 7-test problem).
- We are developing more efficient algorithms.

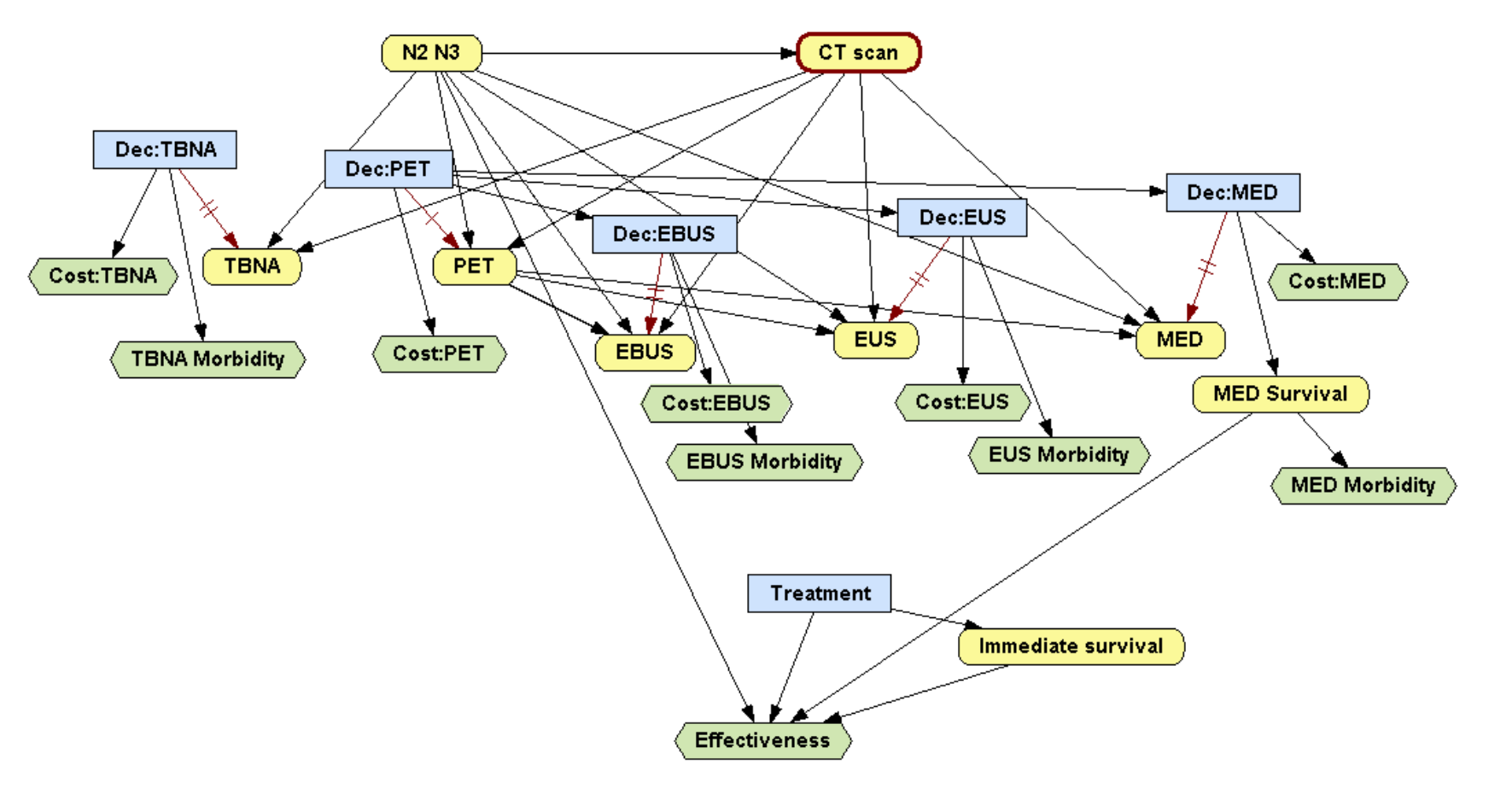
4.2. Examples of decision models for real-world problems

Mediastinet, an ID for lung cancer



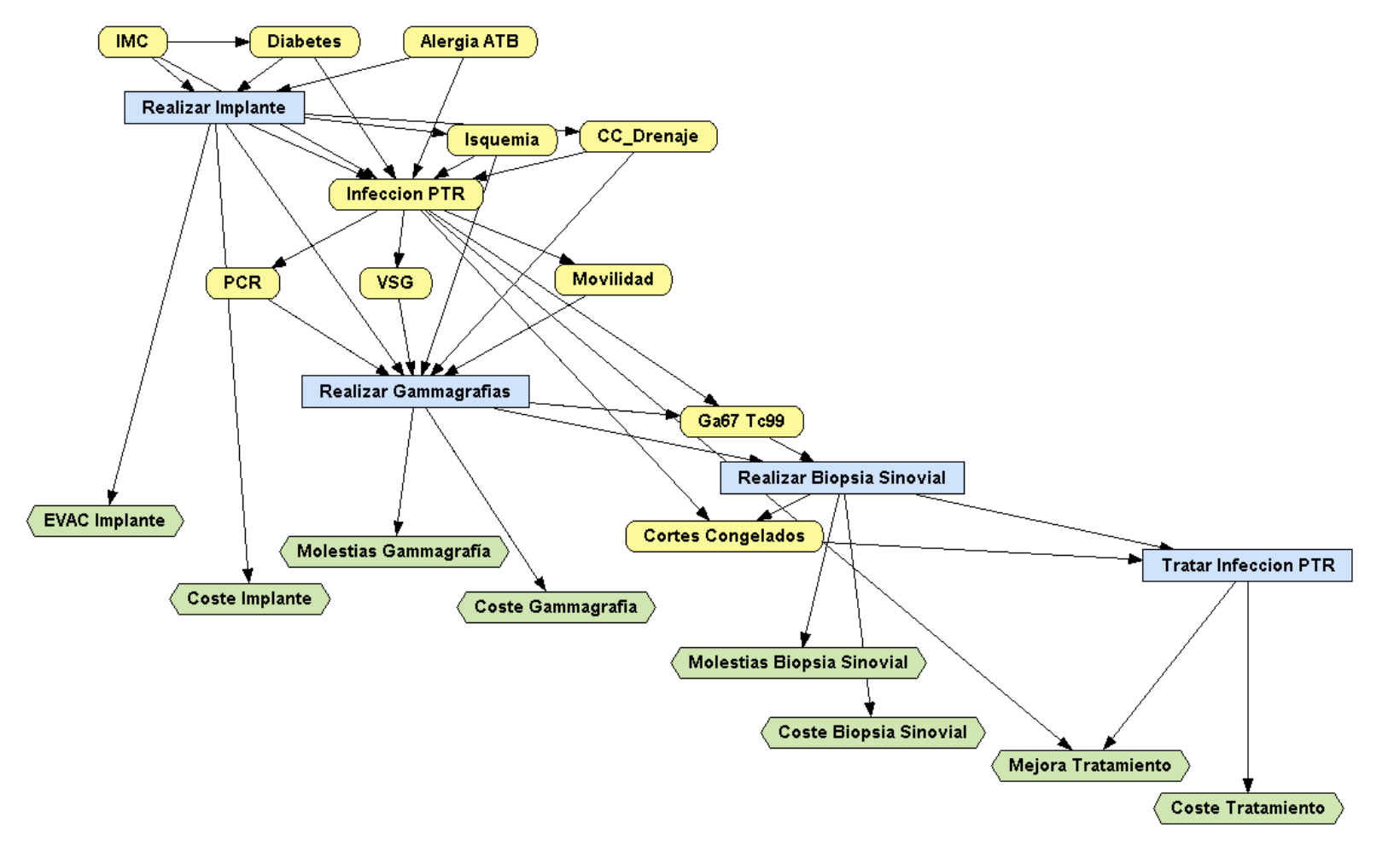
Equivalent to a decision tree containing ~10^4 branches.

Mediastinet (DAN version)



Decisions are partially ordered.

Arthronet, an ID for total knee arthroplasty



Equivalent to a decision tree containing ~10^4 branches.

4.3. Advantages and limitations of influence diagrams

Advantages of influence diagrams (1/3)

- IDs are more compact than decision trees
  - An ID having *n* binary nodes ~ a DT having 2^n branches
- IDs transform automatically into decision trees
  - ... but the reverse is not true (no general algorithm)
  - If you build a decision tree, you only have a decision tree.
  - If you build an ID, you have both.
- IDs are much easier to build than decision trees
  - IDs use direct probabilities (prevalence, sensitivity, specificity...) and costs (mortality, morbidity, economic cost...)
  - ID can use canonical models (noisy OR, noisy AND, etc.)
  - Each parameter appears only once in the ID
    - in many cases it is not necessary to have parametric variables
  - IDs can have several value nodes: more clarity, separate criteria

Advantages of influence diagrams (2/3)

- No external pre-calculation of probabilities is required
- Having all the information, no debugging is usually needed
  - On the contrary, "all trees have bugs" (Primer on MDA, at *MDM* journal)
- IDs are much easier to modify than decision trees
  - Refine the model with new decisions and chance variables
  - Structural sensitivity analysis
  - Can adapt to different regional settings
  - Can adapt to patient's medical characteristics and preferences
- Explicit representation of causality
  - a link indicates causal influence
  - the absence of a link means "no causal influence" (hypothesis)

Advantages of influence diagrams (3/3)

- Two possibilities of evaluation:
  1. expansion of an equivalent decision tree
     - exponential complexity (time and space)
     - equivalent to the brute-force method for Bayesian networks
     - many problems can not be solved with this method
  2. operations on the ID (recursive reduction of the ID)
     - direct manipulation of the graph and/or potentials of the ID
     - similar to the best algorithms for Bayesian networks
     - canonical models and the separation of utility nodes can lead to more efficient evaluations
- More possibilities of explanation of reasoning
  - computation of posterior probabilities on the ID (as if it were a BN)
  - value of information (EVPI and other measures) can be computed easily
  - other methods from Bayesian networks and qualitative prob. networks.

  These methods can be used to debug/refine IDs.

DECISION ANALYSIS
Vol. 2, No. 4, December 2005, pp. 229–231
ISSN 1545-8490 | EISSN 1545-8504 | 05 | 0204 | 0229

informs

DOI 10.1287/deca.1050.0054

The Influence of Influence Diagrams on Artificial Intelligence

Craig Boutilier
Department of Computer Science, University of Toronto, Toronto, Ontario, M5S 3G5 Canada, cebly@cs.toronto.edu

Howard and Matheson's article "Influence Diagrams" has had a substantial impact on research in artificial intelligence (AI). In this perspective, I briefly discuss the importance of influence diagrams as a model for decision making under uncertainty in the AI research community; but I also identify some of the less direct, but no less important, influences this work has had on the field.

Key words: influence diagrams; decision theory; artificial intelligence; value of information; graphical models; perspective, the focus on graphical modeling research
History: Received on November 14, 2005. Accepted by Eric Horvitz on November 23, 2005, without revision.

Howard and Matheson's (1984/2005) "Influence Diagrams" has had a profound impact on developments in artificial intelligence. Some of these influences have been quite direct; others are more indirect, but in many ways, more substantial. The paper itself is representative of developments that had been

vision (Binford and Levitt 2003), dialog management, user interface design, multiagent systems, and game theory (Koller and Milch 2003), to name but a few.

Another reasonably direct impact of "Influence Diagrams" derives from its role in the development of graphical models for probabilistic modeling and

DECISION ANALYSIS
Vol. 2, No. 4, December 2005, pp. 238–244
ISSN 1545-8490 | EISSN 1545-8504 | 05 | 0204 | 0238

informs

DOI 10.1287/deca.1060.0060

The Influence of Influence Diagrams in Medicine

Stephen G. Pauker, John B. Wong
Division of Clinical Decision Making, Informatics and Telemedicine, Department of Medicine,
Tufts–New England Medical Center, Tufts University School of Medicine, 750 Washington St., NEMC 302,
Boston, Massachusetts 02111 {spauker@tufts-nemc.org, jwong@tufts-nemc.org}

Although influence diagrams have used medical examples almost from their inception, that graphical representation of decision problems has disseminated surprisingly slowly in the medical literature and among clinicians performing decision analyses. Clinicians appear to prefer decision trees as their primary modeling metaphor. This perspective examines the use of influence diagrams in medicine and offers explanations and suggestions for accelerating their dissemination.

Key words: decision analysis; influence diagrams; clinical decision making; medicine
History: Received December 12, 2005. Accepted by Eric Horvitz on January 5, 2006, after 1 revision.

Introduction

Two decades after Howard's landmark paper (Howard and Matheson 1984/2005) that introduced the concept of the influence diagram and three decades since Miller's initial report (Miller et al. 1976), *Decision Analysis* reproduced that paper in 2005 and solicited a set of commentaries. This paper

modeling paradigm slowly spread from Stanford, both with courses offered at meetings of the Society for Medical Decision Making (Society for Medical Decision Making 2005) and with the development of software that could conveniently capture and evaluate such models.

Clinical practice guidelines (CPGs)

- Construction of CPGs
  - Usually: expert opinion or consensus of experts
  - Another possibility: probabilistic graphical models
    - Sanders, Nease, Owens: several papers on building CPGs from IDs.
- Advantages of a PGM wrt a traditional CPG
  - explicit decision model
    - combines expert opinions and evidence (statistical data)
    - helps in difficult cases, in which the policy is not evident for experts
  - flexibility: can be extended and adapted, as mentioned above
  - can include patients' preferences
  - the physician plays an active role,
    he/she is not a passive user of CPGs developed by others.

A proverb

- Don’t give a man a fish;
  give him a rod
  and teach him how to fish.

- Don’t give a doctor a written CPG;
  give him/her a DAN
  and teach him/her how to use OpenMarkov.

IDs in the literature on MDM (1/3)

- Books that mention decision trees but do not mention IDs
  - Weinstein, Fineberg. *Clinical Decision Making*. 1980.
  - Sloan (ed.). *Valuing Health Care*. 1995.
  - Gold et al. *Cost-Effectiveness in Health and Medicine*. 1996.
  - Sackett et al. *Evidence-Based Medicine*. 1997 (and three other books on EBM).
  - Petitti. *Meta-Analysis, Decision Analysis and CEA*. 2nd ed., 2000.
  - Drummond, McGuire (eds.). *Economic Eval. in Health Care Programs*. 2001.
  - Levin and McEwan. *Cost-Effectiveness Analysis*. 2nd ed., 2001.
  - Parmigiani. *Modelling in Medical Decision* Making. 2002.
  - Haddix et al. *Prevention Effectiveness*. 2nd ed., 2003.
  - Fox-Rushby and Cairns. *Economic Evaluation.* 2005.
  - Briggs et al. *Decision Modelling for Health Economic Evaluation*, 2006.
  - Alemi and Gustafson. *Decision Analysis for Healthcare Managers*, 2006.
  - Arnold. *Pharmacoeconomics: From Theory to Practice*. 2009.
  - Kassirer et al. *Learning Clinical Reasoning*. 2nd ed., 2010.
  - Mushlin and Greene. *Decision Making in Medicine*. 3rd ed., 2010.

(cont'd)

IDs in the literature on MDM (2/3)

- Books that mention decision trees but do not mention IDs (cont.)
  - Gray et al. *Applied Methods of CEA in Health Care*, 2011.
  - Alfaro-LeFevre. *Critical Thinking, Clinical Reasoning…* 5th ed., 2013.
  - Morris et al. *Economic Analysis in Healthcare*. 2nd ed., 2012.
  - Rascati. *Essentials of Pharmacoeconomics*. 2nd ed., 2013.
  - Sox et al. *Medical Decision Making*. Latest ed., 2013.
  - Hunink et al. *Decision Making in Health and Medicine*. 2nd ed., 2014.
  - Drummond et al. *Methods for the Economic Evaluation of…* 4th ed. 2015.
  - Edlin et al. *Cost Effectiveness Modelling for HTA…* 2015.
  - Neumann et al. *Cost-Effectiveness in Health and Medicine*. 2016
  - Caro et al. *Discrete Event Simulation for HTA*. 2016

- One book that mentioned IDs
  - Muennig. *Designing and Conducting Cost-Effectiveness Analyses in Medicine and Health Care*. 2002, page 242:

    "An influence diagram (also known as a tornado diagram) ..."

    The 2nd edition (2007) and the 3rd (2016) do not mention them.

IDs in the literature on MDM (3/3)

- Three books that describe IDs
  - Chapman and Sonnenberg (eds.). *Decision Making in Health Care*. 2000 (5 pages out of 421, in a chapter authored by Mark Roberts)
  - Schwartz and Bergus. *Medical Decision Making. A Physician's Guide*. 2008. (2 pages out of 230)
  - Kattan. *Encyclopedia of Medical Decision Making*. 2009 (4 pages out of 1200+).
- Summary of the informal survey of books on MDM and EBM
  - 26 books published after 1984
  - All of them explain DTs but only 3 describe IDs, very briefly.
- Some books on medical informatics mention IDs:
  - Shortliffe and Cimino. *Biomedical Informatics*. 4th ed., 2013 (2.5 pages out of 991).
  - Kalet. *Principles of Biomedical Informatics*. 2nd ed., 2013 (3 pages out of 708).
- Why are IDs so little known in health sciences after 30+ years?

Limitations of IDs

1. The “reasoning” of an ID is not easy to understand
2. The evaluation returns large policy tables
3. IDs can only model symmetric problems
   - IDs require a total ordering of the decisions
   - IDs cannot represent incompatibilities between values
     - Non-standard versions of IDs partially solve this problem, but none of the alternatives was completely satisfactory.
4. Algorithms could only evaluate unicriterion IDs
   - They could not perform cost-effectiveness analysis
5. Temporal reasoning was not possible with IDs
   - Dynamic IDs are computationally unfeasible.

Solutions we have proposed

1. Explanation in influence diagrams
   - showing the posterior probabilities and expected values
   - introduction of evidence
   - hypothetical reasoning (what if) by means of imposed policies
2. Synthesizing the optimal intervention
   - in the form of a compact tree
3. Decision analysis networks
   - an alternative to IDs for asymmetric decision problems.
4. Cost-effectiveness analysis with IDs
5. Markov influence diagrams
   - including cost-effectiveness analysis

International Journal of Approximate Reasoning 96 (2018) 1–17

Contents lists available at ScienceDirect

International Journal of Approximate Reasoning

www.elsevier.com/locate/ijar

Decision analysis networks

Francisco Javier Díez *, Manuel Luque, Iñigo Bermejo

Dept. Artificial Intelligence, Universidad Nacional de Educación a Distancia (UNED), Juan del Rosal 16, 28040 Madrid, Spain

ARTICLE INFO

Article history:
Received 7 July 2017
Received in revised form 15 December 2017
Accepted 21 February 2018
Available online 27 February 2018

Keywords:
Decision analysis
Decision trees
Influence diagrams
Probabilistic graphical models
Asymmetric decision problems

ABSTRACT

This paper presents decision analysis networks (DANs) as a new type of probabilistic graphical model. Like influence diagrams (IDs), DANs are much more compact and easier to build than decision trees and can represent conditional independencies. In fact, for every ID there is an equivalent symmetric DAN, but DANs can also represent asymmetric problems involving partial orderings of the decisions (order asymmetry), restrictions between the values of the variables (domain asymmetry), and conditional observability (information asymmetry). Symmetric DANs can be evaluated with the same algorithms as IDs. Every asymmetric DAN can be evaluated by converting it into an equivalent decision tree or, much more efficiently, by decomposing it into a tree of symmetric DANs. Given that DANs can solve symmetric problems as easily and as efficiently as IDs, and are more appropriate for asymmetric problems—which include virtually all real-world problems—DANs might replace IDs as the standard type of probabilistic graphical model for decision support and decision analysis. We also argue that DANs compare favorably with other formalisms proposed for asymmetric decision problems. In practice, DANs can be built and evaluated with OpenMarkov, a Java open-source package for probabilistic graphical models.

1. Introduction

The two formalisms most widely used for the representation and analysis of decision problems are decision trees (DTs) [31] and influence diagrams (IDs) [15]. DTs have the advantage of almost absolute flexibility, but also have three drawbacks: their size grows exponentially with the number of variables, they cannot represent conditional independencies, and they

DANs vs. IDs

- A DAN is symmetric if:
  - it has no restrictions
  - if a value of X reveals Y, then every value of X reveals Y
- DANs can replace IDs as the standard decision analysis tool (in AI, MDM, operations research…) because:
  - For every ID there is an equivalent symmetric DAN
    - but for many DANs there is no equivalent ID
  - Virtually all real-world problems are asymmetric.
  - There many problems that cannot be modeled with IDs.
  - Even if a problem can be modeled with an ID, a DAN is usually better because it does not need dummy states.

DANs vs. IDs

- DANs can replace IDs as the standard decision analysis tool (in AI, MDM, operations research…) because:
  - For every ID there is an equivalent symmetric DAN
    - but for many DANs there is no equivalent ID
  - Virtually all real-world problems are asymmetric.
  - There many problems that cannot be modeled with IDs.
  - Even if a problem can be modeled with an ID, a DAN is usually better because it does not need dummy states.

5. Multicriteria decision making

5.1. Effectiveness and utility in medicine

Economic evaluation in medicine

- Objective:

  to decide whether the benefit of an intervention outweighs its economic cost.

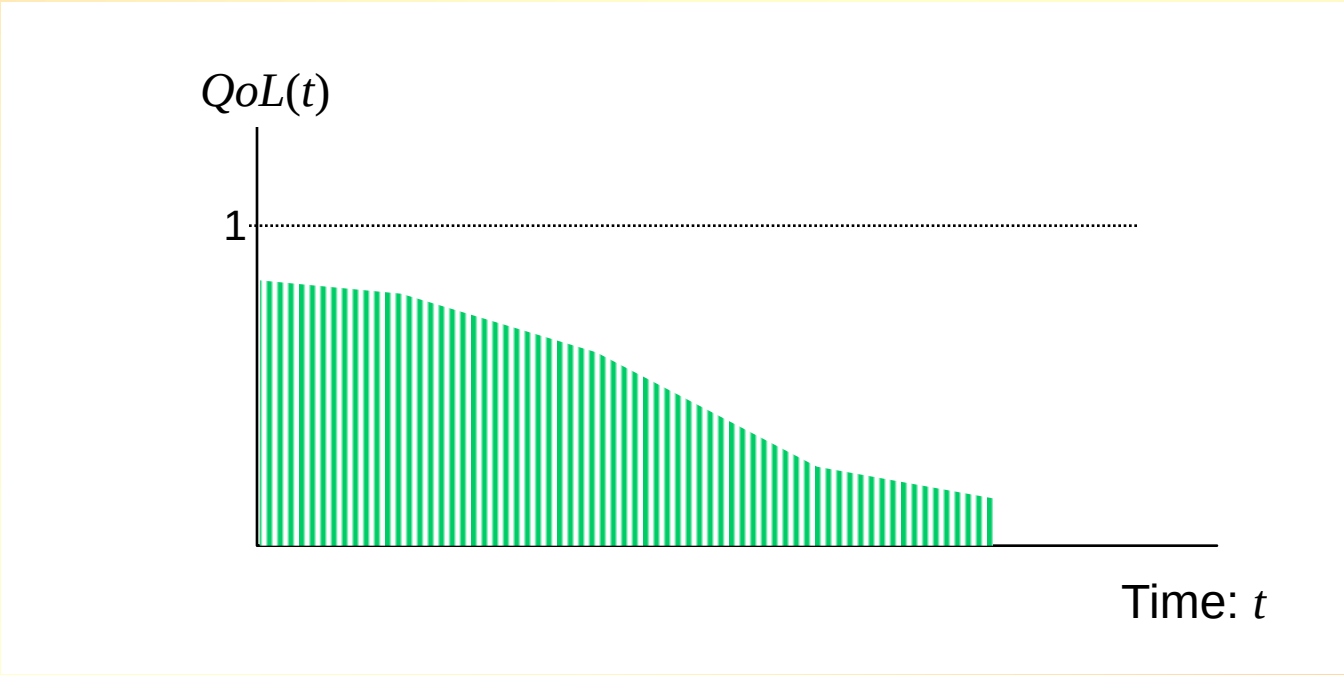
- Three types of analysis:

  - **Cost-benefit**
    - Health benefits are converted into monetary units
  - **Cost-effectiveness**
    - Benefits are measured in medical units, such as lives saved, life years gained, detected cases, etc.
  - **Cost-utility**
    - Benefit is measured in quality-adjusted life years (QALYs).

Quantity and quality of life

- Effectiveness (in cost-utility studies):

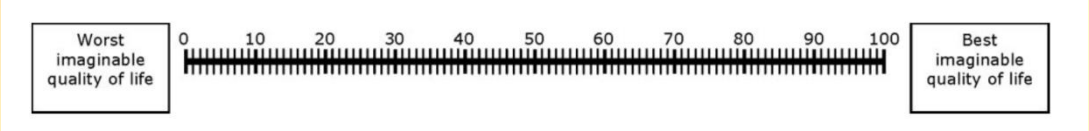
$$eff = \int QoL(t) \cdot dt$$



  - 1 QALY = effectiveness accrued in one year of perfect health

QoL is subjective: how can we measure it?

- Visual analog scale (VAS)

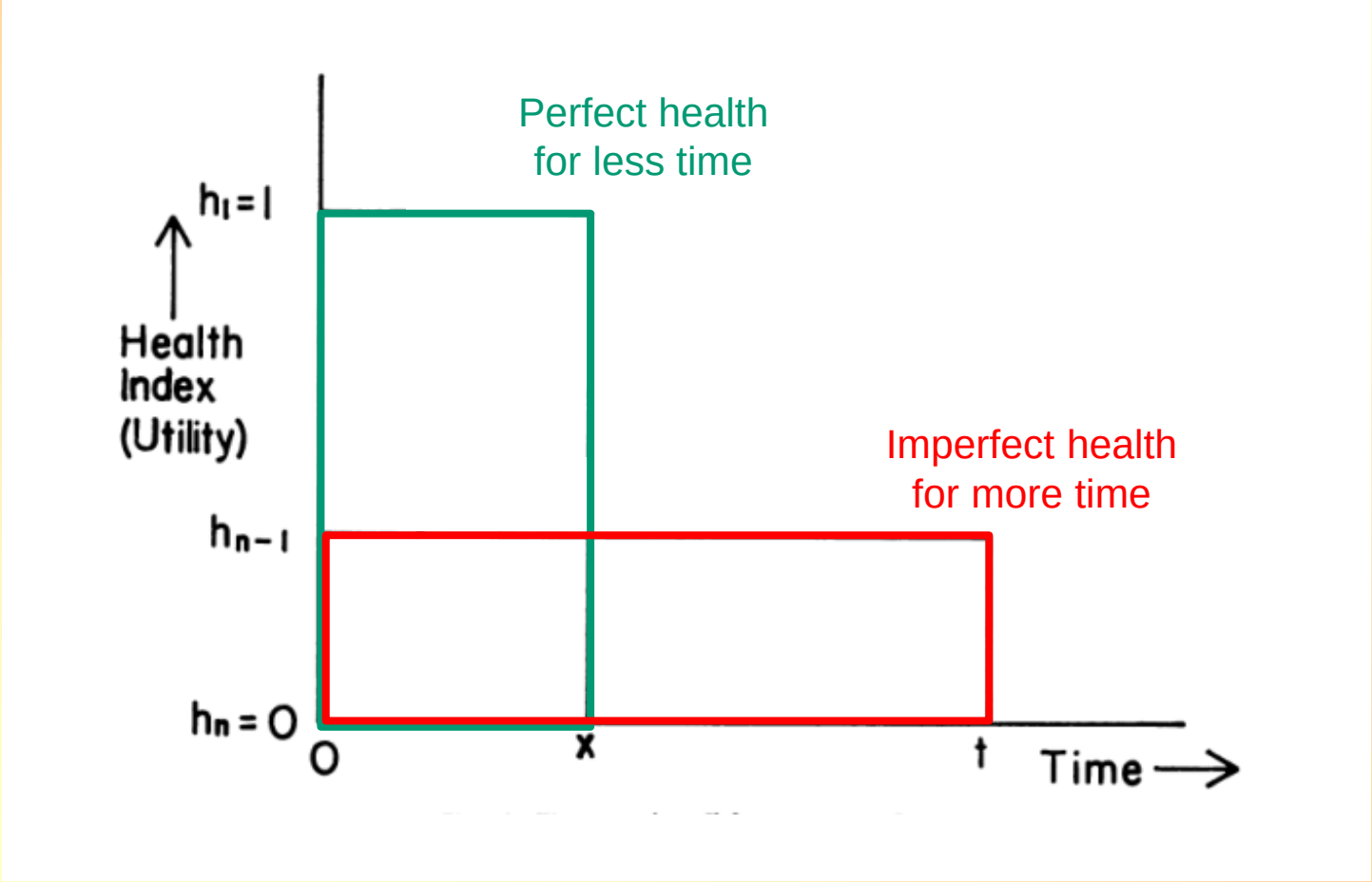


  - does not measure quantitative preferences
  - cannot be directly used in cost-utility analyses

- Standard gamble
  - "Do you prefer to live in state s or to enter a lottery with probability p of recovering perfect health and $(1- p)$ of dying?"

- Time trade-off (TTO)
  - "Do you prefer to live in state s for 50 years or do you prefer to live with perfect health for 45 years?"
  - "Imagine you are in state s and your life expectancy is 50 years. How many years of your life expectancy would you give up to recover perfect health?

Trade-off between quantity and quality of life



Torrance, Thomas and Sackett (1972)

Quality of life indexes (indices)

- Every index considers a reduced number of attributes (dimensions)
  - HUI-3 has 8 attributes: sight, hearing, ability to converse, ability to walk, manual dexterity, emotional status, cognitive ability and pain
- Every attribute has a limited number of states
  - "Hearing" in the HUI-3 has 6 states (see next slide)
- Every individual is characterized by a tuple of states
  - In HUI-3, it is an 8-tuple. Example: (4,5,5,6,4,3,4,5).
- A mathematical function maps each configuration onto a number
  - $f(4,5,5,6,4,3,4,5) = 0.742$.
- Function f, specific for each index, is calibrated using a preference-elicitation method: standard gamble or time trade-off.

QoL tables

Health state	Utility
Healthy (reference state)	1.00
Life with menopausal symptoms (judgment)	0.99
Side effects of hypertension treatment (judgment)	0.95–0.99
Mild angina (judgment)	0.90
Kidney transplant (TTO, Hamilton, patients with transplants)	0.84
Moderate angina (judgment)	0.70
Some physical and role limitation with occasional pain (TTO)	0.67
Hospital dialysis (TTO, Hamilton, dialysis patients)	0.59
Hospital dialysis (TTO, St John's, dialysis patients)	0.57
Hospital dialysis (TTO, general public)	0.56
Severe angina (judgment)	0.50
Anxious/depressed and lonely much of the time (TTO)	0.45
Being blind or deaf or dumb (TTO)	0.39
Hospital confinement (TTO)	0.33
Mechanical aids to walk and learning disabled (TTO)	0.31
Dead (reference state)	0.00
Quadriplegic, blind and depressed (TTO)	<0.00
Confined to bed with severe pain (ratio)	<0.00
Unconscious (ratio)	<0.00

From: Torrance (1987). Utility approach to measuring…

An example with two criteria

- Two therapies
  - Effectiveness (QALY)

	No therapy	Therapy 1	Therapy 2
Disease present	1.2	4.0	6.5
Disease absent	10	9.9	9.3

  - Therapy 1 cost = 20,000 €
  - Therapy 2 cost = 70,000 €
- Questions:
  - What therapy to apply when the disease is present
  - What therapy to apply when the disease is absent
- Problem: how to compare health and money

5.2. Combining cost and effectiveness into a single criterion

Net benefit

- Net monetary benefit: $NMB = \lambda \cdot E - C$
  - *E* = effectiveness, usually measured in QALYs (utility)
  - *C* = cost, in monetary units (€, £, $…)
  - **λ = willingness to pay = cost-effectiveness threshold**
  - λ is usually measured in $/QALY, €/QALY, £/QALY…
  - It converts effectiveness into monetary units
  - It is specific for each decision maker
- When comparing two or more interventions/strategies, which one is more beneficial?
  – It may depend on λ

NMB as a function of λ

- If λ = 6,000 €/QALY:

NMB	No therapy	Therapy 1	Therapy 2
Disease present	7,200 €	4,000 €	–31,000 €
Disease absent	60,000 €	39,400 €	–14,200 €

- If λ = 15,000 €/QALY:

NMB	No therapy	Therapy 1	Therapy 2
Disease present	18,000 €	40,000 €	27,500 €
Disease absent	150,000 €	128,500 €	69,500 €

- If λ = 30,000 €/QALY:

NMB	No therapy	Therapy 1	Therapy 2
Disease present	36,000 €	100,000 €	125,000 €
Disease absent	300,000 €	277,000 €	209,000 €

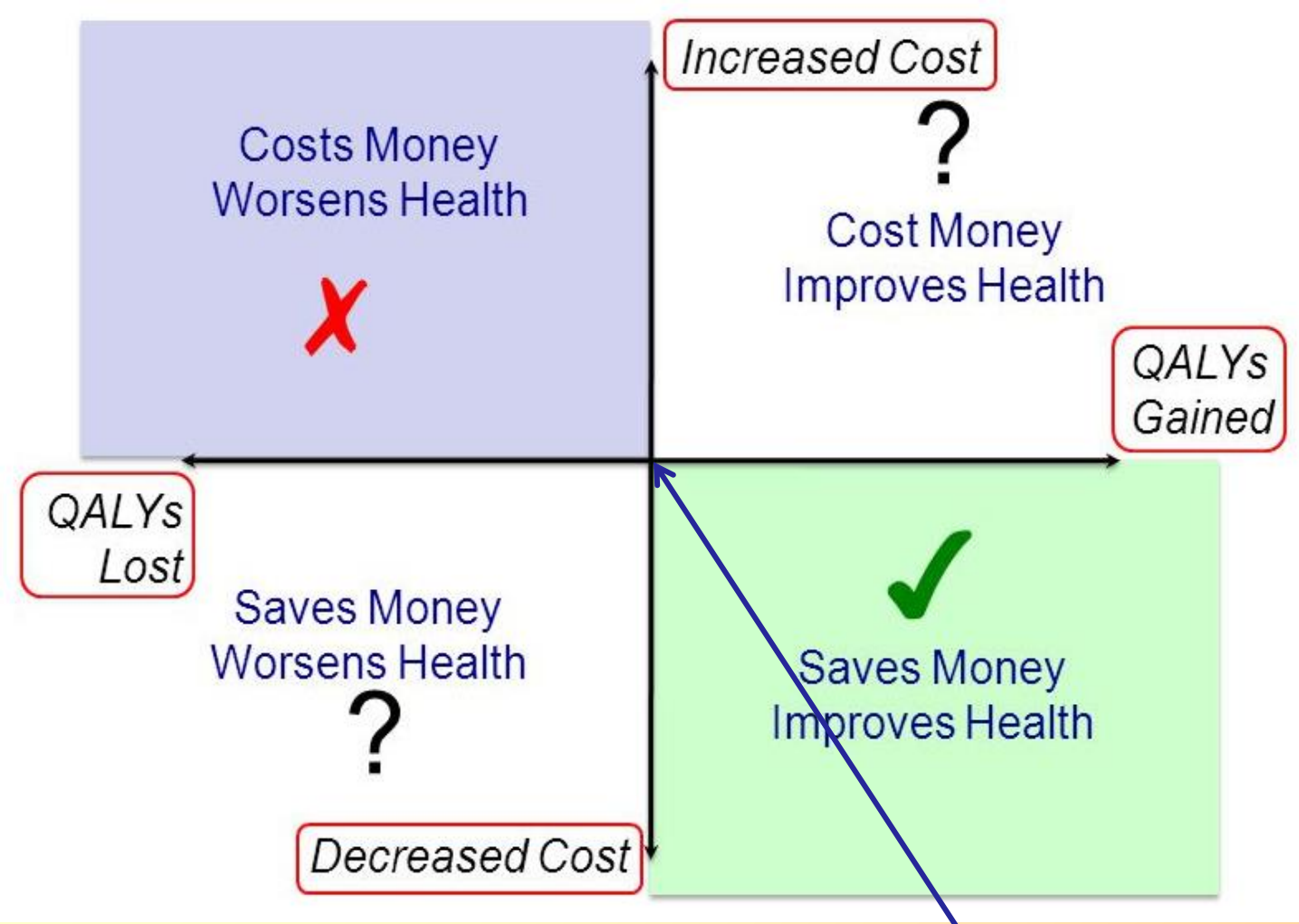
Problem: difficult to estimate the WTP

- λ is different for each decision maker:
  - USA $50,000-100,000 / QALY
  - UK £20,000-30,000 / QALY
  - Spain, Italy ~ €30,000/QALY
  - Norway ~ €70,000/QALY
  - WHO ~3 × (annual per capita GDP) / DALY
  - In some countries the range of variation is very wide.
- How to estimate it?
  1. Shadow threshold: what interventions are covered in a country
  2. Econometric methods
  - No consensus among health economists
- What value of λ should we use in our analyses?
- Solution (partial solution): cost-effectiveness analysis

5.3. Cost-effectiveness analysis

5.3.1. Deterministic CEA

Cost-effectiveness plane



Incremental cost-effectiveness ratio (ICER)

- One intervention is more effective but more expensive

$$NHB_1 = \lambda \times E_1 - C_1$$

$$NHB_2 = \lambda \times E_2 - C_2$$

$$NHB_1 > NHB_2 \Leftrightarrow \frac{C_2 - C_1}{E_2 - E_1} < \lambda$$

- Def.: Incremental cost-effectiveness ratio (ICER)

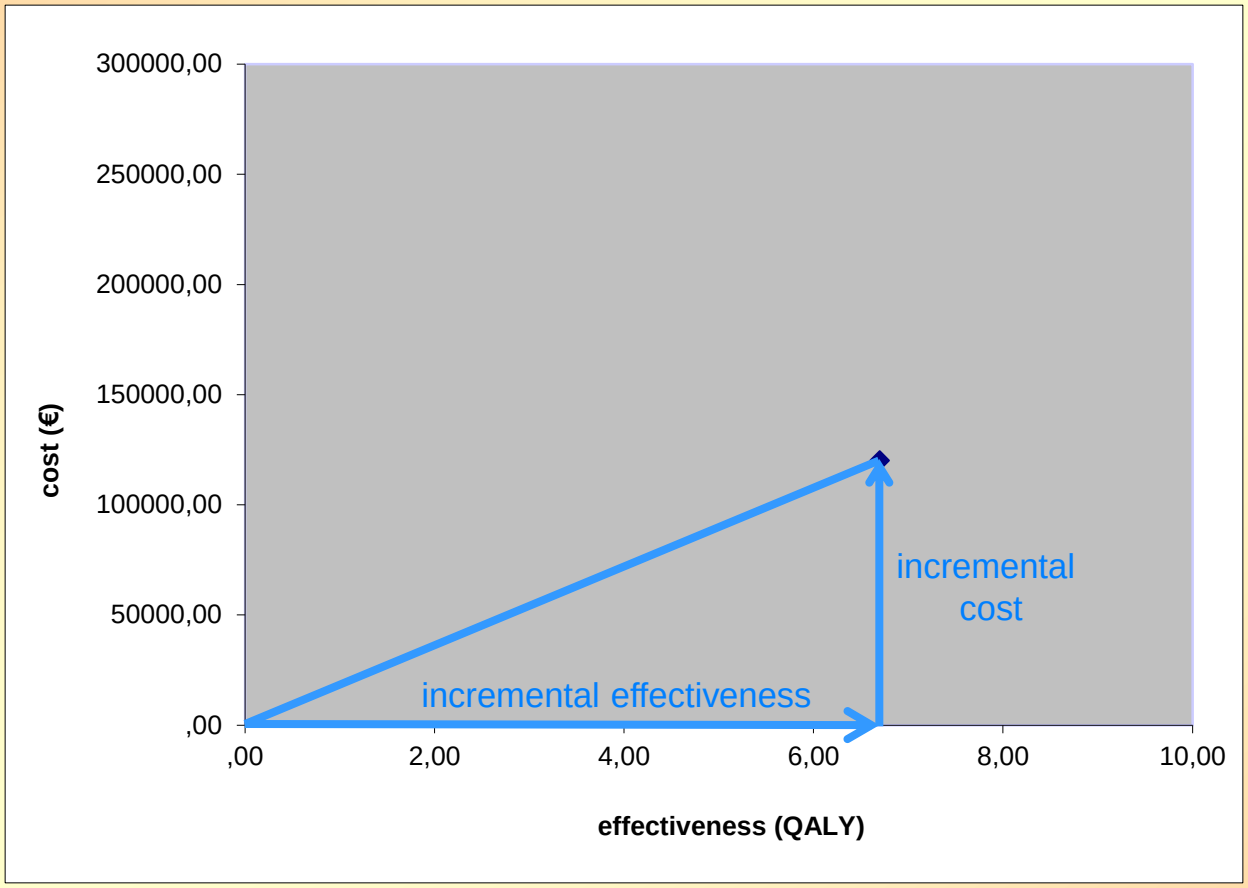
$$ICER_{2,1} = \frac{C_2 - C_1}{E_2 - E_1}$$

- Conclusion

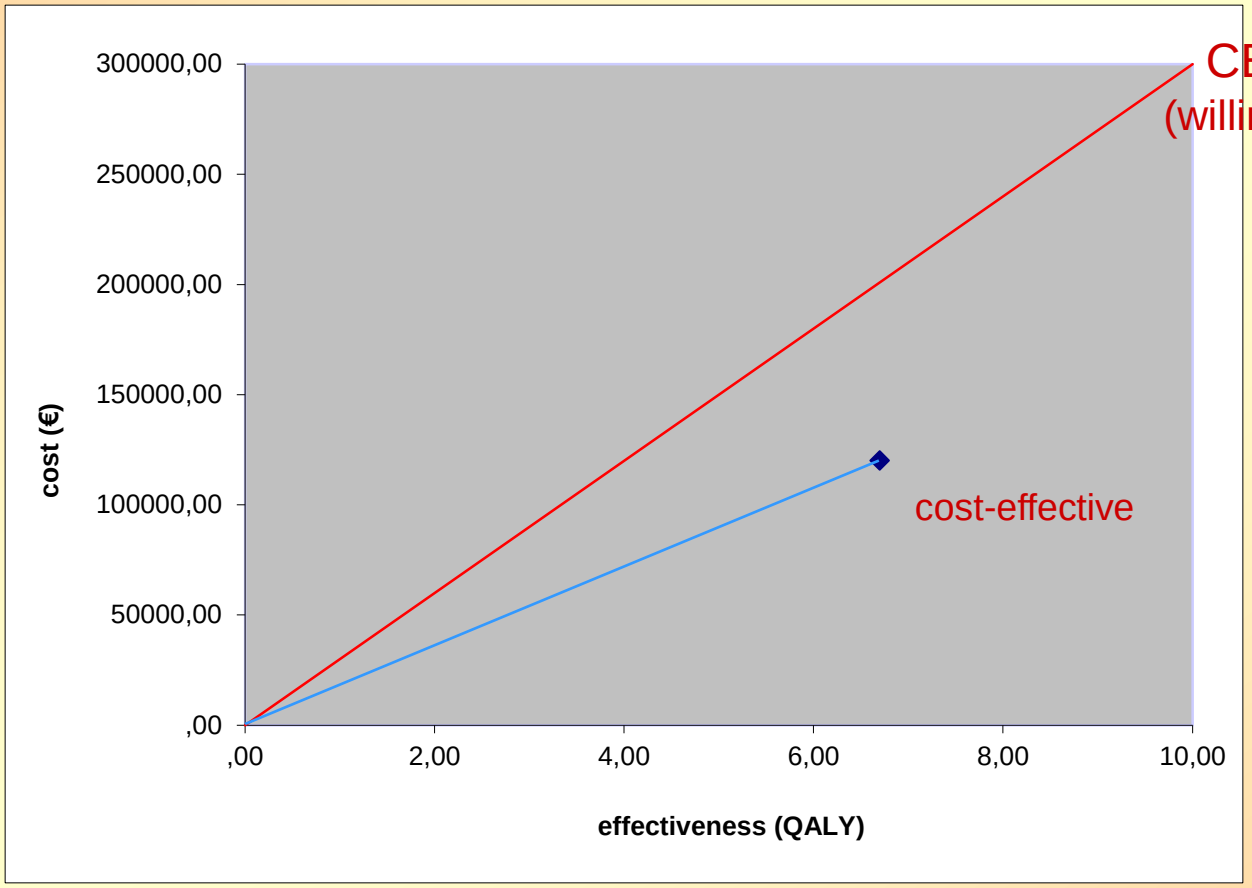
$$NHB_2 > NHB_1 \Leftrightarrow ICER_{2,1} < \lambda$$

- λ, the WTP, determines which option is more beneficial

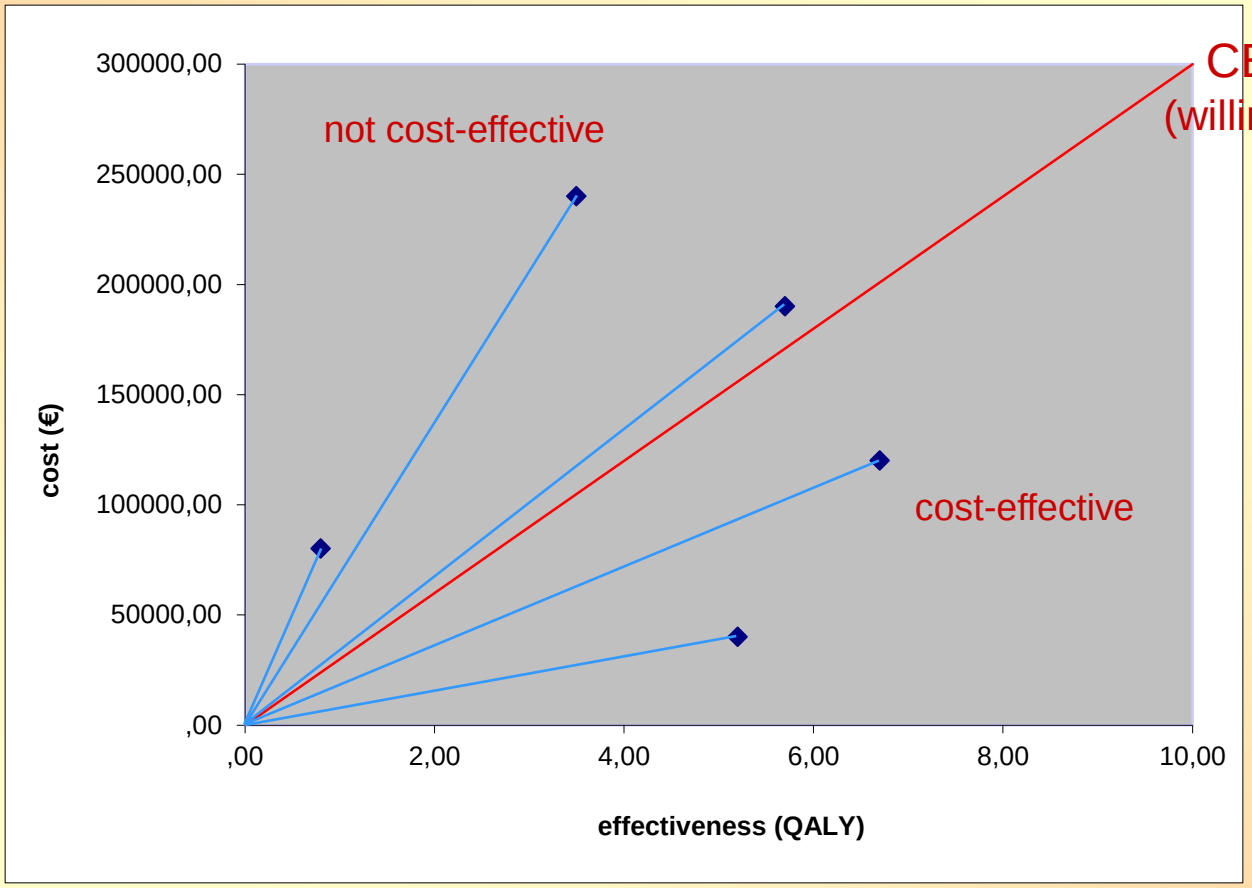
Why does the ICER matter?



Why does the ICER matter?

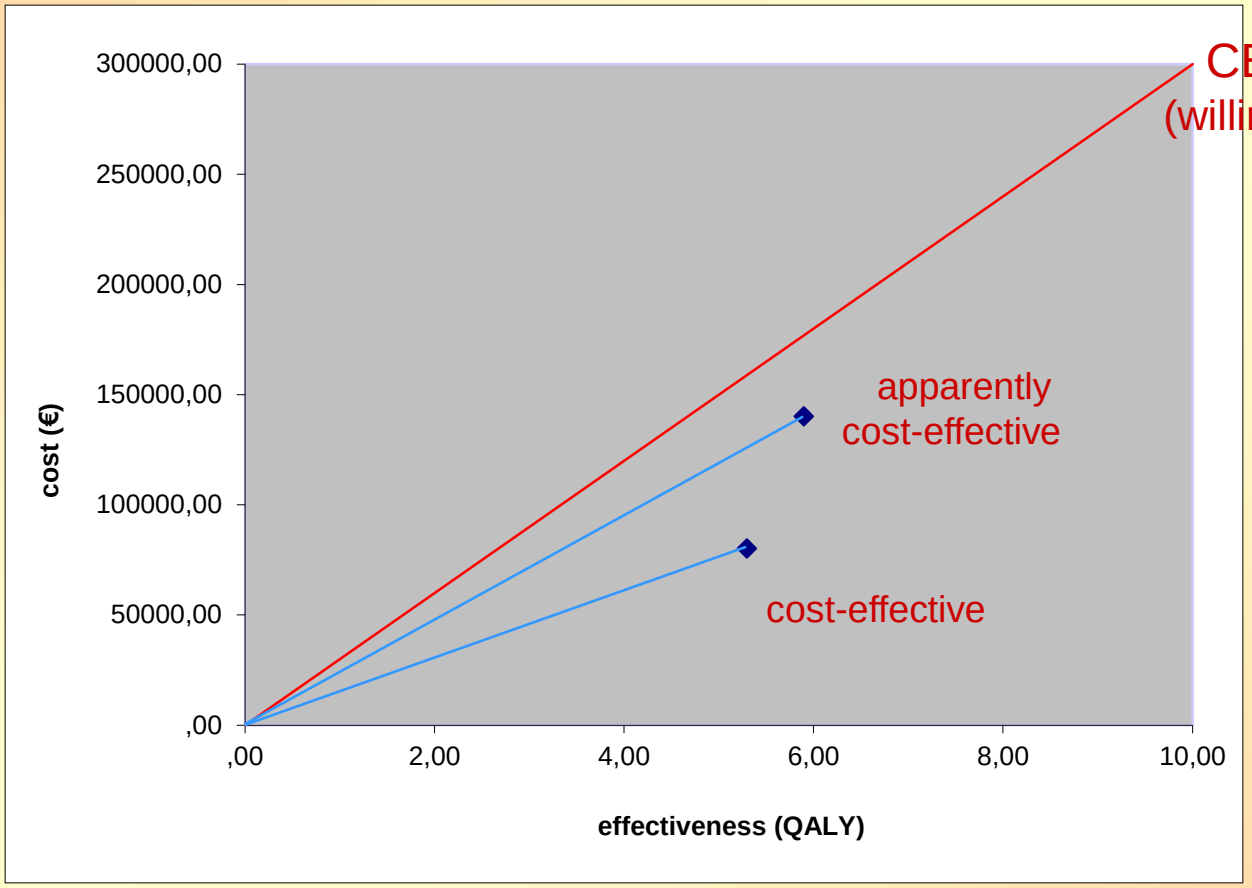


Why does the ICER matter?



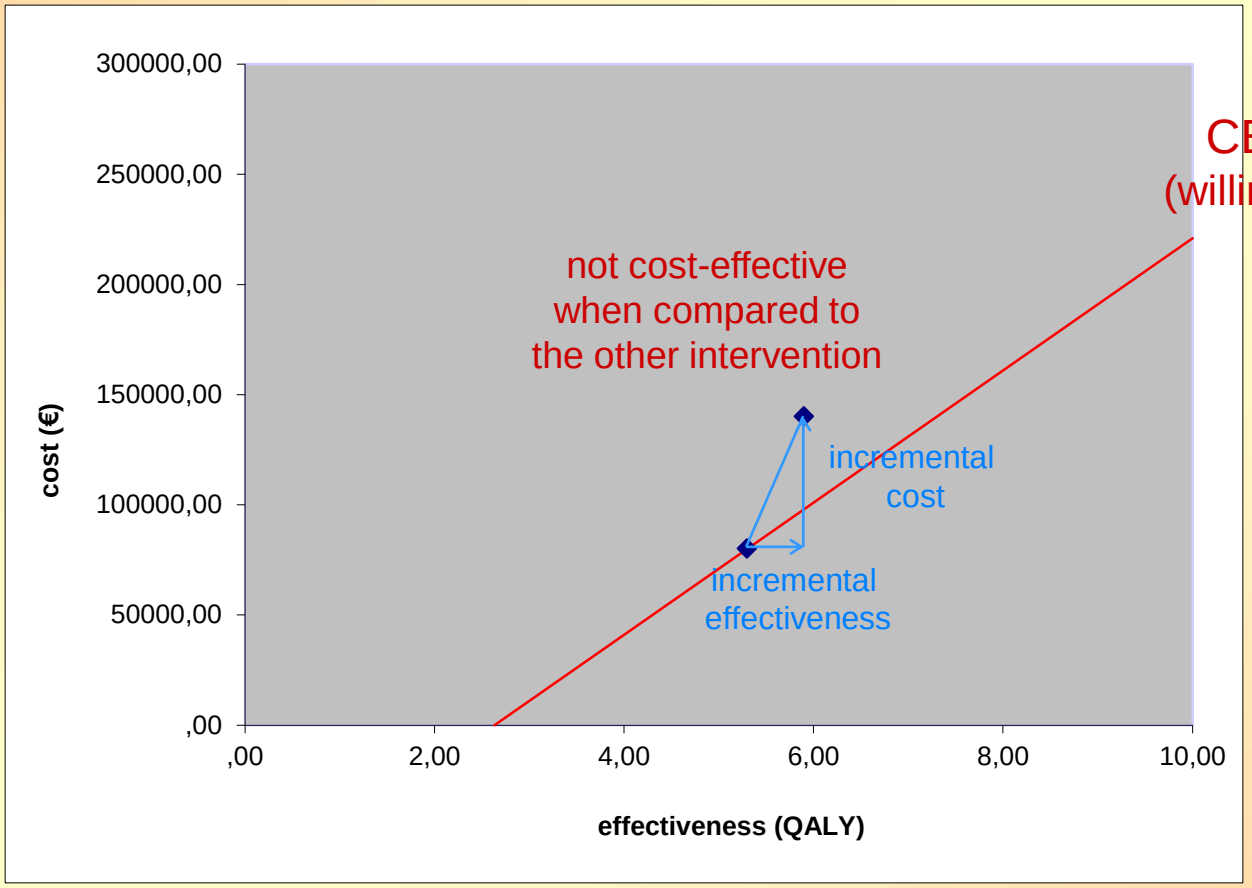
- Two of the interventions are cost-effective wrt no intervention.

Why does the ICER matter?



- If they are mutually exclusive, they must be compared pairwise.

Why does the ICER matter?



- It is essential to select the right **comparator**.

5.3.1.1. Deterministic CEA for the previous example

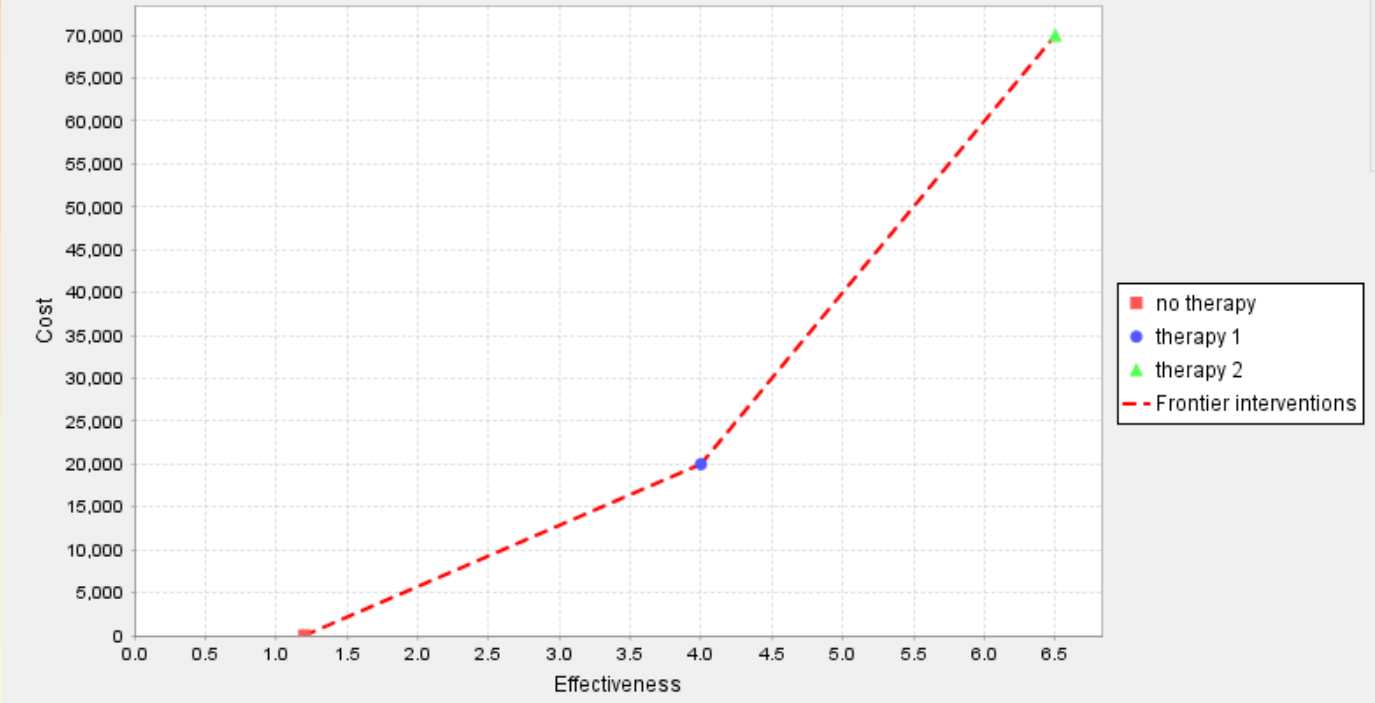
The example we are considering

- Two therapies
  - Effectiveness (QALY)

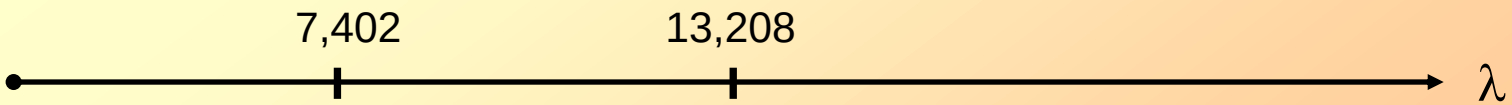
	No therapy	Therapy 1	Therapy 2
Disease present	1.2	4.0	6.5
Disease absent	10	9.9	9.3

  - Therapy 1 cost = 20,000 €
  - Therapy 2 cost = 70,000 €
- Questions:
  - What therapy to apply when the disease is present
  - What therapy to apply when the disease is absent
- Problem: how to compare health and money

When we know that the disease is present



Interval for λ	Cost	Effect.	Best therapy
(0, 7,143)	0	1.2	no-therapy
(7,143, 13,208)	20.000	4.0	therapy-1
(13,208, +∞)	70.000	6.5	therapy-2



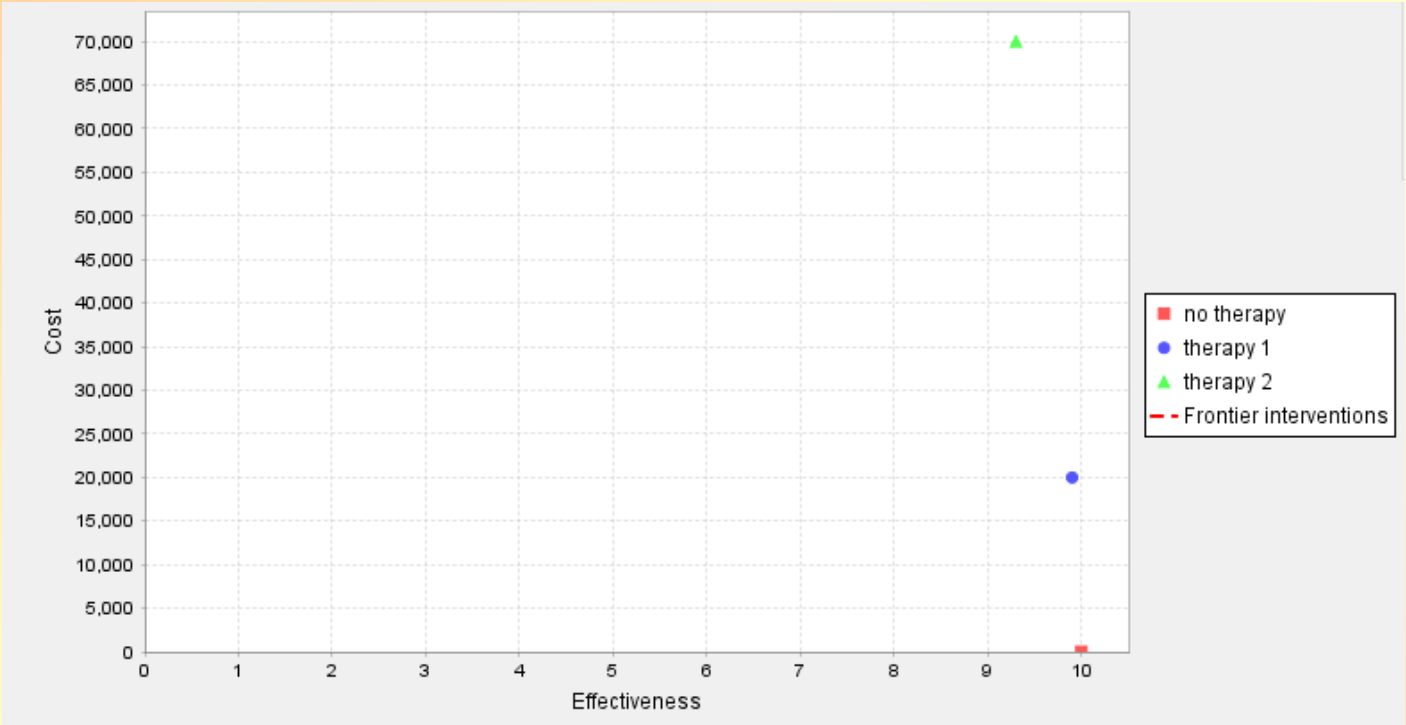
The example we are considering

- Two therapies
  - Effectiveness (QALY)

	No therapy	Therapy 1	Therapy 2
Disease present	1.2	4.0	6.5
Disease absent	10	9.9	9.3

  - Therapy 1 cost = 20,000 €
  - Therapy 2 cost = 70,000 €
- Questions:
  - What therapy to apply when the disease is present
  - What therapy to apply when the disease is absent
- Problem: how to compare health and money

When we know that the disease is absent



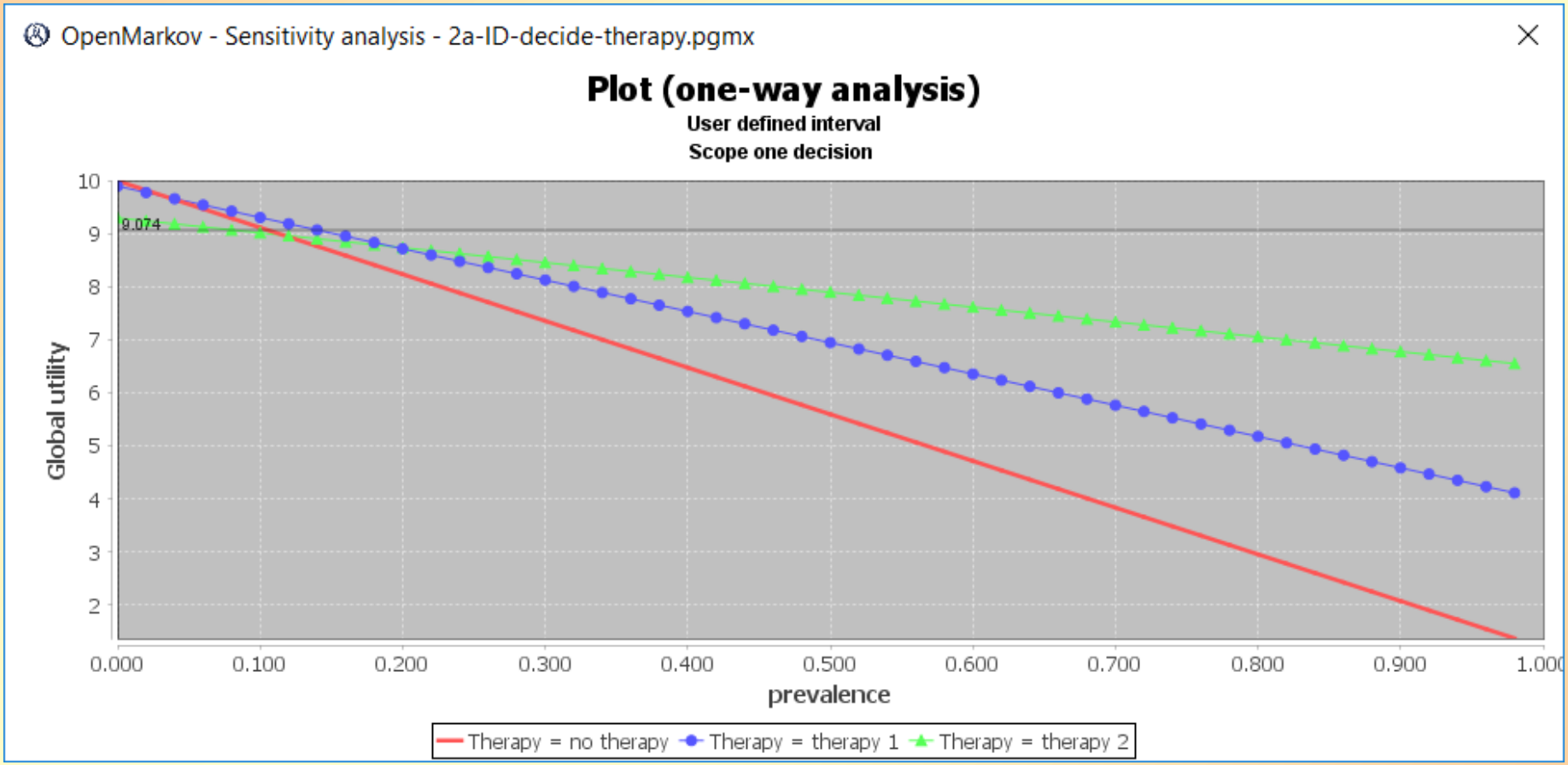
Interval for λ	Cost	Effect.	Best therapy
(0, +∞)	0	10	no-therapy

5.3.2. CEA with uncertain outcomes

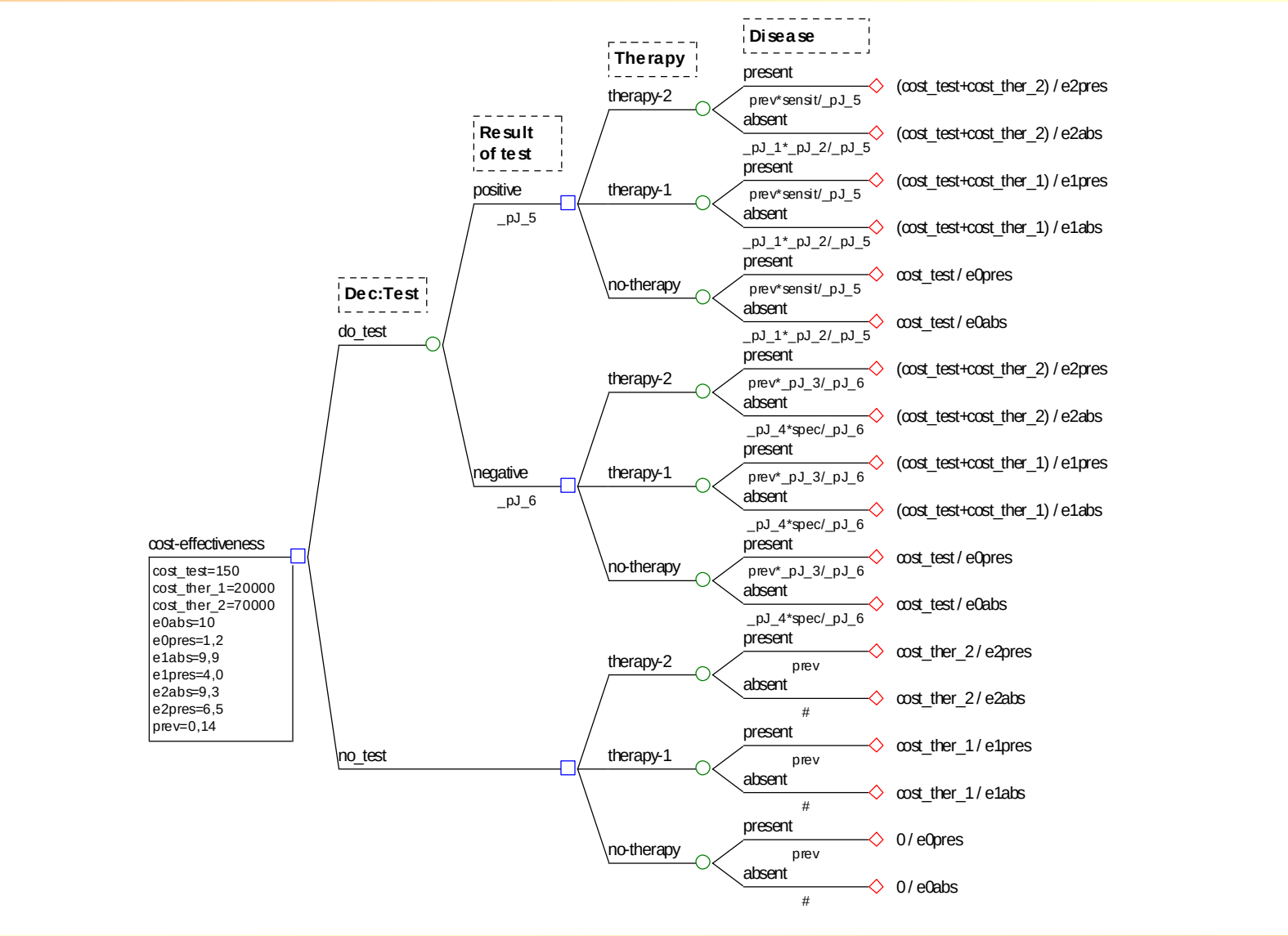
Example with uncertain outcomes: cost-effectiveness of a test

- The costs and effectiveness of the two therapies are the same as in the previous example
- But there is uncertainty (probabilities):
  - prevalence of the disease: 0.14
  - test: sensitivity 0.90
    specificity 0.93
- Also the test has a cost: 150 €
- Questions:
  - When is the test cost-effective? = What is its ICER?
  - What is the most beneficial therapy for each value of λ?

Effectiveness as a function of prevalence



A decision tree for this example



Problem: the standard algorithm only works for the unicriterion case

A warning and a (rudimentary) solution

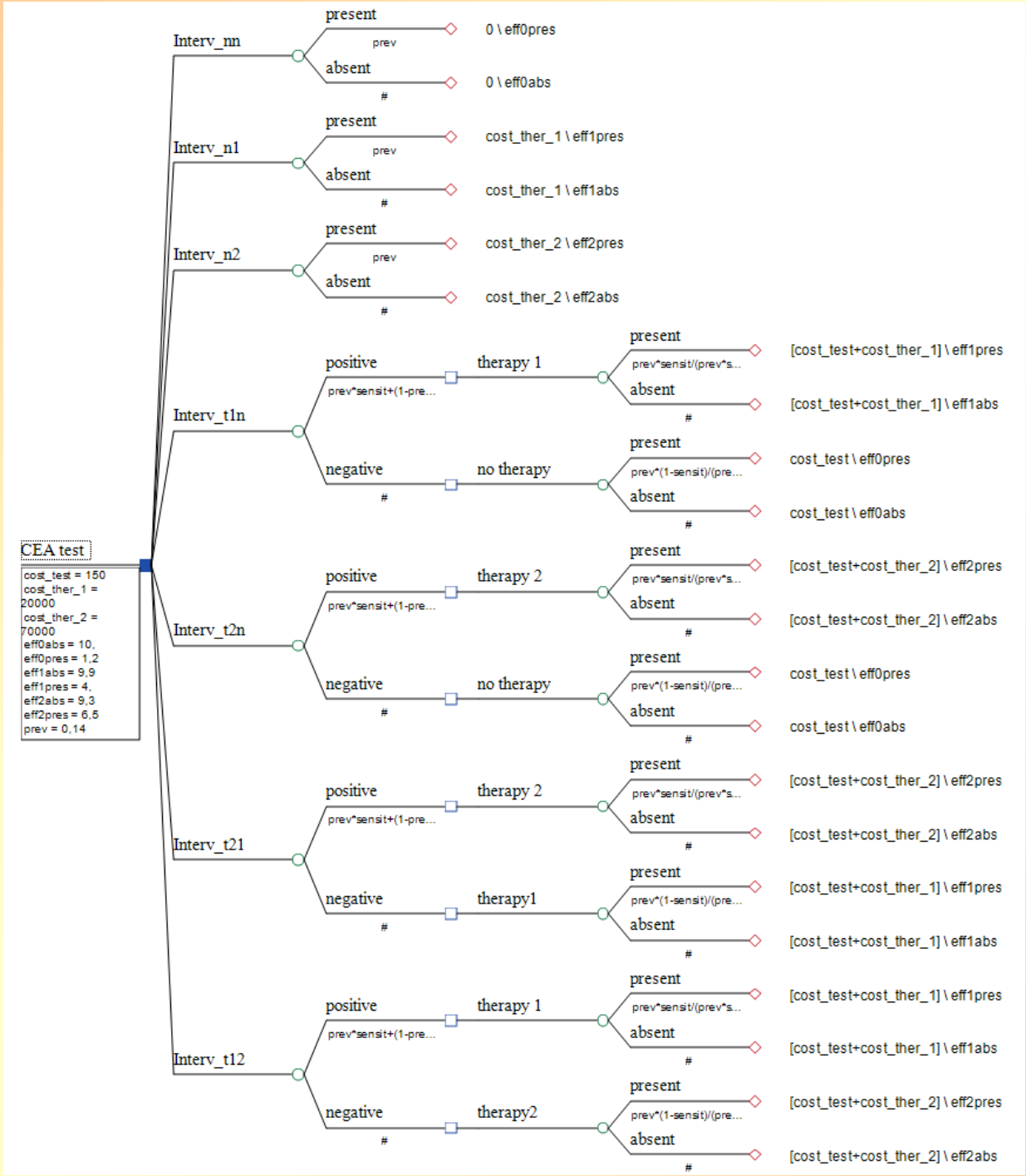
"Embedded, or downstream, decision nodes are not useful in cost-effectiveness analysis because the optimal branch cannot be determined when folding back the tree without an explicit decision rule for comparing costs and consequences.

Cost-effectiveness analyses can be performed with a decision tree that has one decision node at the root. The branches of the initial decision node represent all of the strategies that are to be compared."

Kuntz and Weinstein [2001]

How many strategies for this example?

- Without testing
  - No therapy in any case
  - Always therapy 1
  - Always therapy 2
- With testing
  - ~~If positive, no therapy; if negative, no therapy.~~
  - ~~If positive, no therapy; if negative, therapy 1.~~
  - ~~If positive, no therapy; if negative, therapy 2.~~
  - If positive, therapy 1; if negative, no therapy.
  - ~~If positive, therapy 1; if negative, therapy 1.~~
  - If positive, therapy 1; if negative, therapy 2.
  - If positive, therapy 2; if negative, no therapy.
  - If positive, therapy 2; if negative, therapy 1.
  - ~~If positive, therapy 2; if negative, therapy 2.~~


CEA test
cost_test = 150
cost_ther_1 = 20000
cost_ther_2 = 70000
eff0abs = 10,
eff0pres = 1,2
eff1abs = 9,9
eff1pres = 4,
eff2abs = 9,3
eff2pres = 6,5
prev = 0,14
Interv_nn
present
prev
0 \ eff0pres
absent
#
0 \ eff0abs
Interv_n1
present
prev
cost_ther_1 \ eff1pres
absent
#
cost_ther_1 \ eff1abs
Interv_n2
present
prev
cost_ther_2 \ eff2pres
absent
#
cost_ther_2 \ eff2abs
Interv_t1n
positive
prev*sensit+(1-pre...
therapy 1
present
prev*sensit/(prev*s...
[cost_test+cost_ther_1] \ eff1pres
absent
#
[cost_test+cost_ther_1] \ eff1abs
negative
#
no therapy
present
prev*(1-sensit)/(pre...
cost_test \ eff0pres
absent
#
cost_test \ eff0abs
Interv_t2n
positive
prev*sensit+(1-pre...
therapy 2
present
prev*sensit/(prev*s...
[cost_test+cost_ther_2] \ eff2pres
absent
#
[cost_test+cost_ther_2] \ eff2abs
negative
#
no therapy
present
prev*(1-sensit)/(pre...
cost_test \ eff0pres
absent
#
cost_test \ eff0abs
Interv_t21
positive
prev*sensit+(1-pre...
therapy 2
present
prev*sensit/(prev*s...
[cost_test+cost_ther_2] \ eff2pres
absent
#
[cost_test+cost_ther_2] \ eff2abs
negative
#
therapy1
present
prev*(1-sensit)/(pre...
[cost_test+cost_ther_1] \ eff1pres
absent
#
[cost_test+cost_ther_1] \ eff1abs
Interv_t12
positive
prev*sensit+(1-pre...
therapy 1
present
prev*sensit/(prev*s...
[cost_test+cost_ther_1] \ eff1pres
absent
#
[cost_test+cost_ther_1] \ eff1abs
negative
#
therapy2
present
prev*(1-sensit)/(pre...
[cost_test+cost_ther_2] \ eff2pres
absent
#
[cost_test+cost_ther_2] \ eff2abs

PharmacoEconomics (2014) 32:1141–1145
DOI 10.1007/s40273-014-0195-1

RESEARCH LETTER

The Problem of Embedded Decision Nodes in Cost-Effectiveness Decision Trees

Manuel Arias · Francisco Javier Díez

Published online: 31 July 2014

1 Introduction

Cost-effectiveness analysis (CEA) is increasingly used to inform health policies. Decision trees are the standard method for decision analysis in non-temporal domains. A decision node that is not the root of the tree is said to be embedded.

All books on medical decision analysis discuss both CEA *and* decision trees [1–11], but few explain how to conduct a CEA *with* decision trees [1, 2, 10, 11], and only build a decision tree with one decision node at the root, which represents all the strategies to be evaluated, as proposed by Kuntz and Weinstein; the other is to apply the algorithm presented in Arias and Díez [13].

As a case study, we consider the common problem of finding the incremental cost-effectiveness ratio (ICER) of a test:

Example 1 For a disease with a prevalence of 0.14, there are two possible therapies, the effectiveness of which depends on whether or not the disease is present, as shown

Technical Report

M. Arias and F. J. Díez. **Cost-effectiveness analysis with sequential decisions.** Technical Report CISIAD-11-01, UNED, Madrid, 2011.

26 pages. PDF (859 KB), zip version (827 KB), BibTeX entry.

Abstract

In this paper we present a new method for performing cost-efectiveness analysis of problems that involve multiple decisions and probabilistic outcomes. This problem has been ignored by most of the literature on medical decision making, and the few solutions proposed so far are either wrong or unfeasible except for very small problems. The method proposed in this paper consists of building a decision tree with several decision nodes and evaluating it with a modified roll-back algorithm that operates with partitions of intervals.

Decision trees

See the technical report for an explanation of these examples.

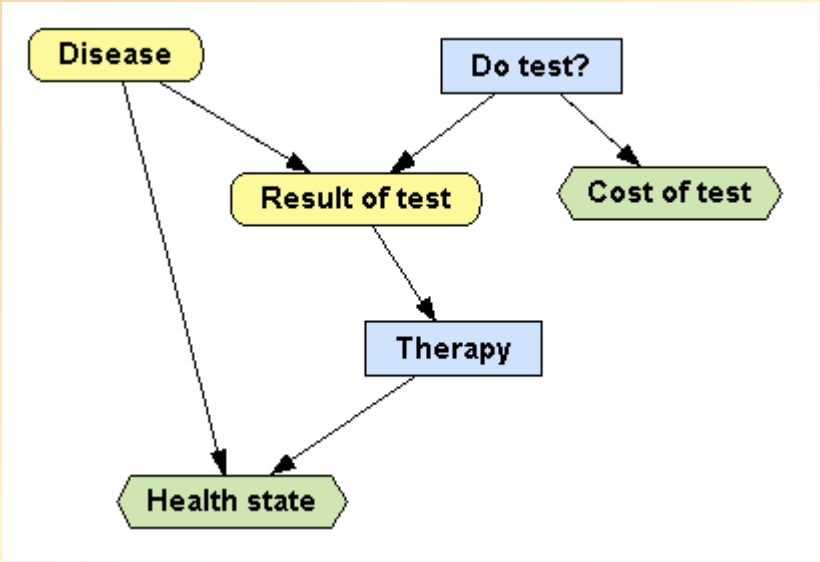
- natural tree (WinDM)
- natural tree (TreeAge Pro)
- all-strategies tree (TreeAge Pro)

Additional information

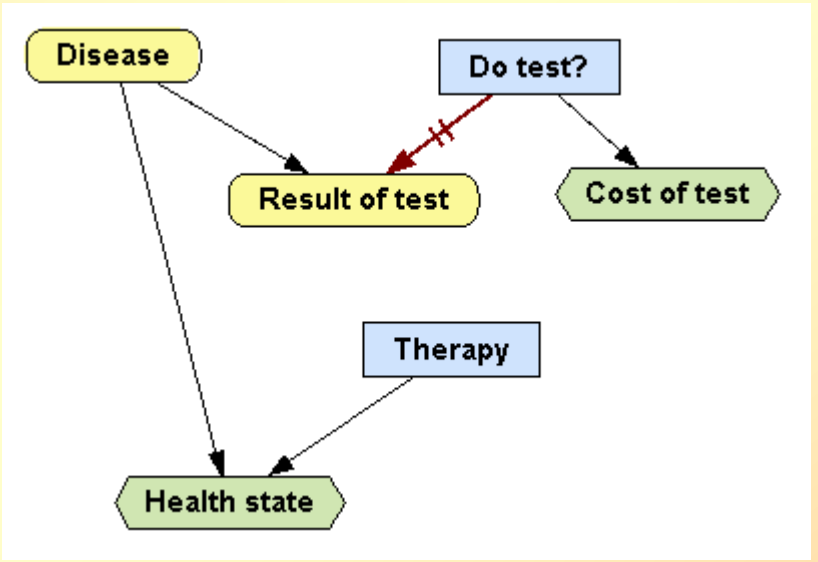
- Slides presented at SMDM-2007.
- Cost-effectiveness analysis in OpenMarkov.

5.3.3. CEA with IDs and DANs

Influence diagram



DAN



- The same structure as in the unicriterion case
- but now we have <u>two criteria</u>: cost and effectiveness

Methods of Information in Medicine 2015;54:353-358.

Cost-effectiveness Analysis with Influence Diagrams*

M. Arias; F. J. Díez

Department of Artificial Intelligence, UNED, Madrid, Spain

Keywords

Cost-benefit analysis, cost-effectiveness analysis, decision trees, influence diagrams

Summary

Background: Cost-effectiveness analysis (CEA) is used increasingly in medicine to determine whether the health benefit of an intervention is worth the economic cost. Decision trees, the standard decision modeling technique for non-temporal domains, can only perform CEA for very small problems.

Objective: To develop a method for CEA in problems involving several dozen variables.

Methods: We explain how to build influence diagrams (IDs) that explicitly represent cost and effectiveness. We propose an algorithm for evaluating cost-effectiveness IDs directly, i.e., without expanding an equivalent deci-

Results: The evaluation of an ID returns a set of intervals for the willingness to pay – separated by cost-effectiveness thresholds – and, for each interval, the cost, the effectiveness, and the optimal intervention. The algorithm that evaluates the ID directly is in general much more efficient than the brute-force method, which is in turn more efficient than the expansion of an equivalent decision tree. Using OpenMarkov, an open-source software tool that implements this algorithm, we have been able to perform CEAs on several IDs whose equivalent decision trees contain millions of branches.

Conclusion: IDs can perform CEA on large problems that cannot be analyzed with decision trees.

units divided by cost units; for example, in dollars per death avoided or euros per quality-adjusted life year (QALY) [4]. As the willingness to pay is different for each decision maker, CEA must consider all its possible values. The result of the analysis is usually a set of intervals for λ, each one having an optimal intervention.

When the consequences of the interventions are not deterministic, it is necessary to model the probability of each outcome. Decision trees are the tool used most frequently for this task, especially in medicine [5]. Their main drawback is that their size grows exponentially with the number of variables[b]. In the medical literature, trees usually have 3 or 4 variables and between 6 and 10 leaf nodes. A tree of 5 variables typically contains around 20 leaf nodes,

UNED

SMDM Society for Medical Decision Making — Better Health through Better Decisions.
39TH ANNUAL NORTH AMERICAN MEETING
October 22 - 25, 2017 | Pittsburgh, PA

Cost-effectiveness analysis with decision analysis networks

Manuel Arias Manuel Luque Jorge Pérez-Martín Francisco Javier Díez

Universidad Nacional de Educación a Distancia (UNED), Madrid, Spain

1 Elements of a DAN

- Structural information: graph
  - three types of nodes: chance, decision, value
  - links connecting nodes; usually represent causality
- Quantitative information
  - probabilities: prevalence, sensitivities, specificities…
  - value functions:
    - effectiveness (life years, QALYs…)
    - economic costs (in $, €, £…)

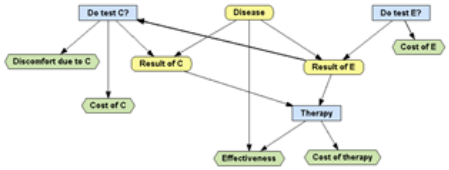
The same as in influence diagrams.

2 Representing the flow of information

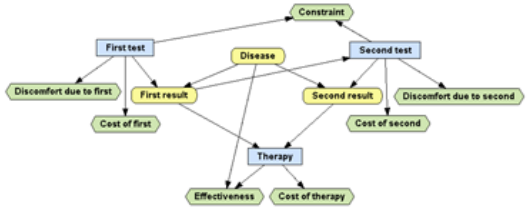
- In influence diagrams (IDs)
  - information links
  - temporal-order links between decisions
  - requisite (by definition): a total ordering of the decisions
- In DANs
  - always observed variables
  - revelation links
  - the decisions may be partially ordered
  - the evaluation algorithm will determine the optimal order

3 Example

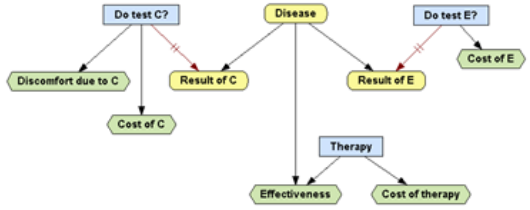
Two IDs with different orderings of the decisions about the tests



An ID that considers both orderings of the tests



A DAN for the same problem



4 Result of evaluating the DAN

- The optimal policy depends on λ, the willingness to pay: 5 ICER thresholds ⇒ 6 intervals

Hands-on exercise 4

Exercise: Optimal stratety for two tests

Test	sensitivity	specificity	discomfort	cost
A	0.60	0.92	0.0003 QALY	$100
B	0.80	0.91	0.0001 QALY	$200

Disease →	absent	present
therapy	38 QALY	30 QALY
no therapy	40 QALY	20 QALY

cost of therapy = $7,000

- The same probabilities and effectiveness as in exercise 3
- but now we also considering economic costs.
- Question: What is the <u>most beneficial</u> strategy?

6. Temporal models

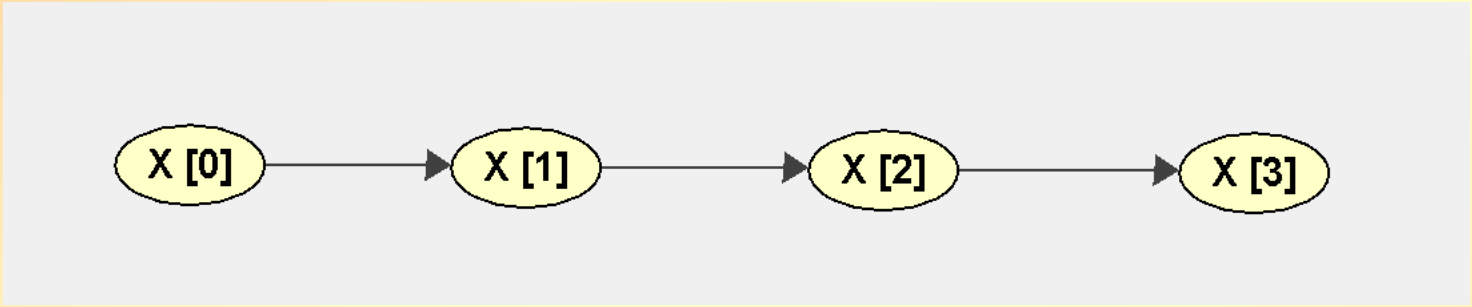
Temporal PGMs

- Markov models
  - The future is independent of the past given the present
    - "Markov models do not have memory"
  - Key concept: state
  - Types of models: Markov chains, HMMs, MDPs, POMDPs, DBNs, MIDs, DLIMIDs…

- Temporal non-Markov models
  - The future is **not** determined by the current state
    - for example, birth occurs around 9 months after conception
  - An type of non-Markov model: event networks
    - Galán, Aguado, Díez, Mira. NasoNet: Modelling the spread of nasopharyngeal cancer with temporal Bayesian networks. *AI in Med*, 2002.

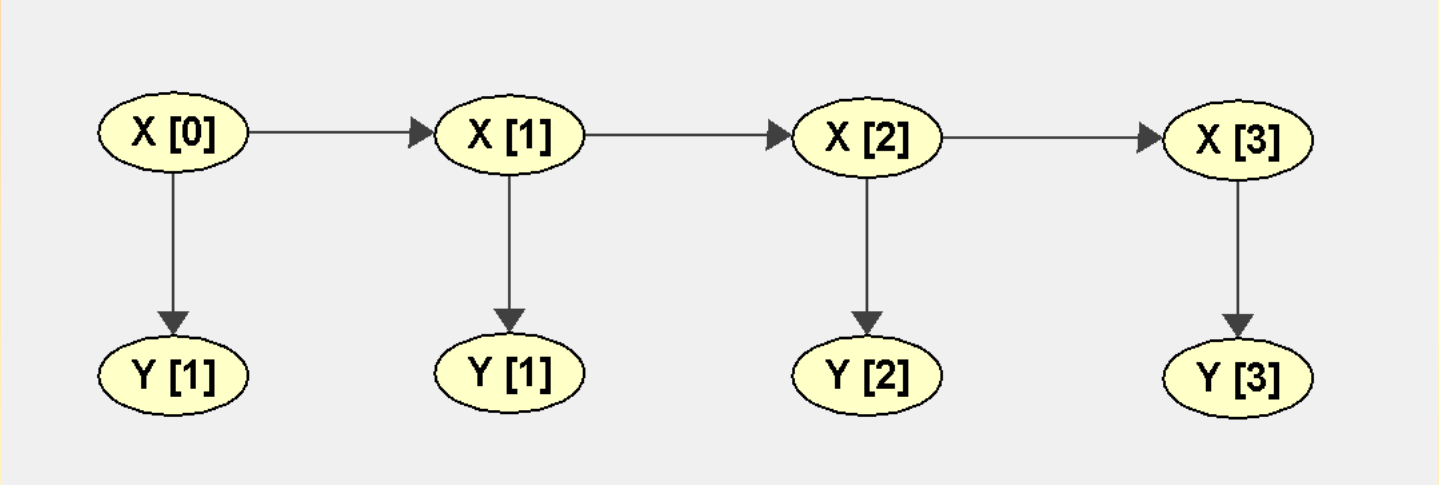
6.1. Types of Markov models

Markov chain



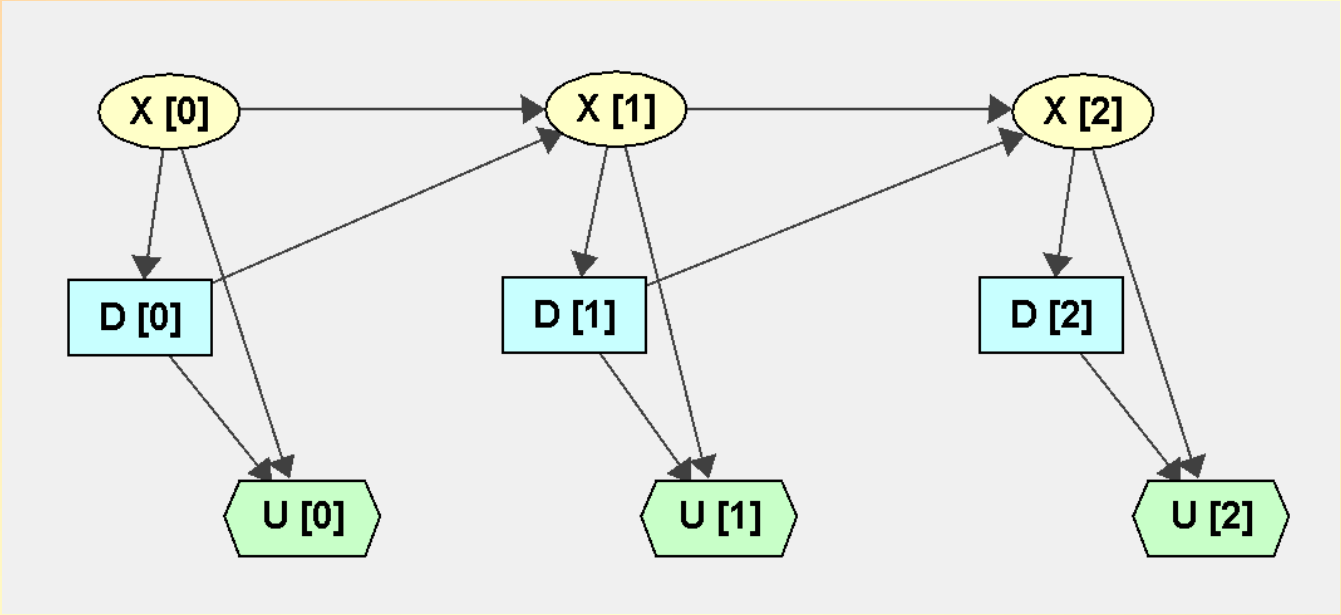
- One variable that evolves over time
- Transition probabilities: $P(x_{i+1}|x_i)$

Hidden Markov model (HMM)



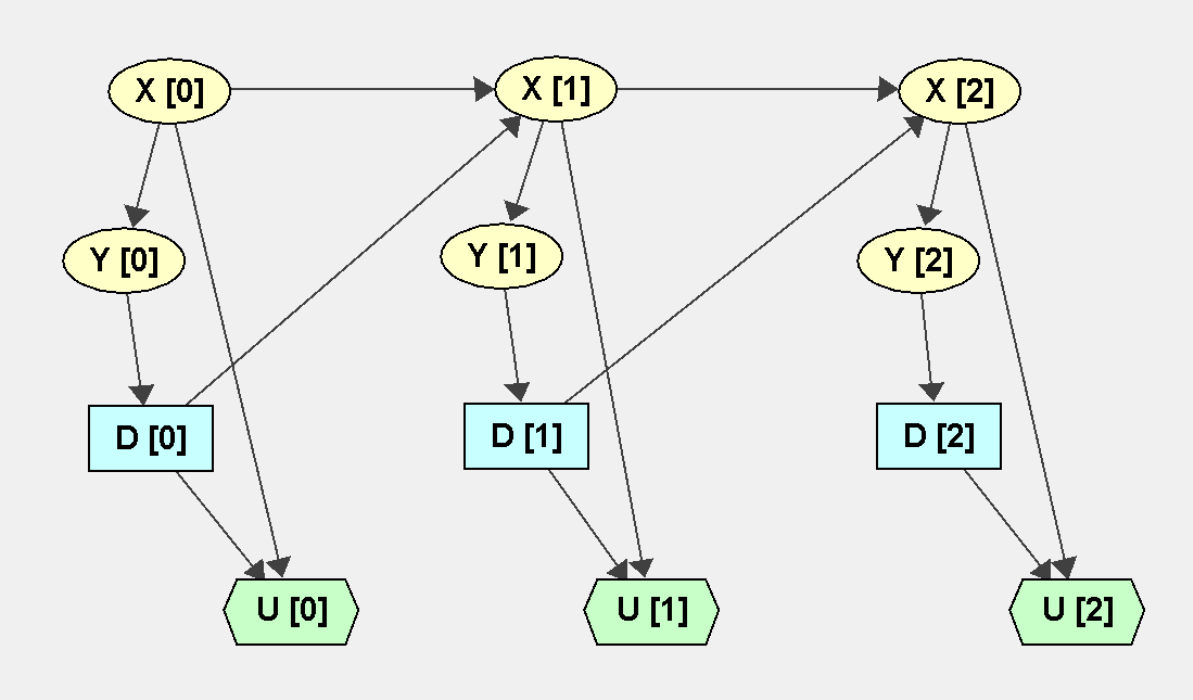
- Observed variable: Y
- Non-observed (hidden) variable: X
- Transition probabilities: $P(x_{i+1}|x_i)$
- Probability of each observation: $P(y_i|x_i)$

Markov decision process (MDP)



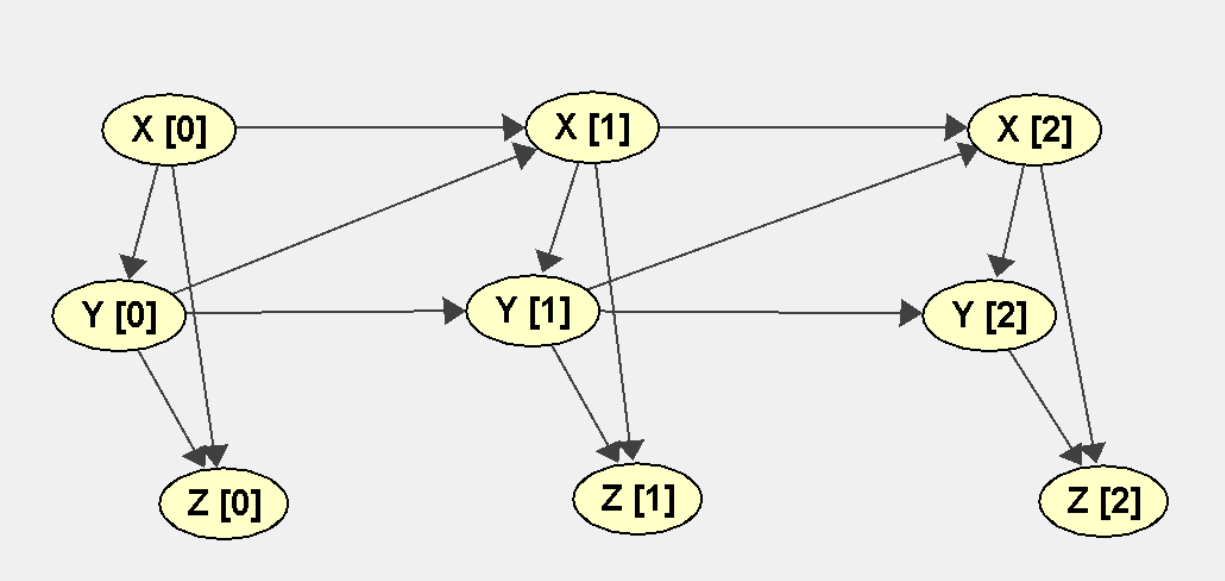
- Observed variable: X
- Decision: D
- Transition probabilities: $P(x_{i+1}|x_i)$
- Reward: $U(x_i, d_i)$

Partially observable MDP (POMDP)



- Hidden variable: X
- Observed variable : Y
- Decision: D
- Observation prob.: $P(y_i|x_i)$
- Transition prob.: $P(x_{i+1}|x_i)$
- Reward: $U(x_i, d_i)$

Dynamic Bayesian network (DBN)

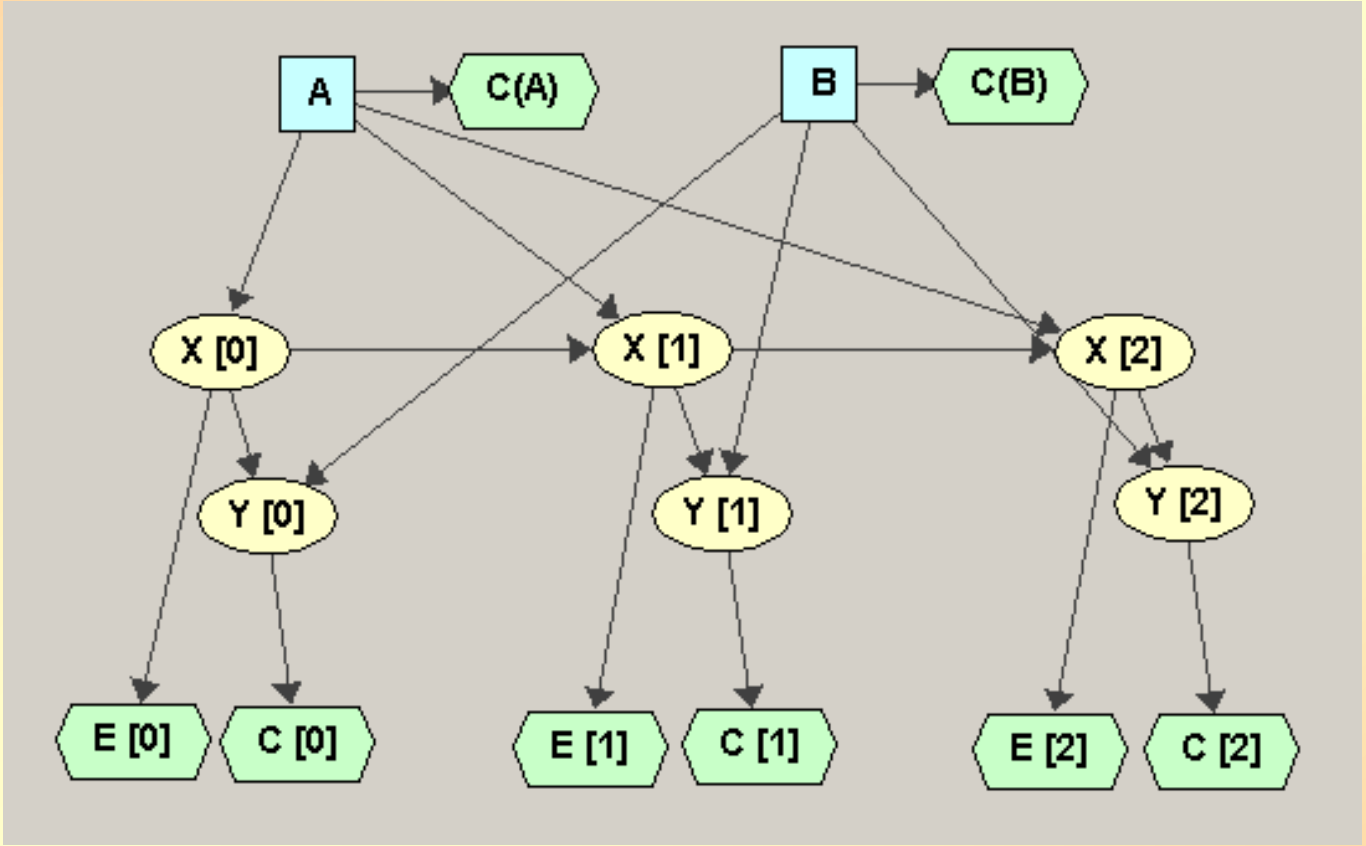


- Markov chain or hidden Markov model:
  - one variable, X
  - one conditional probability: $P(x_{i+1}|x_i)$
- Dynamic Bayesian network:
  - several variables, $\{X, Y, Z\ldots\}$
  - factored probability: $P(y_i|x_i)$, $P(z_i|x_i, y_i)$, $P(x_{i+1}|x_i, y_i)\ldots$

Factored extensions of Markov models

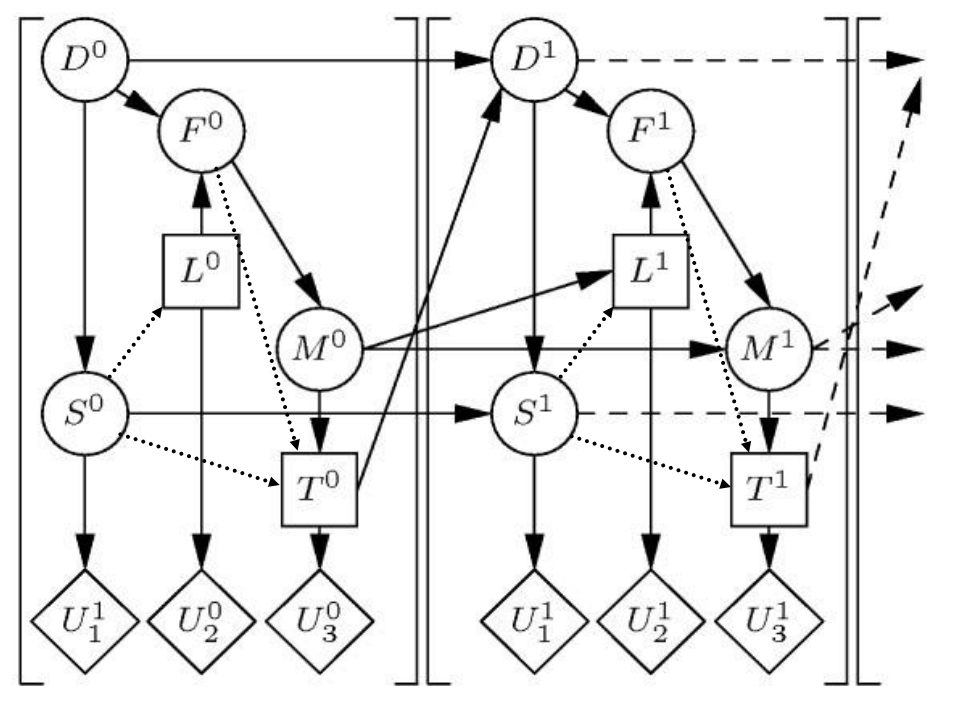
Flat model	Factored model
Markov chain	Dynamic Bayesian network [Dean and Kanazawa, 1989]
Hidden Markov model	
Markov decision process (MDP)	Factored MDP [Boutilier et al., 1995, 2000]
Partially-observable MDP (POMDP)	Factored POMDP [Boutilier and Poole, 1996]

Markov influence diagrams



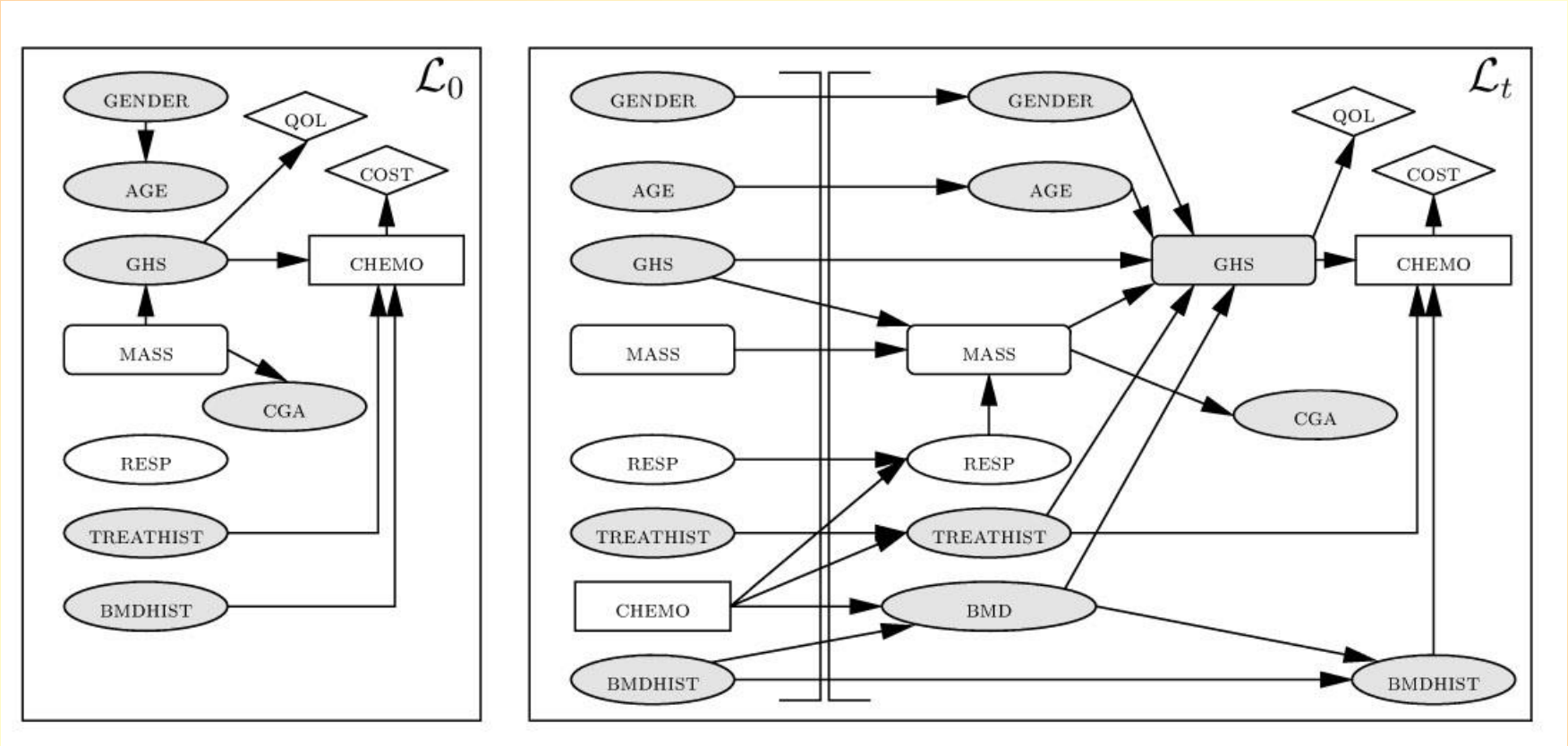
- Can be used for cost-effectiveness analysis

Dynamic limited-memory IDs (DLIMIDs)



- Differences wrt POMDPs
  - Several decisions in each time slice.
  - Limited memory: the decision maker only knows the observations made at the current and the previous time slices
  - Memory variables summarize the past.

A DLIMID for a carcinoid tumors



- Therapy selection for high-grade carcinoid tumors (van Gerven et al., 2007)

IJCAI Workshop *Decision Making in Partially Observable, Uncertain Worlds: Exploring Insights from Multiple Communities*
Barcelona, July 2011

MDPs in Medicine: Opportunities and Challenges

F. J. Díez M. A. Palacios M. Arias
Dept. Artificial Intelligence. UNED
Madrid, Spain

Abstract

In the last three decades hundreds of Markov models have been built for medical applications, but most of them fall under the paradigm of what we call *simple Markov models* (SMMs). Markov decision processes (MDPs) are much more powerful as a decision analysis tool, but they are ignored in medical decision analysis books and the number of medical applications based on them is still very small. In this paper we compare both types of models and discuss the challenges that MDPs must overcome before they can be widely accepted in medicine. We present a software tool, OpenMarkov, that addresses those challenges and has been used to build a Markov model for analyzing the cost-effectiveness of the HPV vaccine.

1 Introduction

Markov models were introduced in the beginning of the 20-th century by the Russian mathematician Andrei Andreyevich Markov [1906]. In the three decades passed since the pioneering work of Beck and Pauker [1983], hundreds of

the emergence of partially observable Markov decision processes (POMDPs) [Åström, 1965], in which the state of the system is not directly observable, but there is a variable that correlates probabilistically with it. POMDPs were developed in the field of automatic control as an extension of MDPs, but currently most of the research about them is carried out in artificial intelligence (AI), again as a tool for planning, especially in robotics [Ghallab *et al.*, 2004]. The main contribution of AI to this field comes from the area of probabilistic graphical models: Bayesian networks [Pearl, 1988] led to the development of dynamic Bayesian networks [Dean and Kanazawa, 1989], which generalize Markov chains and hidden Markov models [Murphy, 2002]. The idea of using several variables to represent the state of the system, instead of only one, led to factored MDPs [Boutilier *et al.*, 1995; 2000] and factored POMDPs [Boutilier and Poole, 1996], which can model efficiently many problems that were unmanageable with flat (i.e., non-factored) representations; correspondingly, there are new algorithms that can solve problems several orders of magnitude bigger than in the recent past [Hoey *et al.*, 1999; Poupart, 2005; Spaan and Vlassis, 2005].

In the rest of the paper, we use the acronym MDPs to refer to both fully observable and partially observable models (FOMDPs and POMDPs, respectively).

6.2. Markov influence diagrams

Medical Decision Making 2017; 37:183-195

ORIGINAL ARTICLE

Markov Influence Diagrams: A Graphical Tool for Cost-Effectiveness Analysis

Francisco J. Díez, PhD, Mar Yebra, MEng, Iñigo Bermejo, PhD, Miguel A. Palacios-Alonso, MSc, Manuel Arias Calleja, PhD, Manuel Luque, PhD, Jorge Pérez-Martín, MEng

Markov influence diagrams (MIDs) are a new type of probabilistic graphical model that extends influence diagrams in the same way that Markov decision trees extend decision trees. They have been designed to build state-transition models, mainly in medicine, and perform cost-effectiveness analyses. Using a causal graph that may contain several variables per cycle, MIDs can model various patient characteristics without multiplying the number of states; in particular, they can represent the history of the patient without using tunnel states. OpenMarkov, an open-source tool, allows the decision analyst to build and evaluate MIDs—including cost-effectiveness analysis and several types of deterministic and probabilistic sensitivity analysis—with a graphical user interface, without writing any code. This way, MIDs can be used to easily build and evaluate complex models whose implementation as spreadsheets or decision trees would be cumbersome or unfeasible in practice. Furthermore, many problems that previously required discrete event simulation can be solved with MIDs; i.e., within the paradigm of state-transition models, in which many health economists feel more comfortable. ***Key words:*** *Markov models; influence diagrams; cost-effectiveness analysis; outcomes research.* ***(Med Decis Making XXXX; XX:xx–xx)***

6.2.1. Example: Chancellor's model for HIV

Case study: HIV/AIDS
(Chancellor et al.,1997)

ORIGINAL RESEARCH ARTICLE

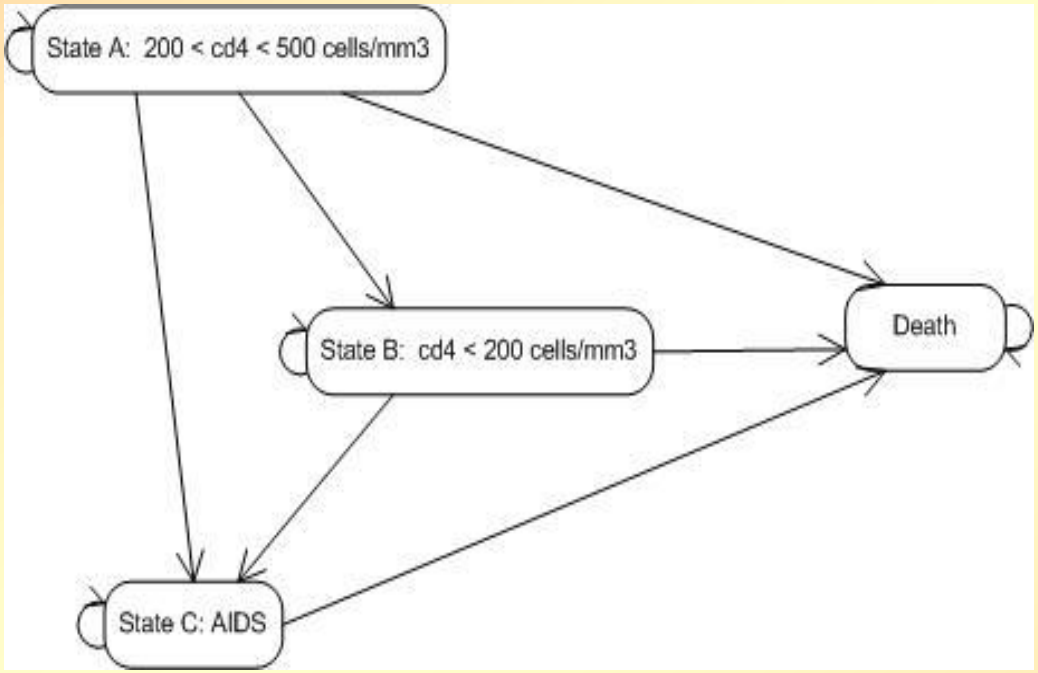
Pharmacoeconomics 1997 Jul 12 (1) 54-66
1170-7690/97/0007-0054/$06.50/0

Modelling the Cost Effectiveness of Lamivudine/Zidovudine Combination Therapy in HIV Infection

Jeremy V. Chancellor,[1] *Andrew M. Hill,*[2] *Caroline A. Sabin,*[3] *Kit N. Simpson*[4] and *Mike Youle*[5]

1 Glaxo Wellcome UK Ltd, Uxbridge, Middlesex, England
2 Glaxo Wellcome Research and Development Ltd, Greenford, Middlesex, England
3 Department of Primary Care and Population Sciences, Royal Free Hospital, London, England
4 University of North Carolina, Chapel Hill, North Carolina, USA
5 HIV/GUM Research Unit, Chelsea and Westminster Hospital, London, England

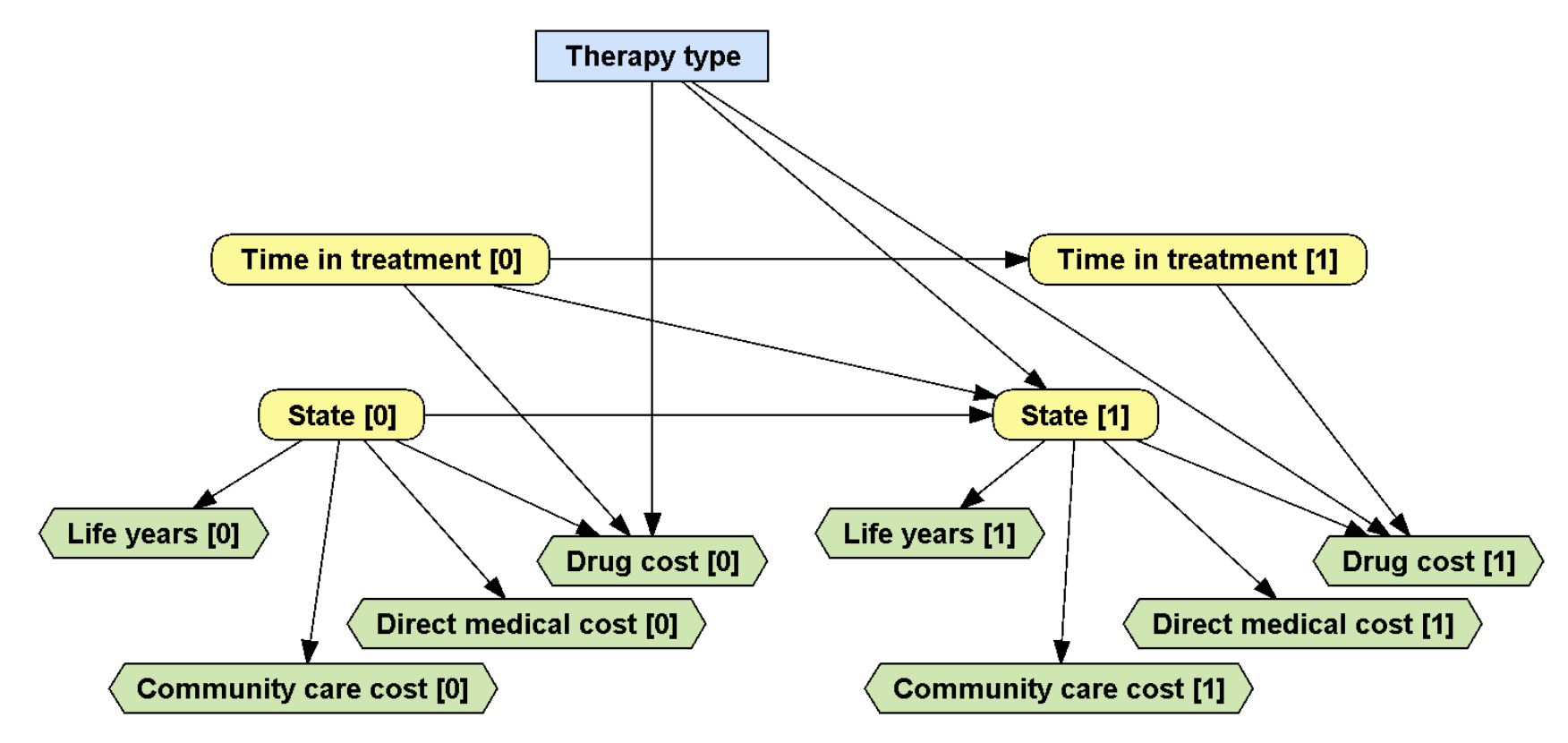
- State-transition diagram: 4 states



- Two therapies:
  - monotherapy: only AZT
  - combined therapy: AZT + lamivudine for 2 years; then only AZT

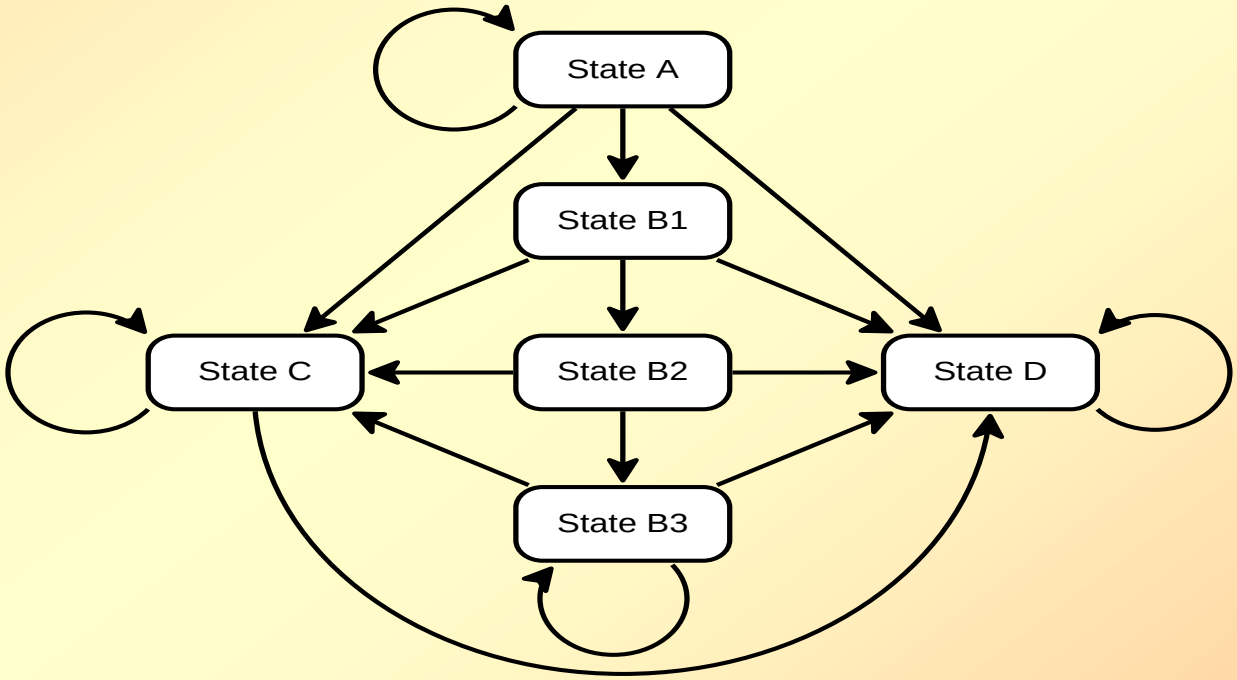
A MID version of the HIV model

[Chancellor et al., 1997]



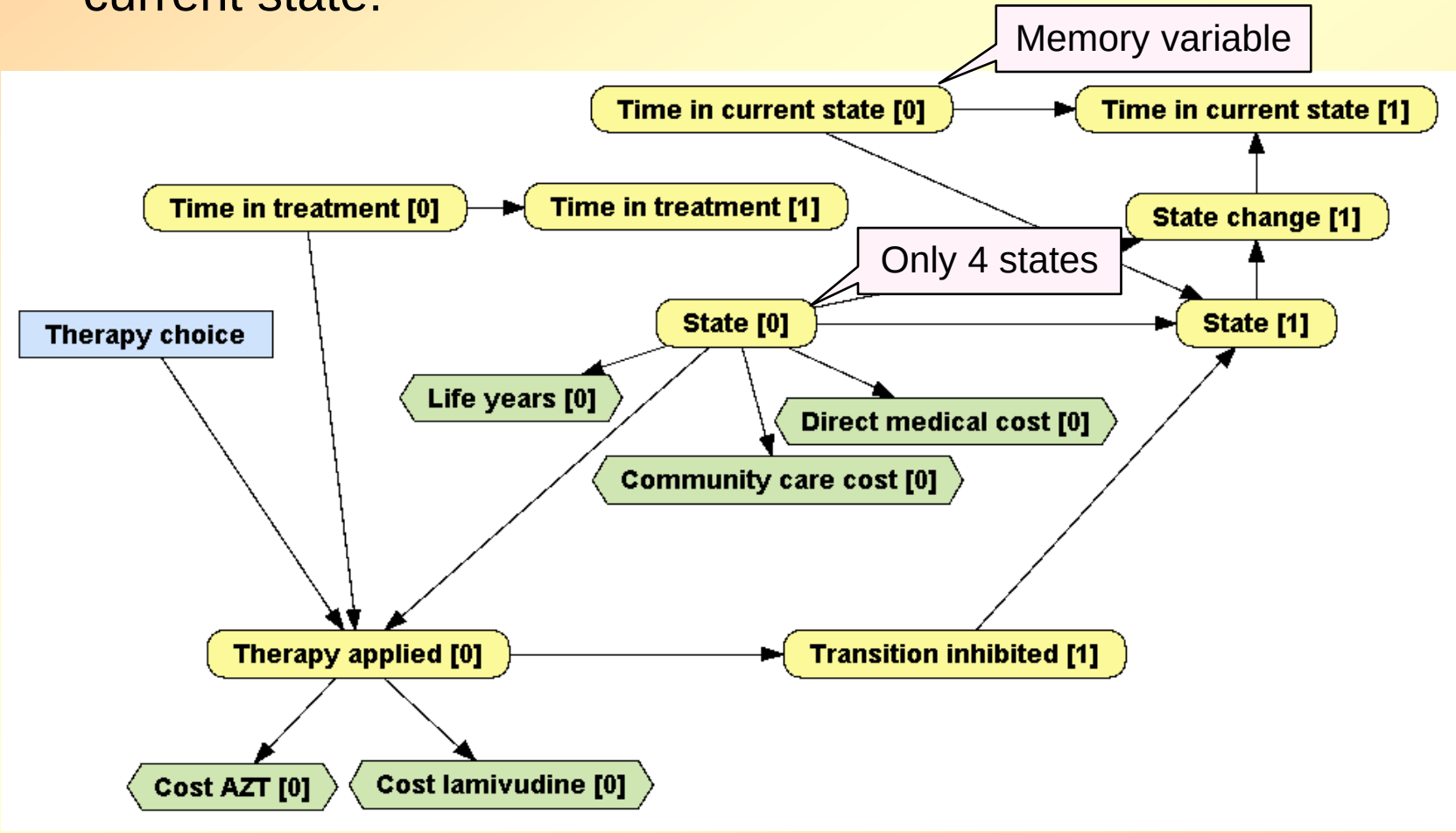
Representing the patient history (1)

- Transition probabilities that depend on the time spent in current state:
  - State-transition model with tunnel states



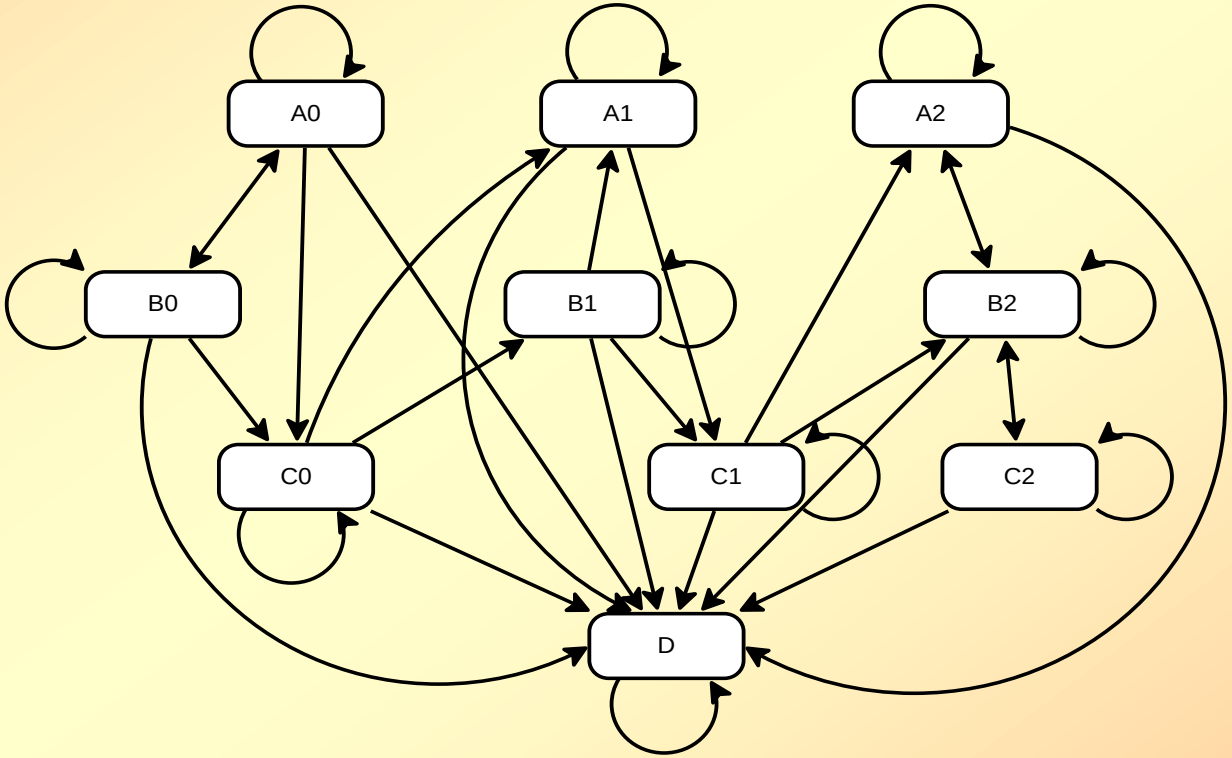
Representing the patient history (1)

- Transition probabilities that depend on the time spent in current state:



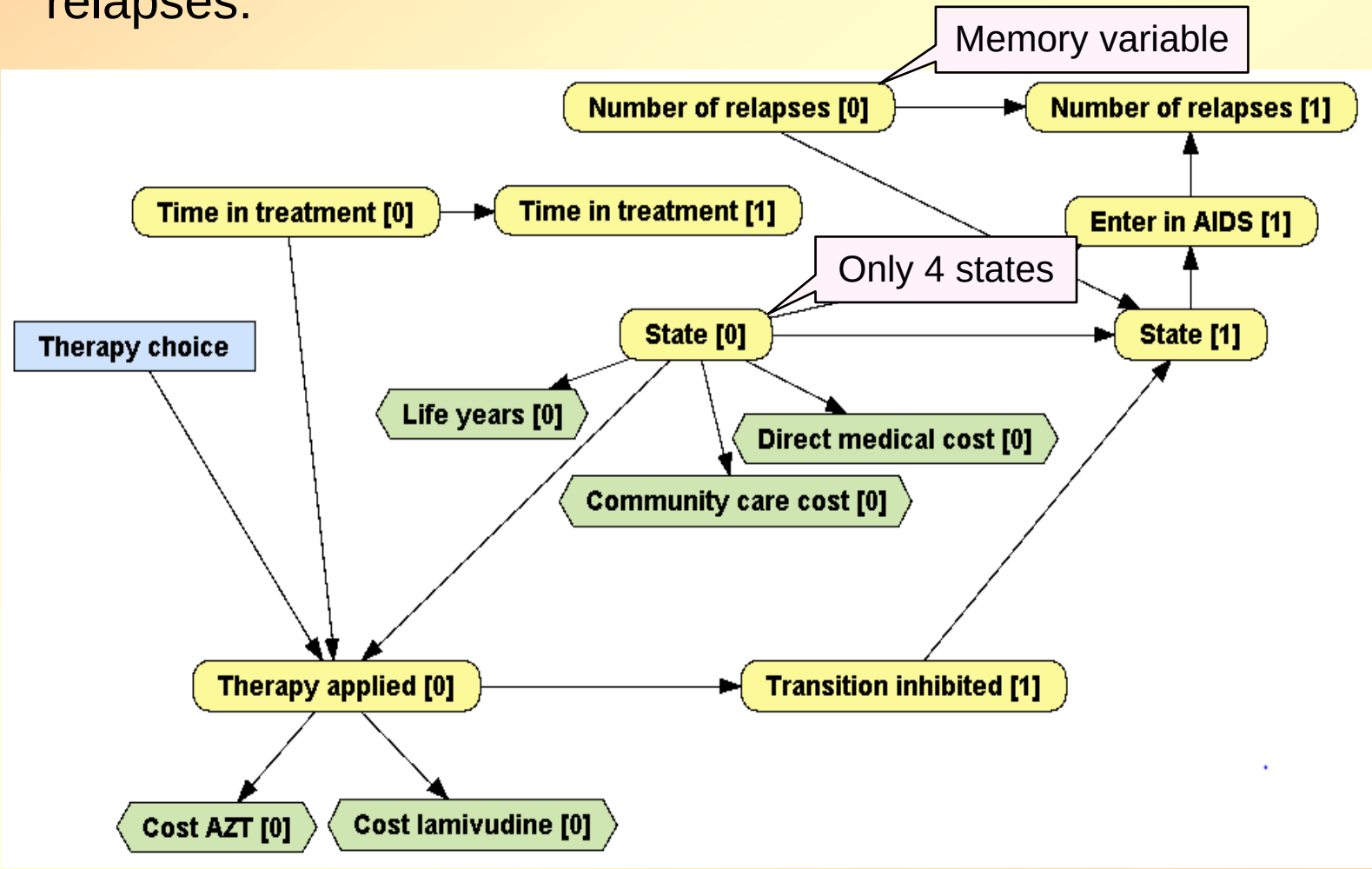
Representing the patient history (2)

- Transition probabilities that depend on the number of relapses:



Representing the patient history (2)

- Transition probabilities that depend on the number of relapses:



6.2.2. Other MIDs for real-world problems

Case study: Hip replacement
(Briggs et al., 2004)

ARTICLE

Appl Health Econ Health Policy 2004; 3 (2): 79-89
1175-5652/04/0002-0079/$31.00/0

The Use of Probabilistic Decision Models in Technology Assessment

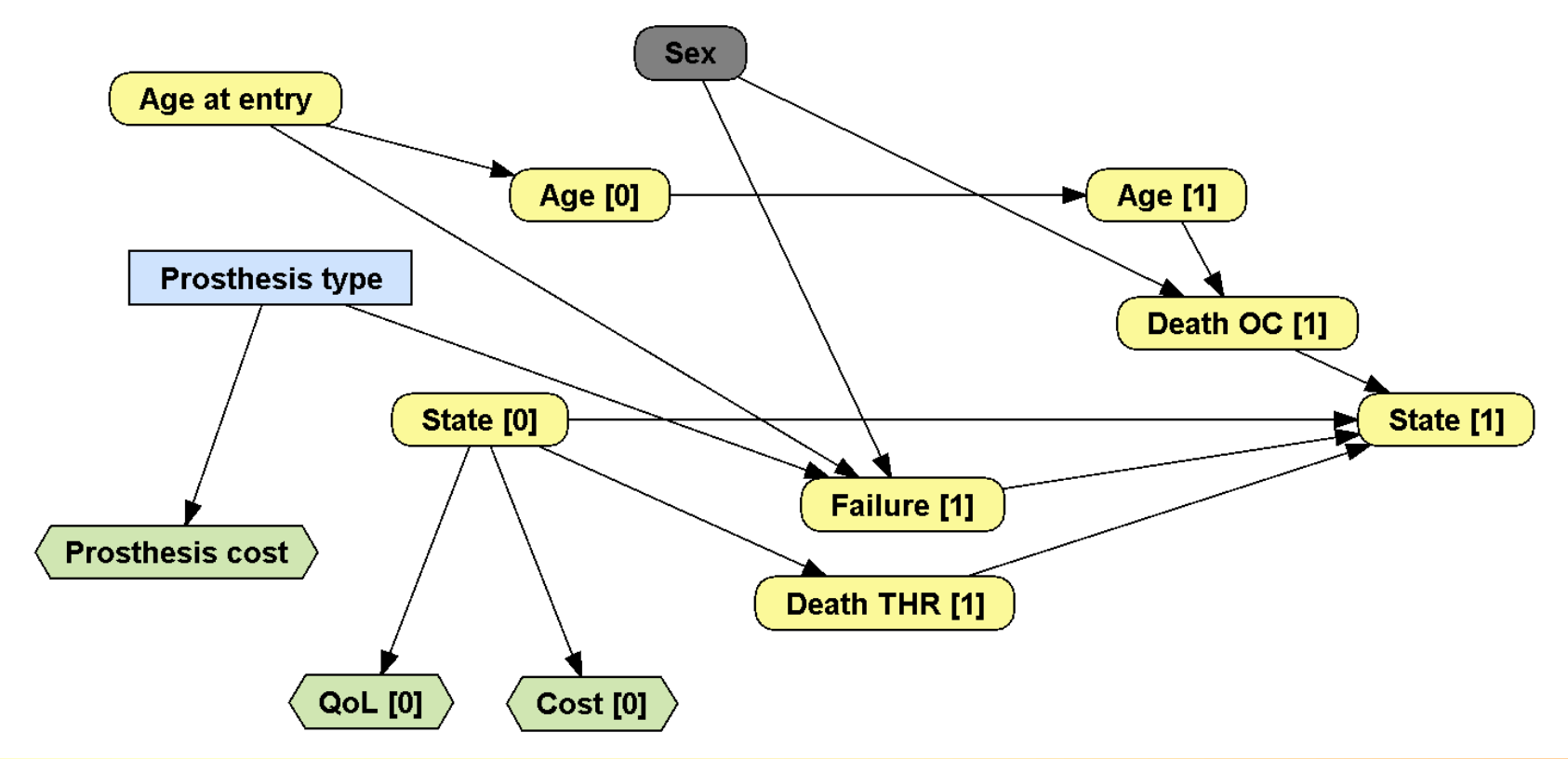
The Case of Total Hip Replacement

Andrew Briggs,[1] *Mark Sculpher*,[2] *Jill Dawson*,[3] *Ray Fitzpatrick*,[4] *David Murray*[5] and *Henrik Malchau*[6]

1 Health Economics Research Centre, Department of Public Health, University of Oxford, Old Road Campus, Headington, Oxford, UK
2 Centre for Health Economics, University of York, Heslington, York, UK
3 School of Health and Social Care, Oxford Brookes University, Oxford, UK
4 Department of Public Health, University of Oxford, Old Road Campus, Headington, Oxford, UK
5 Nuffield Orthopaedic Centre, Headington, Oxford, UK
6 Department of Orthopaedics, Massachusetts General Hospital, Boston, USA

A MID version of the hip replacement model

[Briggs et al., 2004]



Case study: HPV vaccine
(Insinga et al., 2009)

BMC Infectious Diseases

BioMed Central

Research article

Open Access

Epidemiologic natural history and clinical management of Human Papillomavirus (HPV) Disease: *a critical and systematic review of the literature in the development of an HPV dynamic transmission model*

Ralph P Insinga*, Erik J Dasbach and Elamin H Elbasha

Address: Department of Health Economic Statistics, Merck Research Laboratories, North Wales, PA, USA

Email: Ralph P Insinga* - ralph_insinga@merck.com; Erik J Dasbach - erik_dasbach@merck.com; Elamin H Elbasha - elamin_elbasha@merck.com

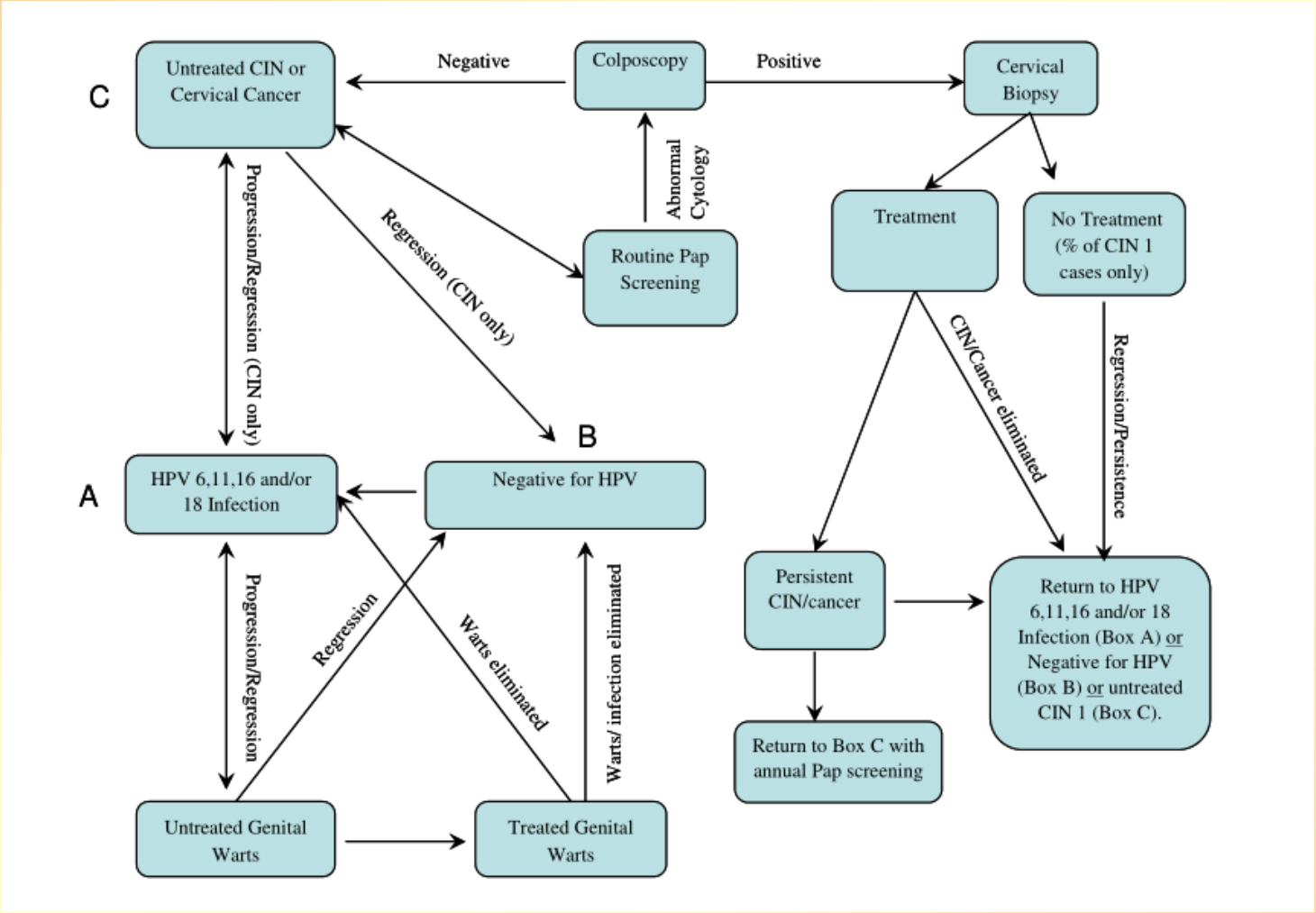
* Corresponding author

Published: 29 July 2009

BMC Infectious Diseases 2009, **9**:119 doi:10.1186/1471-2334-9-119

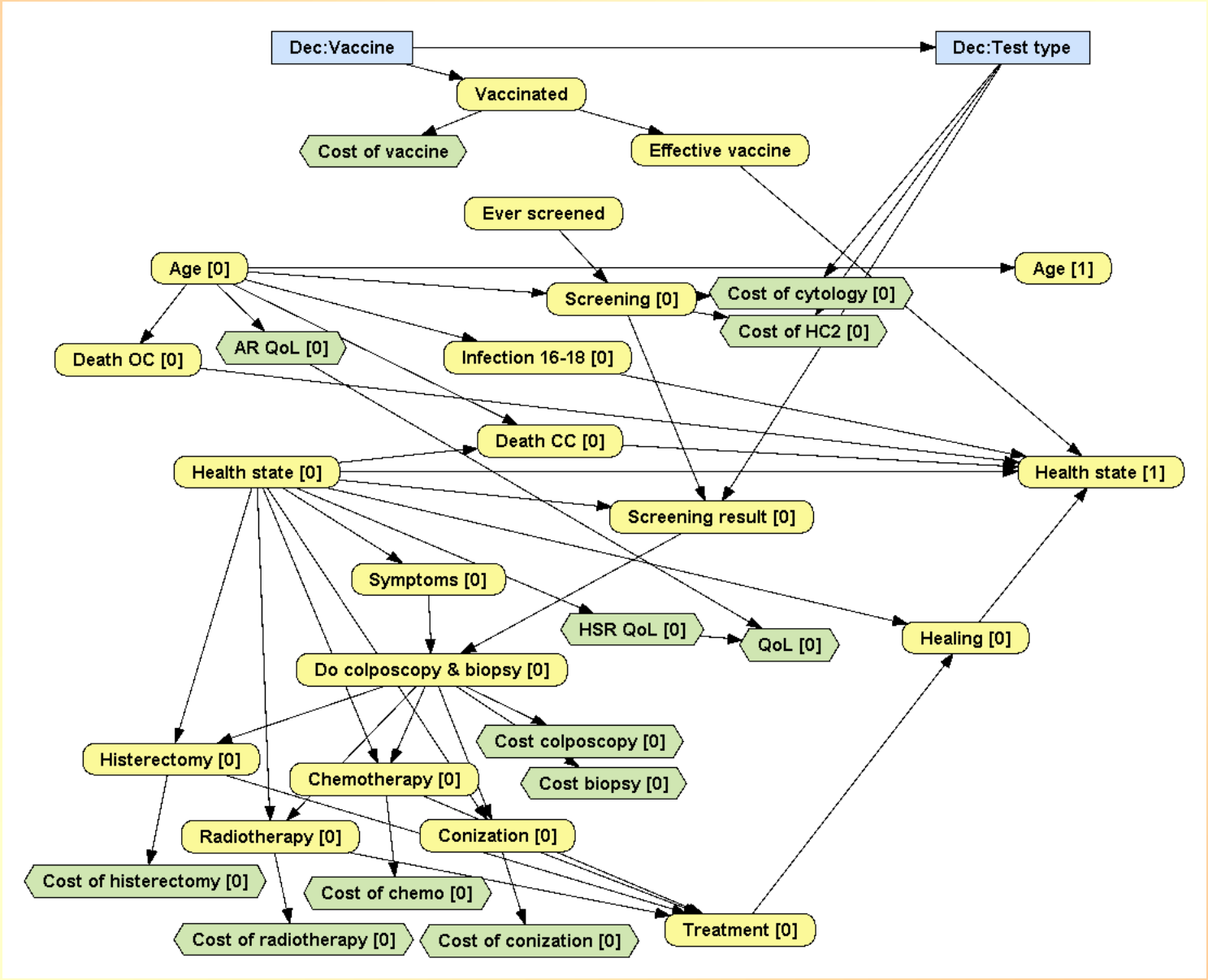
Received: 13 May 2008
Accepted: 29 July 2009

This article is available from: http://www.biomedcentral.com/1471-2334/9/119


C
Untreated CIN or Cervical Cancer
Negative
Colposcopy
Positive
Cervical Biopsy
Abnormal Cytology
Progression/Regression (CIN only)
Regression (CIN only)
Routine Pap Screening
Treatment
No Treatment (% of CIN 1 cases only)
B
A
HPV 6,11,16 and/or 18 Infection
Negative for HPV
CIN/Cancer eliminated
Regression/Persistence
Persistent CIN/cancer
Return to HPV 6,11,16 and/or 18 Infection (Box A) or Negative for HPV (Box B) or untreated CIN 1 (Box C).
Progression/Regression
Regression
Warts eliminated
Warts/ infection eliminated
Return to Box C with annual Pap screening
Untreated Genital Warts
Treated Genital Warts

A MID version of the HPV vaccination model

[Callejo et al., 2010]



Content of one of the Excel cells for this model:

=VLOOKUP($C5;Variables!$A$4:$H$21;8;TRUE)*(((BI5+BJ5)+BK5*uCIN1+SUM(BL5:BP5)*uCIN2_3+(BQ5+BR5)*uLCC+(BS5+BT5)*uRCC+(BU5+BV5)*uDCC)+((BI4+BJ4)+BK4*uCIN1+SUM(BL4:BP4)*uCIN2_3+(BQ4+BR4)*uLCC+(BS4+BT4)*uRCC+(BU4+BV4)*uDCC)*VLOOKUP($C4;Variables!$A$4:$H$21;2;TRUE)+(BQ4+BR4)*uLCC*VLOOKUP($C4;Variables!$A$4:$H$21;4;TRUE)+(BS4+BT4)*uRCC*VLOOKUP($C4;Variables!$A$4:$H$21;5;TRUE)+(BU4+BV4)*uDCC*VLOOKUP($C4;Variables!$A$4:$H$21;2;TRUE))

Case study: AIDS in Africa
(Ryan et al., 2009)

The cost-effectiveness of cotrimoxazole prophylaxis in HIV-infected children in Zambia

Máirín Ryan[a], Susan Griffin[b], Bona Chitah[c], A. Sarah Walker[d], Veronica Mulenga[e], Donald Kalolo[e], Neil Hawkins[b], Concepta Merry[a], Michael G. Barry[a], Chifumbe Chintu[e], Mark J. Sculpher[b] and Diana M. Gibb[d]

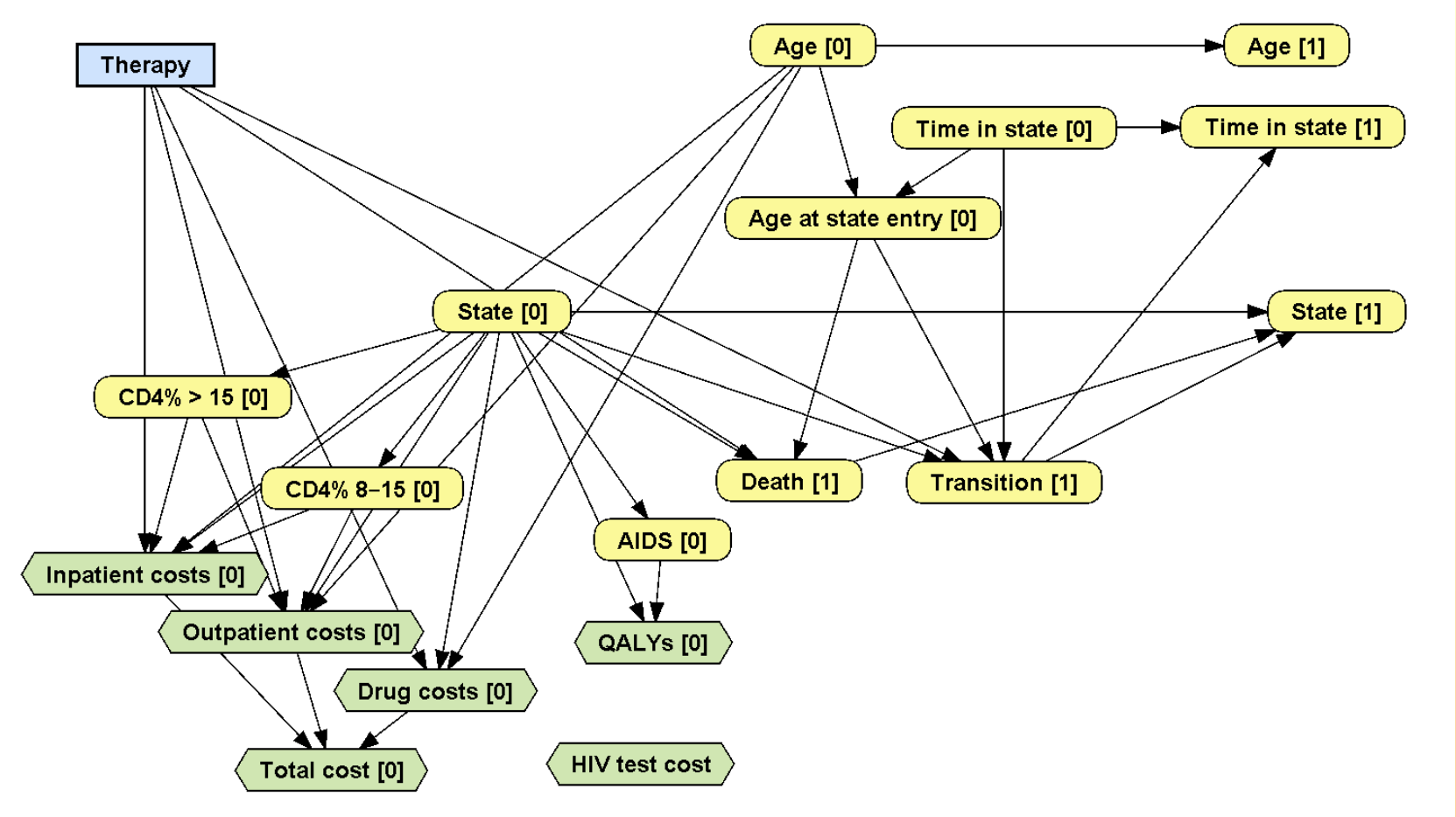
Objective: To assess the cost-effectiveness of cotrimoxazole prophylaxis in HIV-infected children in Zambia, as implementation at the local health centre level has yet to be undertaken in many resource-limited countries despite recommendations in recent updated World Health Organization (WHO) guidelines.

Design: A probabilistic decision analytical model of HIV/AIDS progression in children based on the CD4 cell percentage (CD4%) was populated with data from the placebo-controlled Children with HIV Antibiotic Prophylaxis trial that had reported a 43% reduction in mortality with cotrimoxazole prophylaxis in HIV-infected children aged 1–14 years.

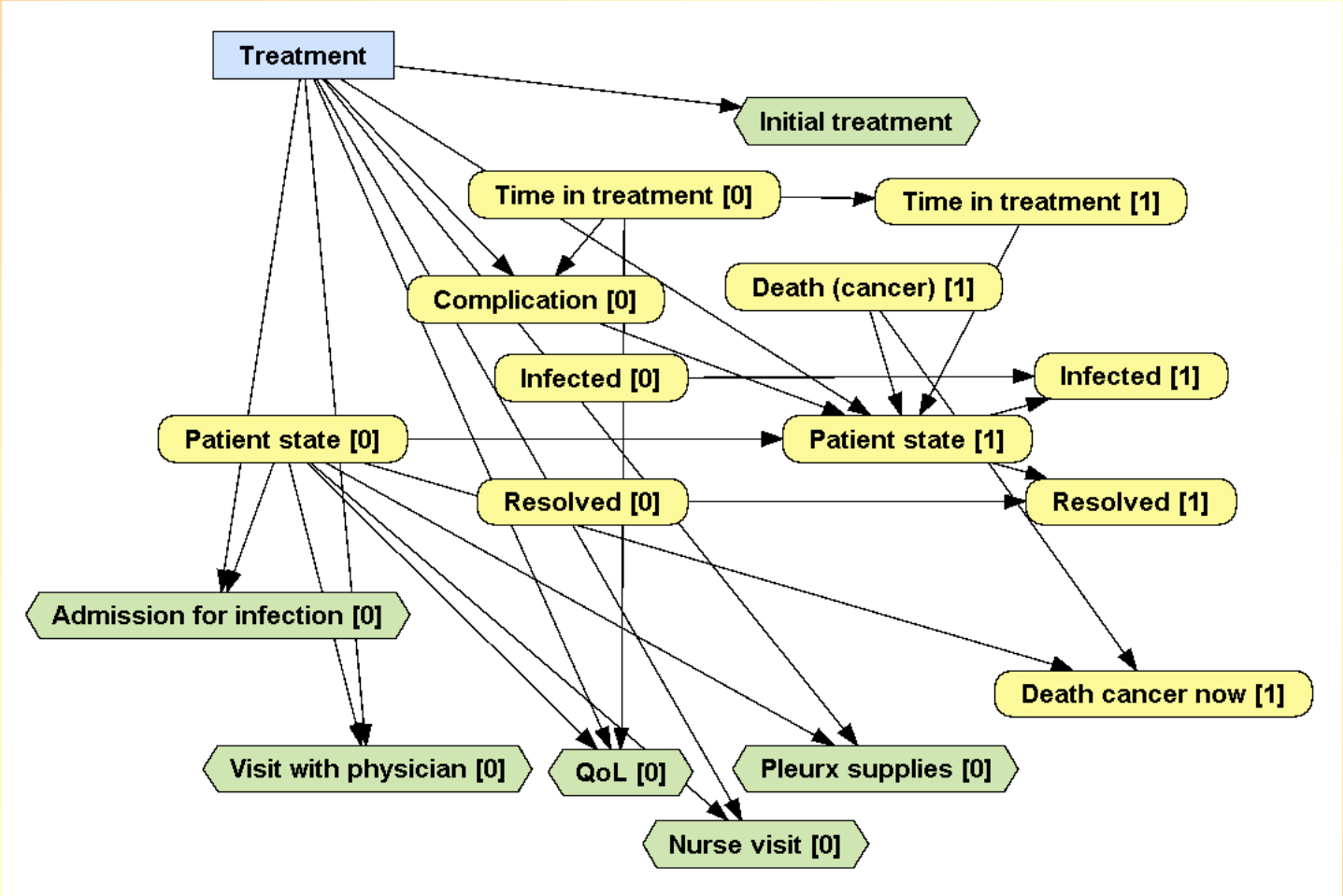
Methods: Unit costs (US$ in 2006) were measured at University Teaching Hospital, Lusaka. Cost-effectiveness expressed as cost per life-year saved, cost per quality adjusted life-year (QALY) saved, cost per disability adjusted life-year (DALY) averted was calculated across a number of different scenarios at tertiary and primary healthcare centres

A MID version of the CHAP model

[Ryan et al., 2008]

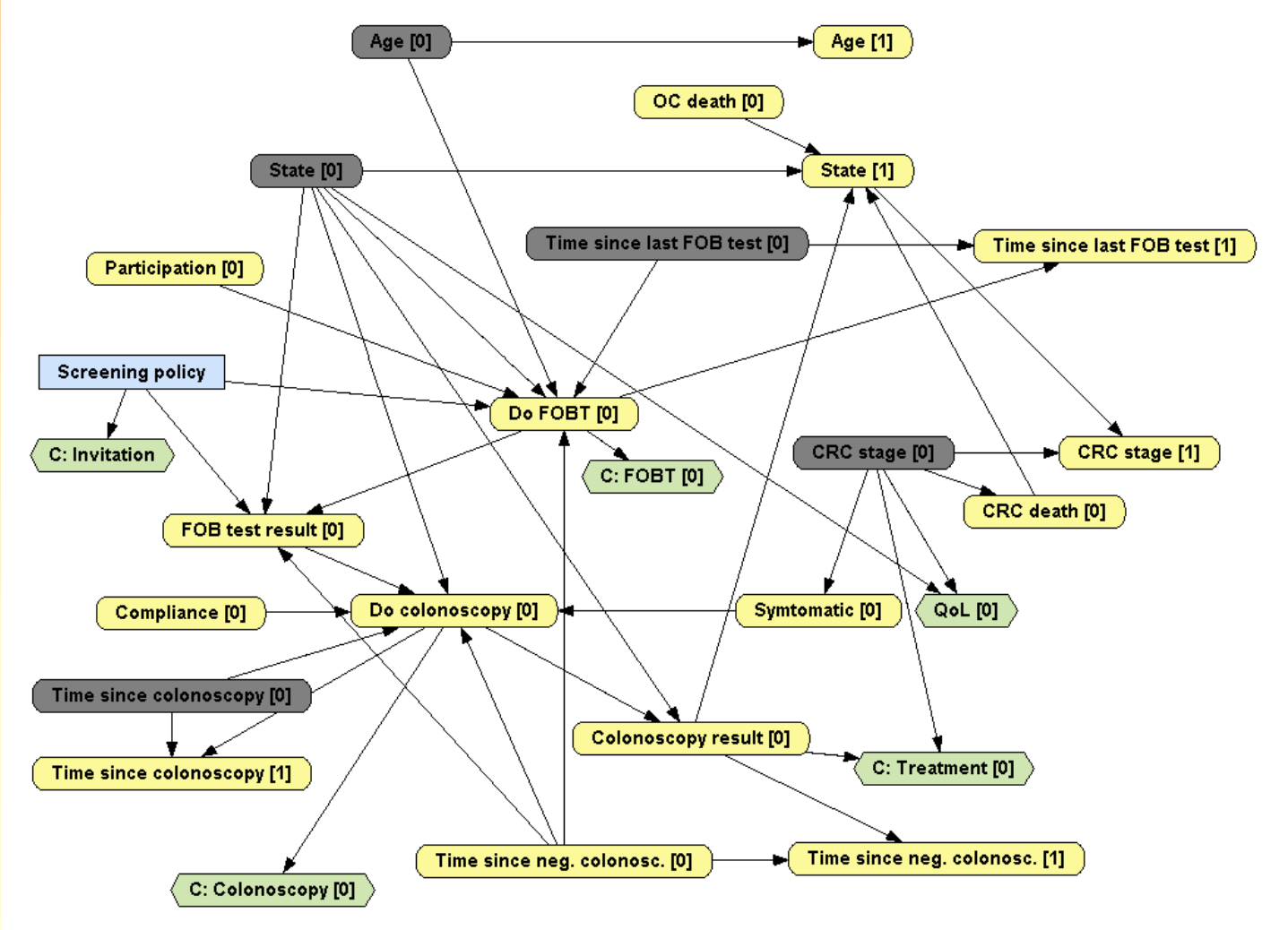


Our model for malignant pleural effusion



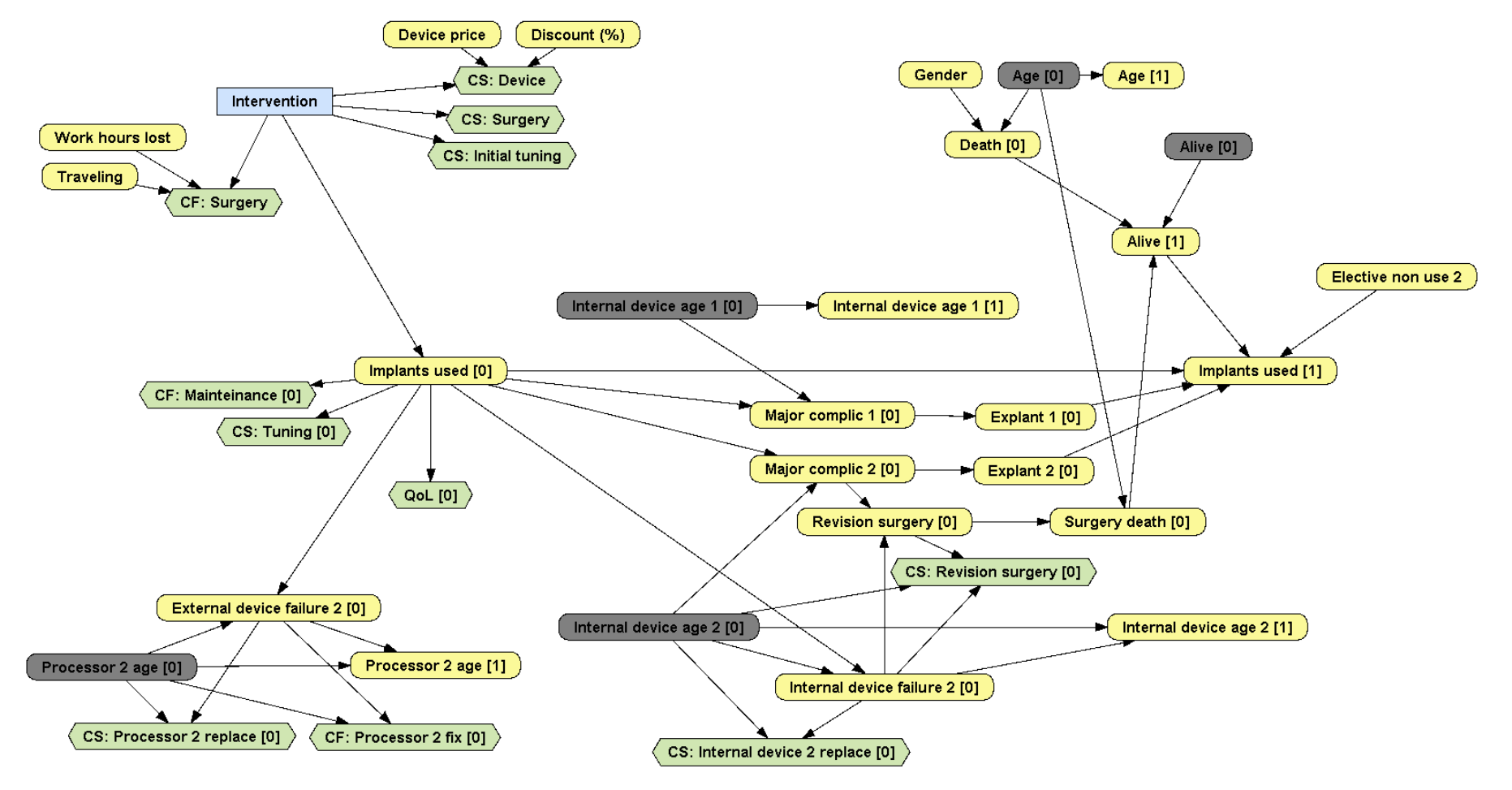
- *Meeting of the Society for Medical Decision Making (SMDM 2015)*, St. Louis, October 2015.

Our model for colorectal cancer screening



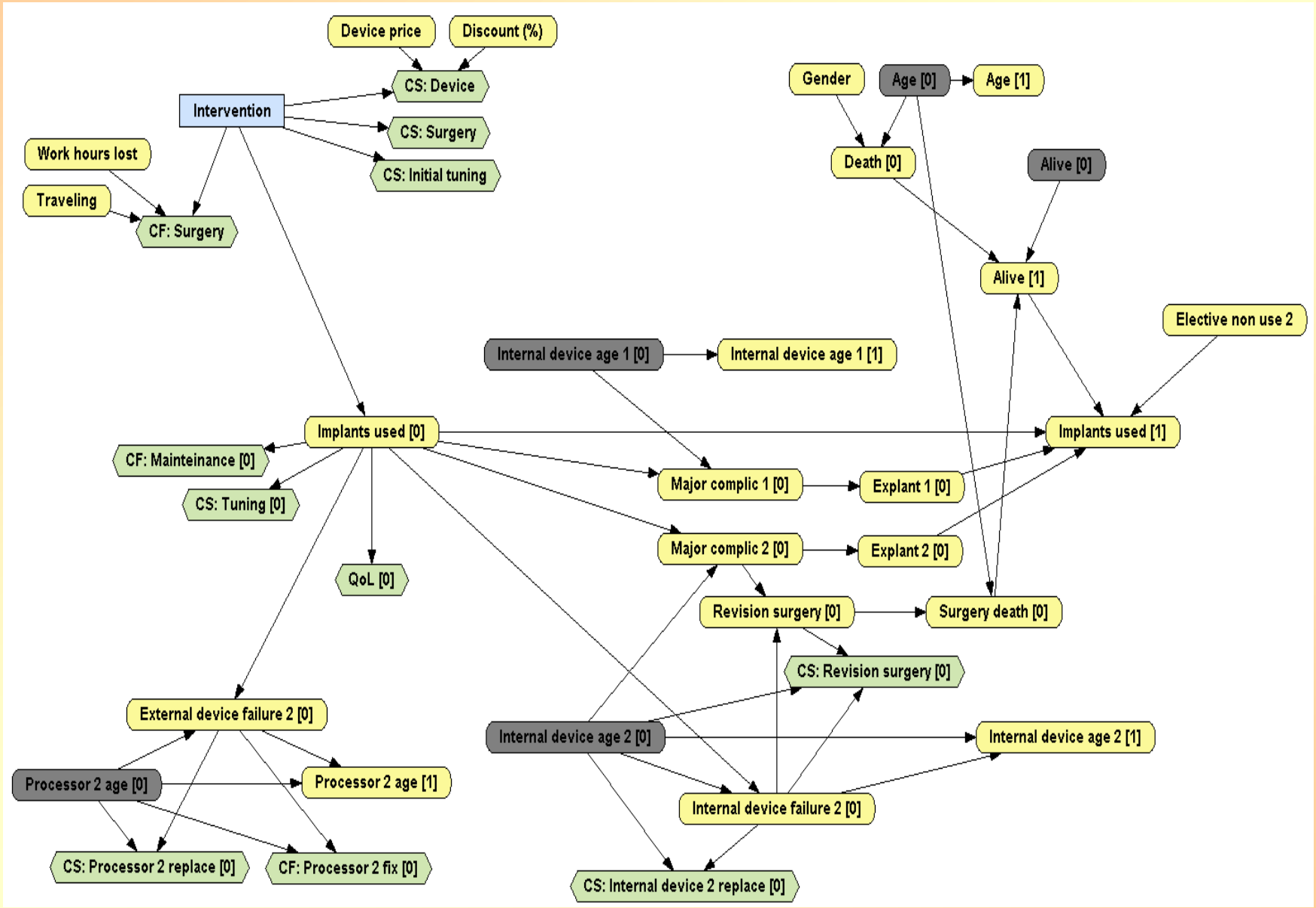
- *European Conference of the Society for Medical Decision Making*, London, UK, June 2015.

Our model for bilateral cochlear implantation

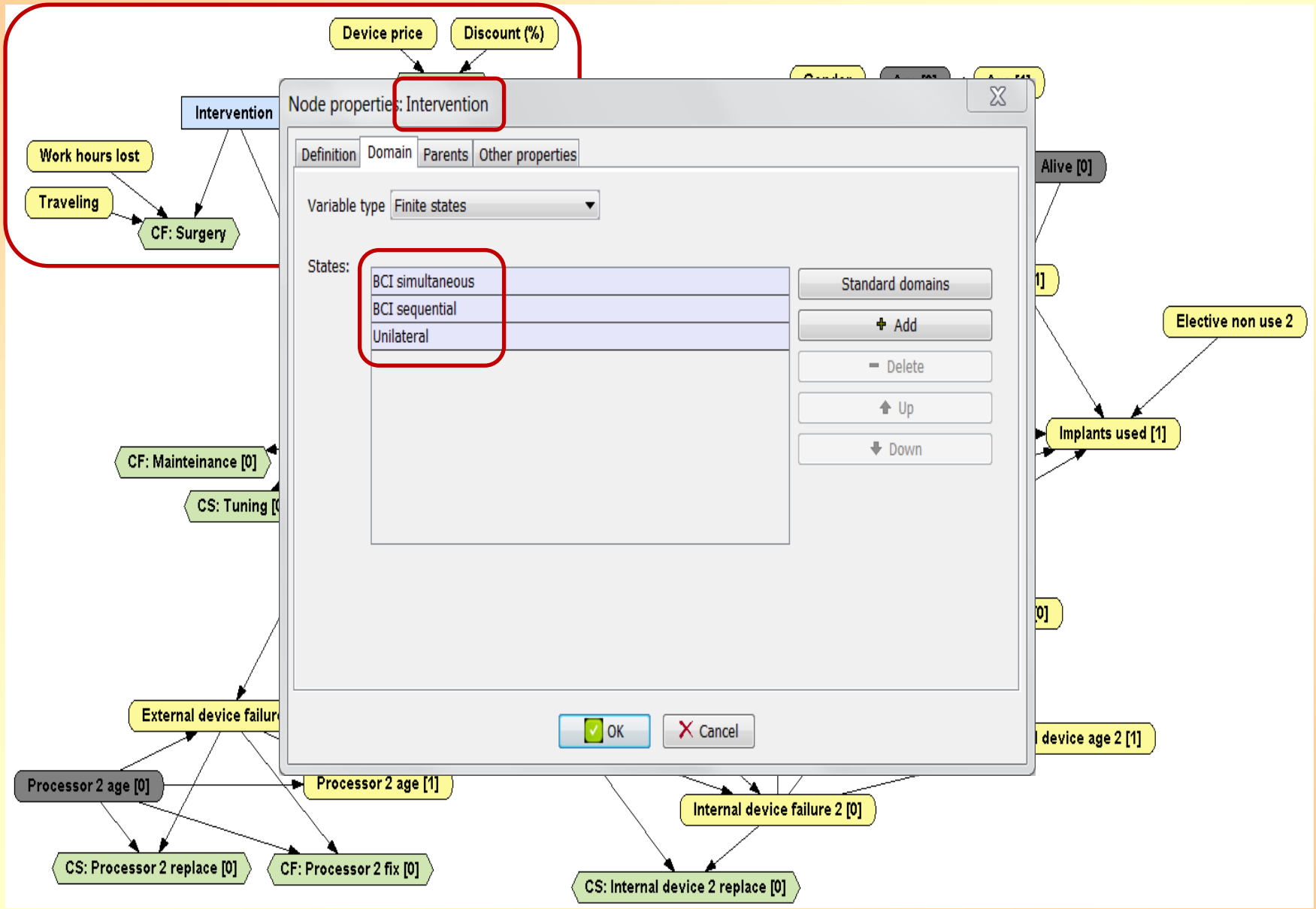


- *Cochlear Implant Symposium*, Washington DC, October 2015.

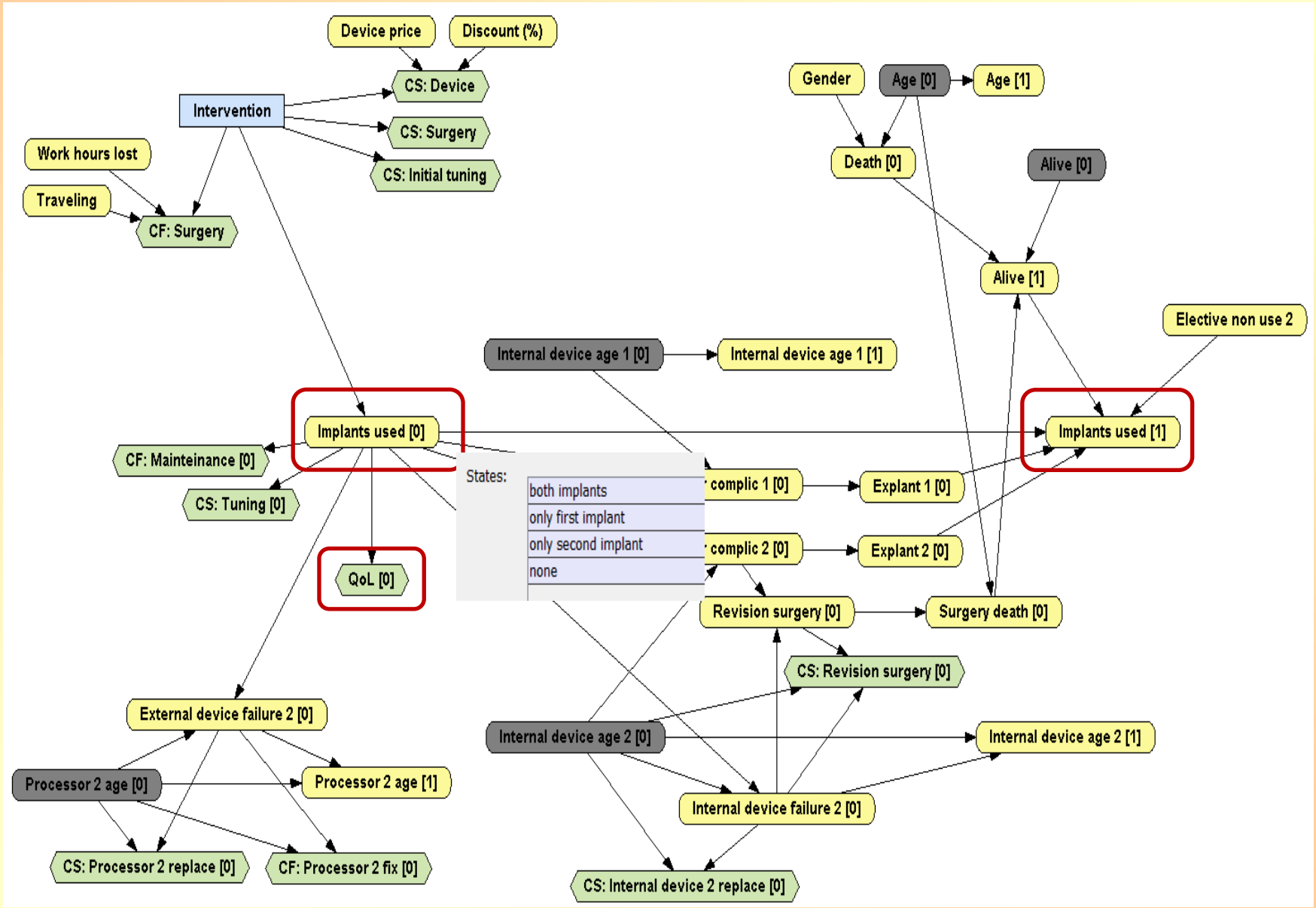
Our model for bilateral cochlear implantation



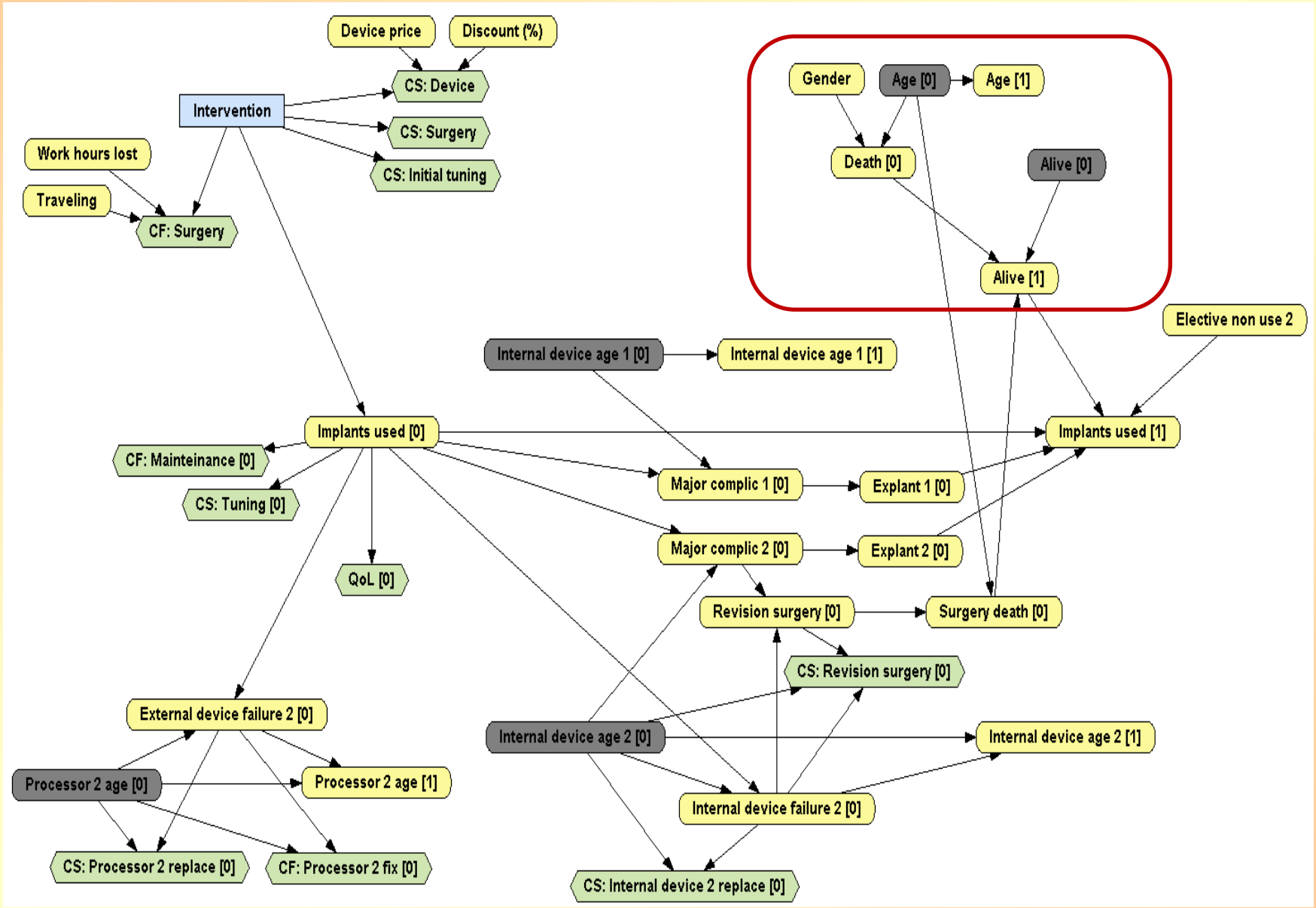
Our model for bilateral cochlear implantation



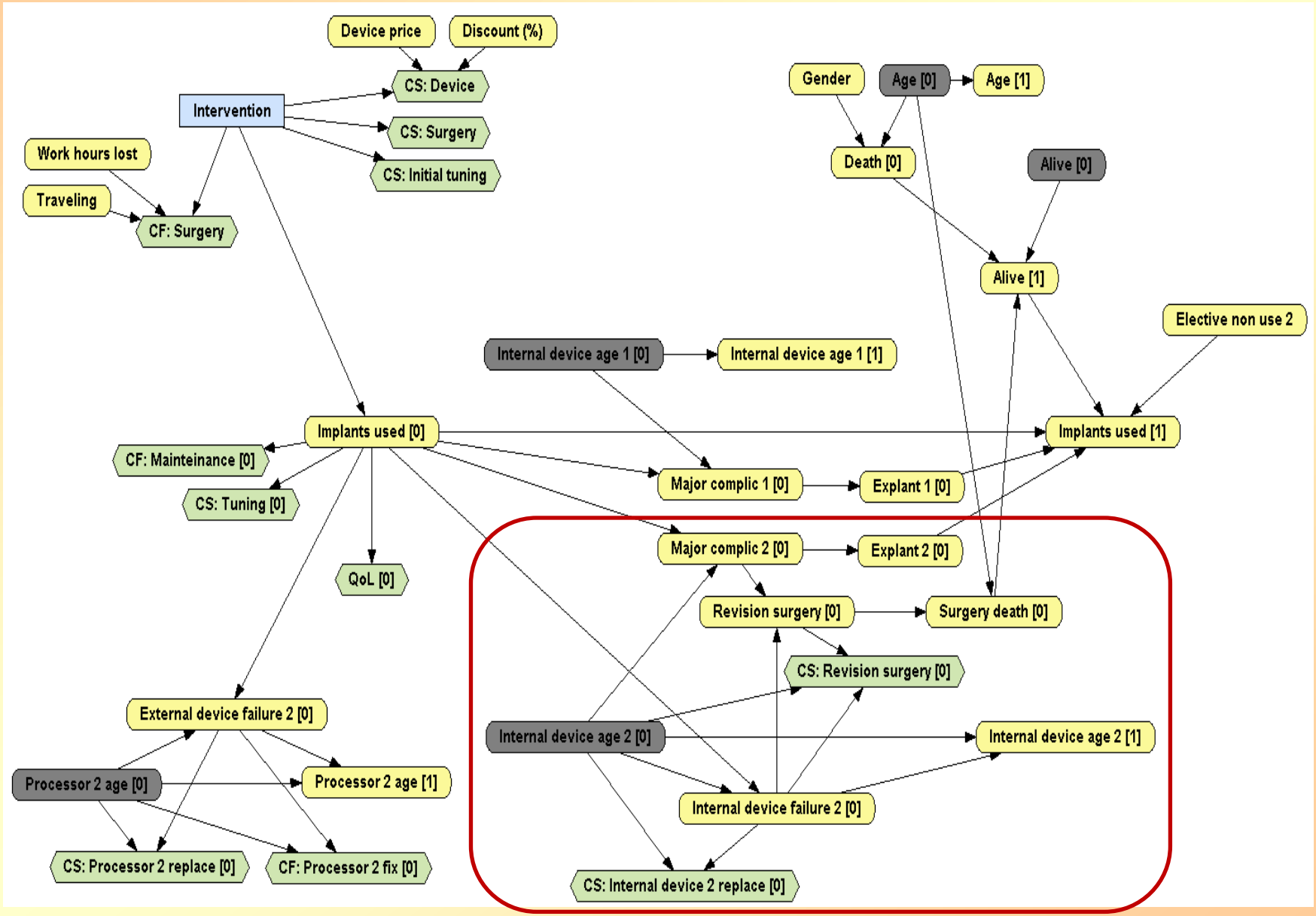
Our model for bilateral cochlear implantation



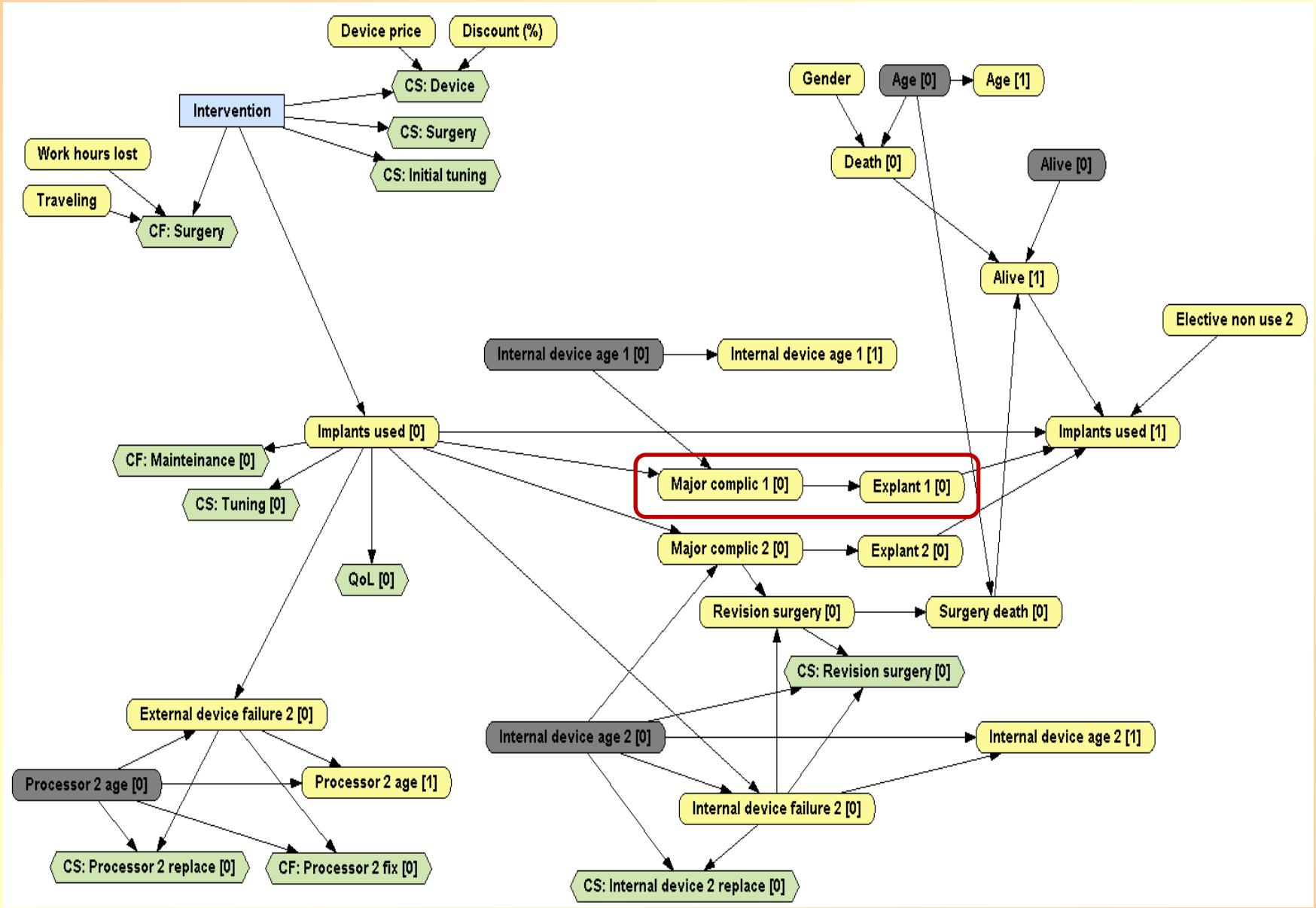
Our model for bilateral cochlear implantation



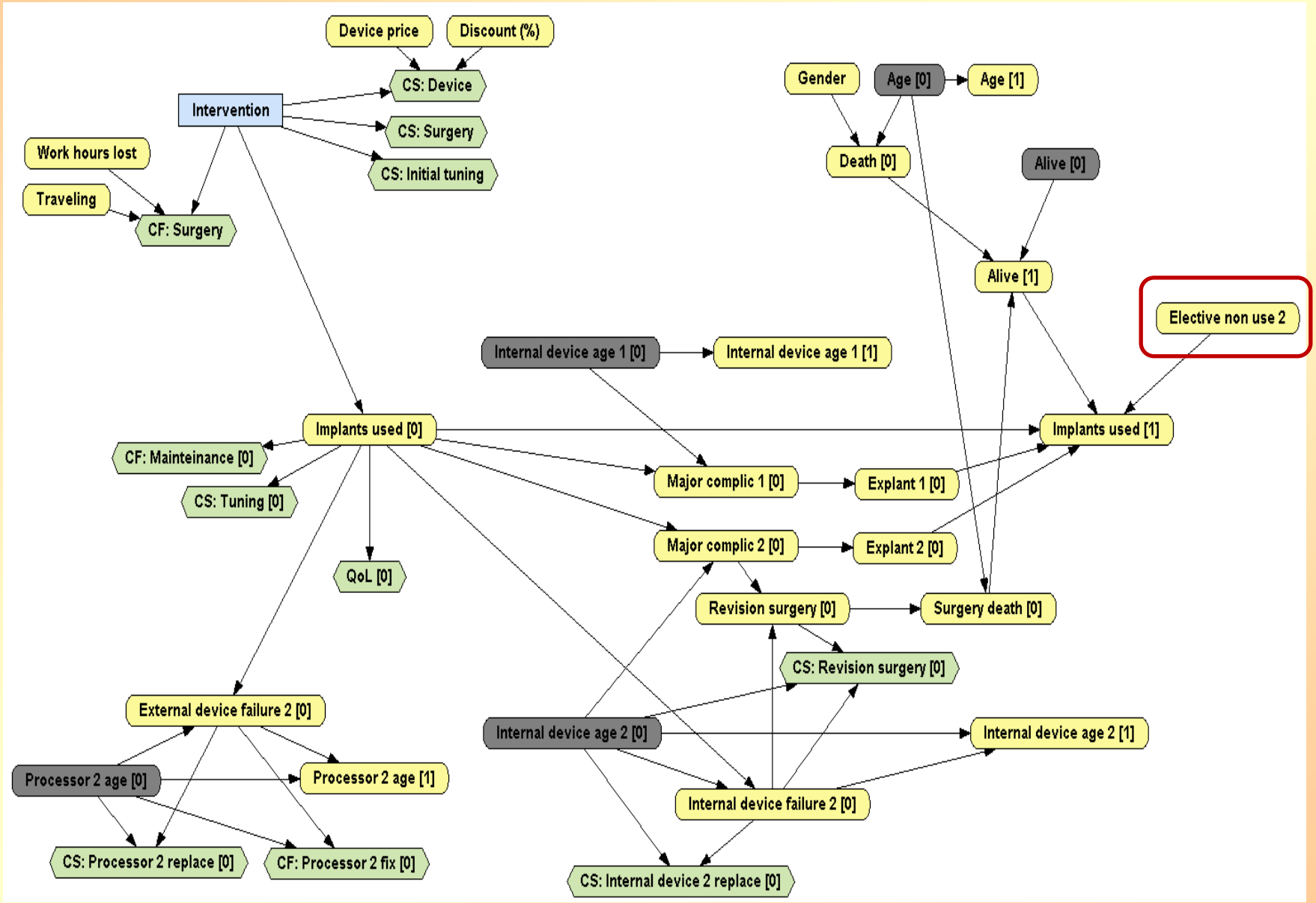
Our model for bilateral cochlear implantation



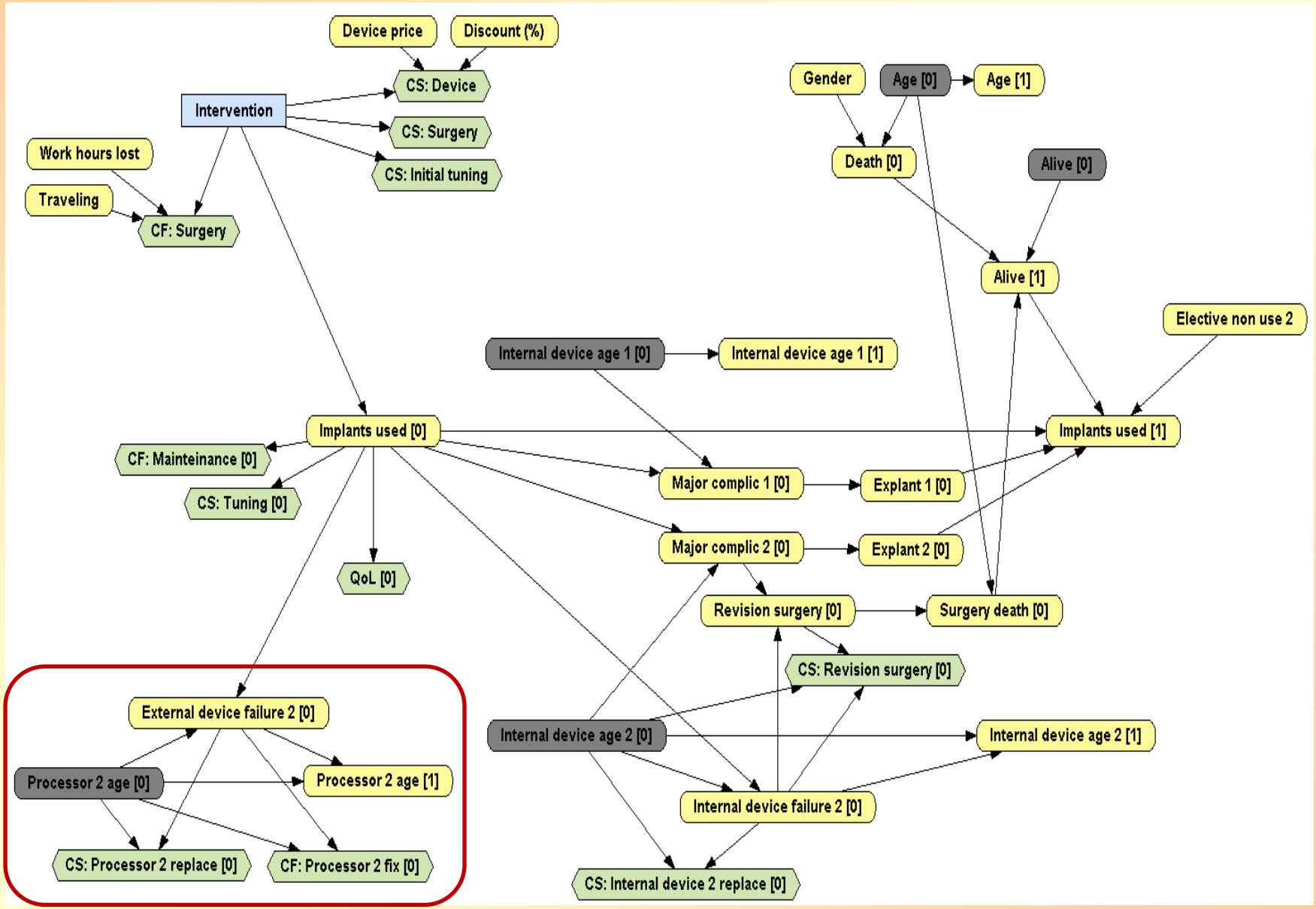
Our model for bilateral cochlear implantation



Our model for bilateral cochlear implantation

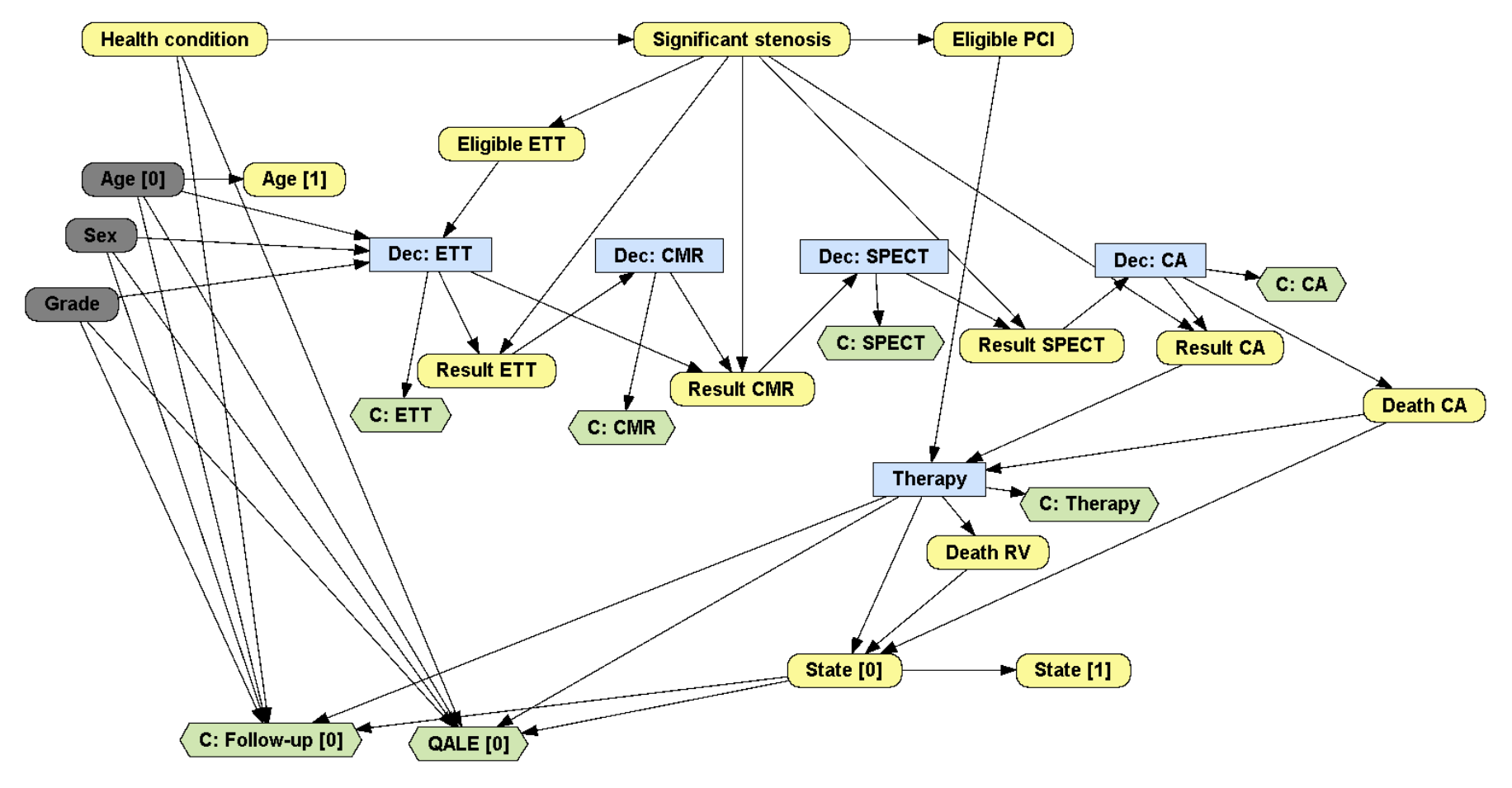


Our model for bilateral cochlear implantation



A MID with several decisions

Adapted from [Walker et al., 2013]



- This model evaluates all the possible interventions.
- It can cope with heterogeneity: sex, age, grade.

ORIGINAL ARTICLE

Cost-effectiveness of cardiovascular magnetic resonance in the diagnosis of coronary heart disease: an economic evaluation using data from the CE-MARC study

Simon Walker,[1] François Girardin,[1,2,3] Claire McKenna,[1] Stephen G Ball,[4] Jane Nixon,[5] Sven Plein,[4] John P Greenwood,[4] Mark Sculpher[1]

▸ Additional material is published online only. To view please visit the journal online (http://dx.doi.org/10.1136/heartjnl-2013-303624).

[1]Centre for Health Economics, University of York, York, UK
[2]Medical Direction, Geneva University Hospitals, Geneva, Switzerland
[3]Division of Clinical Pharmacology and Toxicology, Geneva University Hospitals, Geneva, Switzerland
[4]Multidisciplinary Cardiovascular Research Centre and Leeds Institute of Genetics, Health and Therapeutics, University of Leeds, Leeds, UK
[5]Clinical Trials Research Unit, University of Leeds, Leeds, UK

Correspondence to
Simon Walker, Centre for Health Economics, University of York, Alcuin A Block, Heslington, York YO10 5DD, UK; simon.walker@york.ac.uk

Received 10 January 2013
Revised 15 March 2013
Accepted 17 March 2013

ABSTRACT

Objective To evaluate the cost-effectiveness of diagnostic strategies for coronary heart disease (CHD) derived from the CE-MARC study.

Design Cost-effectiveness analysis using a decision analytic model to compare eight strategies for the diagnosis of CHD.

Setting Secondary care out-patients (Cardiology Department).

Patients Patients referred to cardiologists for the further evaluation of symptoms thought to be angina pectoris.

Interventions Eight different strategies were considered, including different combinations of exercise treadmill testing (ETT), single-photon emission CT (SPECT), cardiovascular magnetic resonance (CMR) and coronary angiography (CA).

Main outcome measures Costs expressed as UK sterling in 2010–2011 prices and health outcomes in quality-adjusted life-years (QALYs). The time horizon was 50 years.

Results Based on the characteristics of patients in the CE-MARC study, only two strategies appear potentially cost-effective for diagnosis of CHD, both including CMR. The choice is between two strategies: one in which CMR follows a positive or inconclusive ETT, followed by CA if CMR is positive or inconclusive (Strategy 3 in the model); and the other where CMR is followed by CA if

INTRODUCTION

Coronary heart disease (CHD) is a leading cause of death and disability worldwide. In the UK, over 2 million people are living with CHD and, in 2007, it was estimated to account for over 94 000 deaths, of which over 31 000 were considered premature.[1]

A variety of investigations may be used to diagnose CHD and identify patients who require coronary revascularisation; all these tests, however, have their limitations. Increasingly, non-invasive imaging has replaced exercise treadmill testing (ETT), with single-photon emission CT (SPECT) being the most commonly used test for myocardial ischaemia worldwide.[2] Cardiovascular magnetic resonance (CMR) imaging is increasingly used for the diagnosis of CHD as a result of its safety (no ionising radiation), high spatial resolution and ability to assess multiple aspects of CHD pathology in both the stable and unstable clinical settings.[3–8]

The diagnosis of CHD has no direct health benefit in itself; instead, any improved accuracy in diagnosis should result in more appropriate treatment which can confer health benefits on patients. The optimal management of patients with CHD continues to be debated, but options include medical therapy, percutaneous coronary intervention (PCI) or coronary artery bypass grafting (CABG). Many patients with

Model structure

To conduct the economic evaluation a decision analytic model was developed. For the initial diagnosis a decision tree allocates patients to the appropriate diagnostic group. The prognostic implications of being in one of these groups are then quantified using three distinct Markov models. An example of the decision tree for Strategy 2 (ETT, followed by CA if ETT is positive or inconclusive) is shown in figure 1.

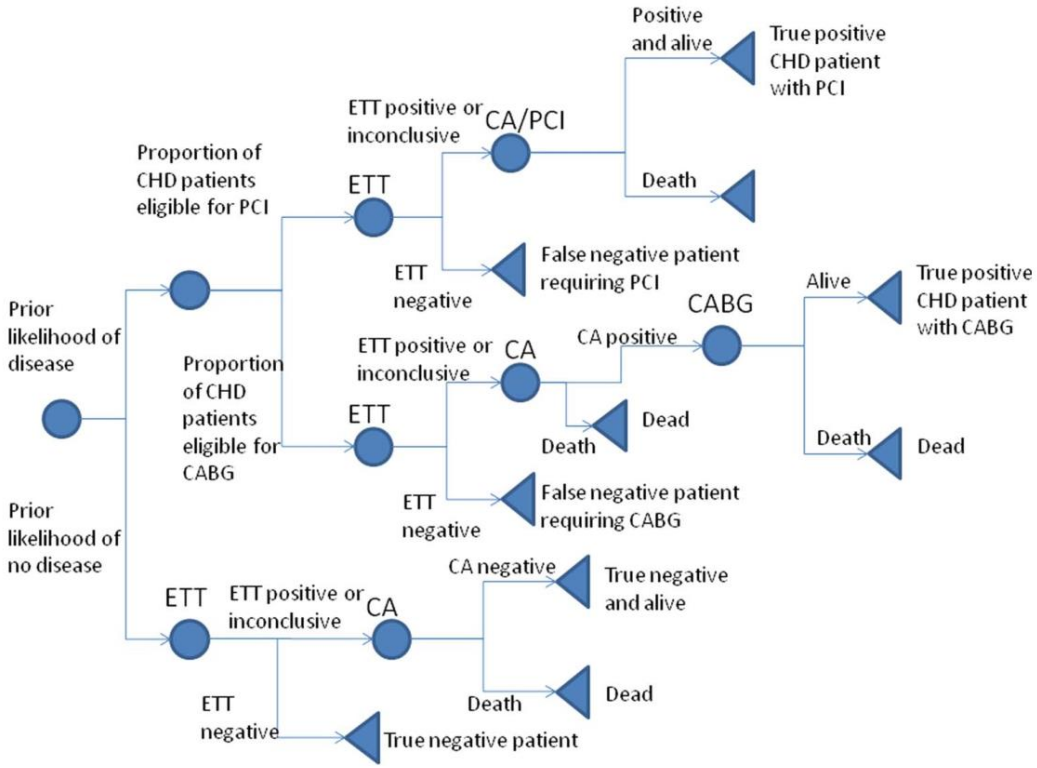


Figure 1 Structure of decision tree using Strategy 2 as an example. CA, coronary angiography; CABG, coronary artery bypass grafting; CHD, coronary heart disease; ETT, exercise treadmill testing; PCI, percutaneous coronary intervention.

Hands-on exercise 5

6.2.3. MIDs vs. other types of models

Advantages of MIDs for CEA

- For model builders
  - No programming is required, not even for sensitivity analysis
  - The construction of the model is much faster and easier.
  - It is possible to accomplish each phase (structure, numeric parameters, deterministic analysis, sensitivity analysis) without thinking of the next one
  - Debugging consists only of refining the knowledge contained in the model: it is not necessary to debug formulas and macros.
- For the recipients of the model (agencies: NICE, etc.)
  - Just by observing the graph it is possible to find out the basic structure of the model its main hypotheses.
  - It is not necessary to check that the code (formulas, macros…) is correct.

Comparison of MIDs with other techniques

- MIDs vs. spreadsheets (Excel)
  - no need to write any formulas nor VisualBasic macros
  - no need to multiply the number of states
- MIDs vs. Markov decision trees
  - much more compact $\Rightarrow$ possible to build much larger models
  - no need to add tracking variables (microsimulation)
- MIDs vs. a programming language (R, C++, MATLAB…)
  - no need to write any code, not even for sensitivity analysis
  - but programming languages are much more flexible
- MIDs vs. discrete event simulation
  - cohort propagation (exact algorithm) is often much faster
- MIDs vs. all the others: may contain several decisions.

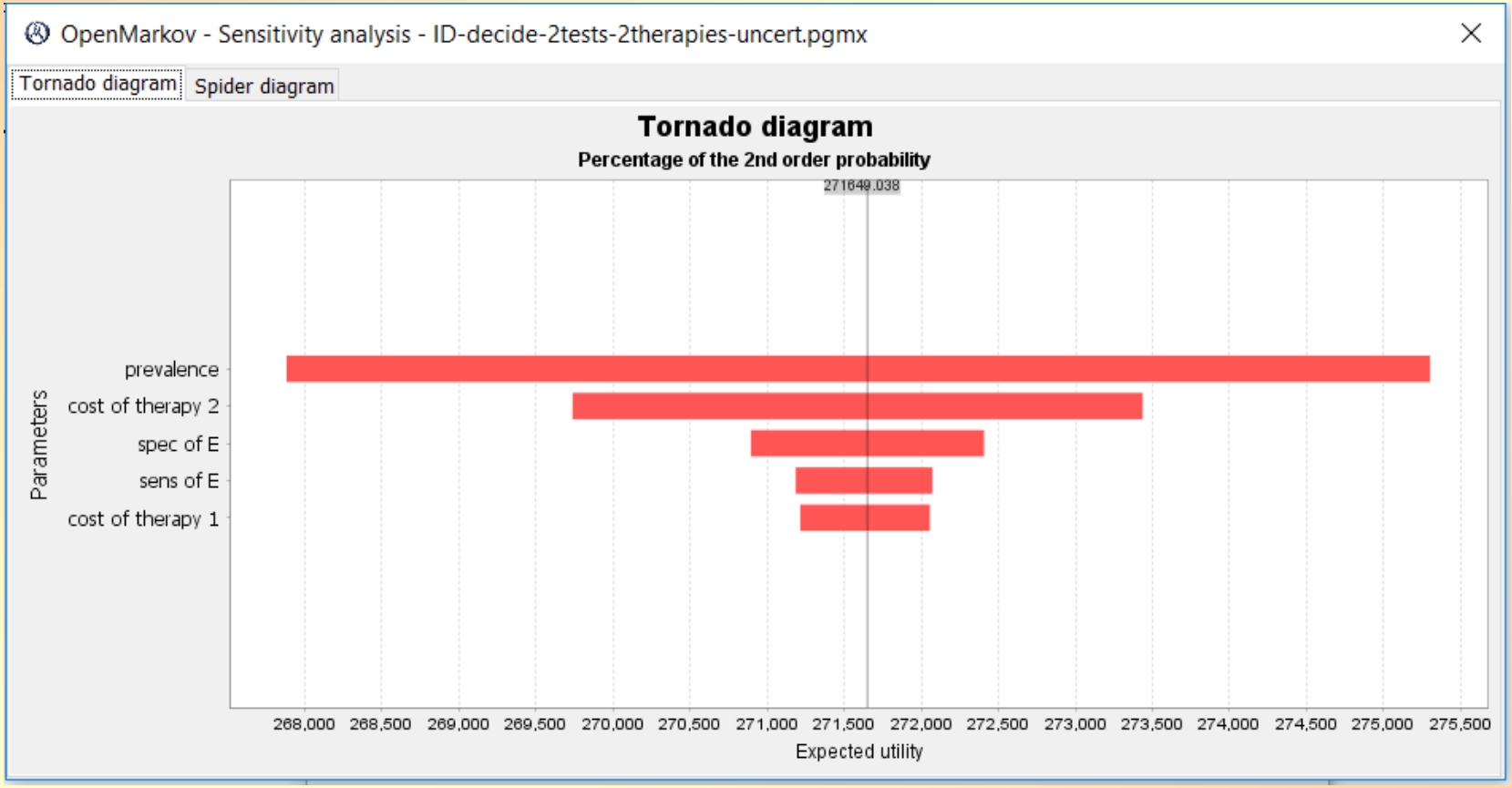
7. Sensitivity analysis

Types of sensitivity analysis

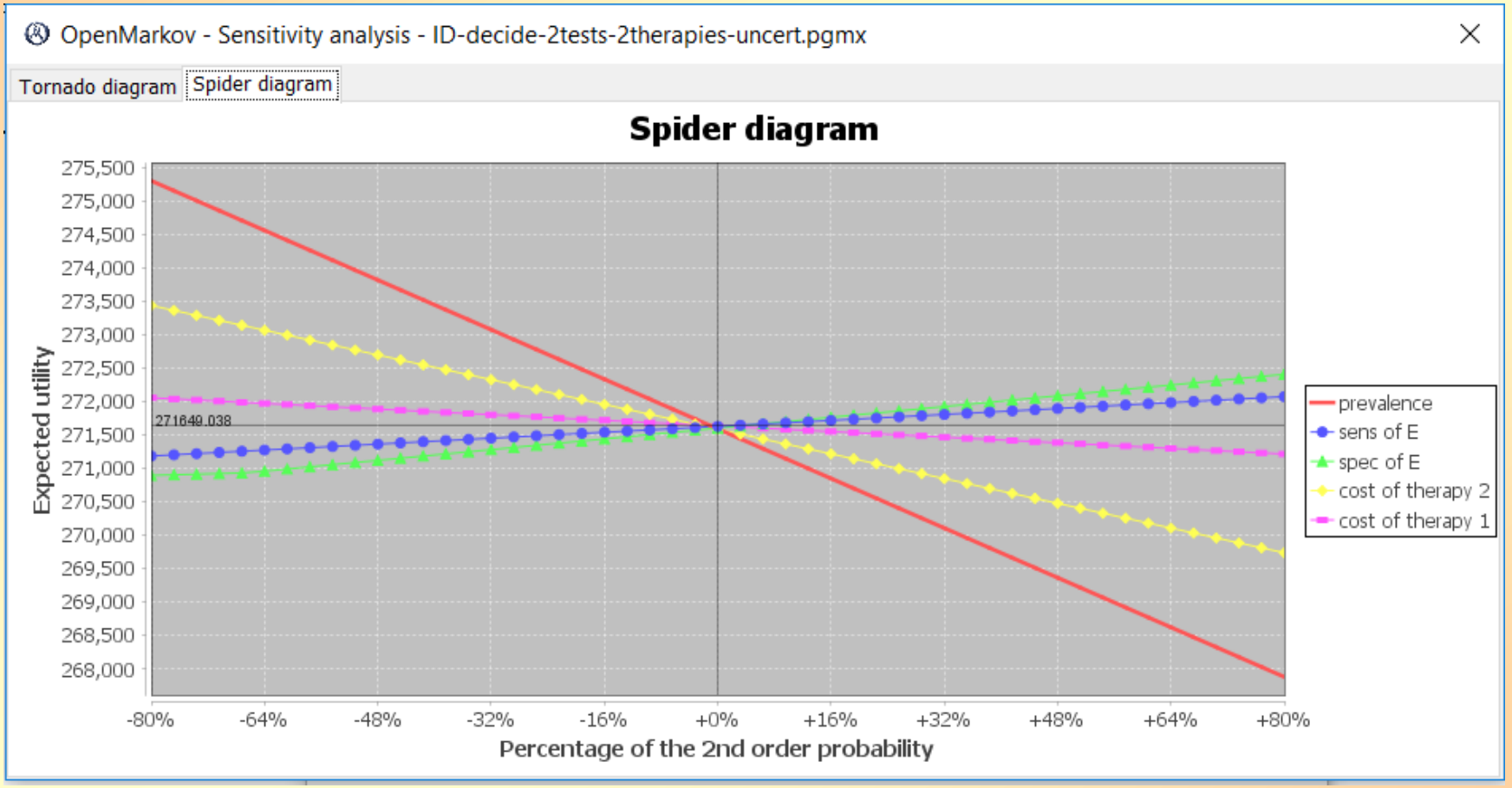
- Two main types
  - structural (qualitative)
  - parametric (quantitative)
- Depending on the effect analyzed
  - analysis of utility
  - analysis of decisions / policies
- Depending on how many parameters are varied
  - one-way analysis
  - *n*-way analysis (independent or join analysis)
- Depending on how the parameters are varied
  - range (interval)
  - probability distribution
  - look for thresholds

7.1. Unicriterion sensitivity analysis

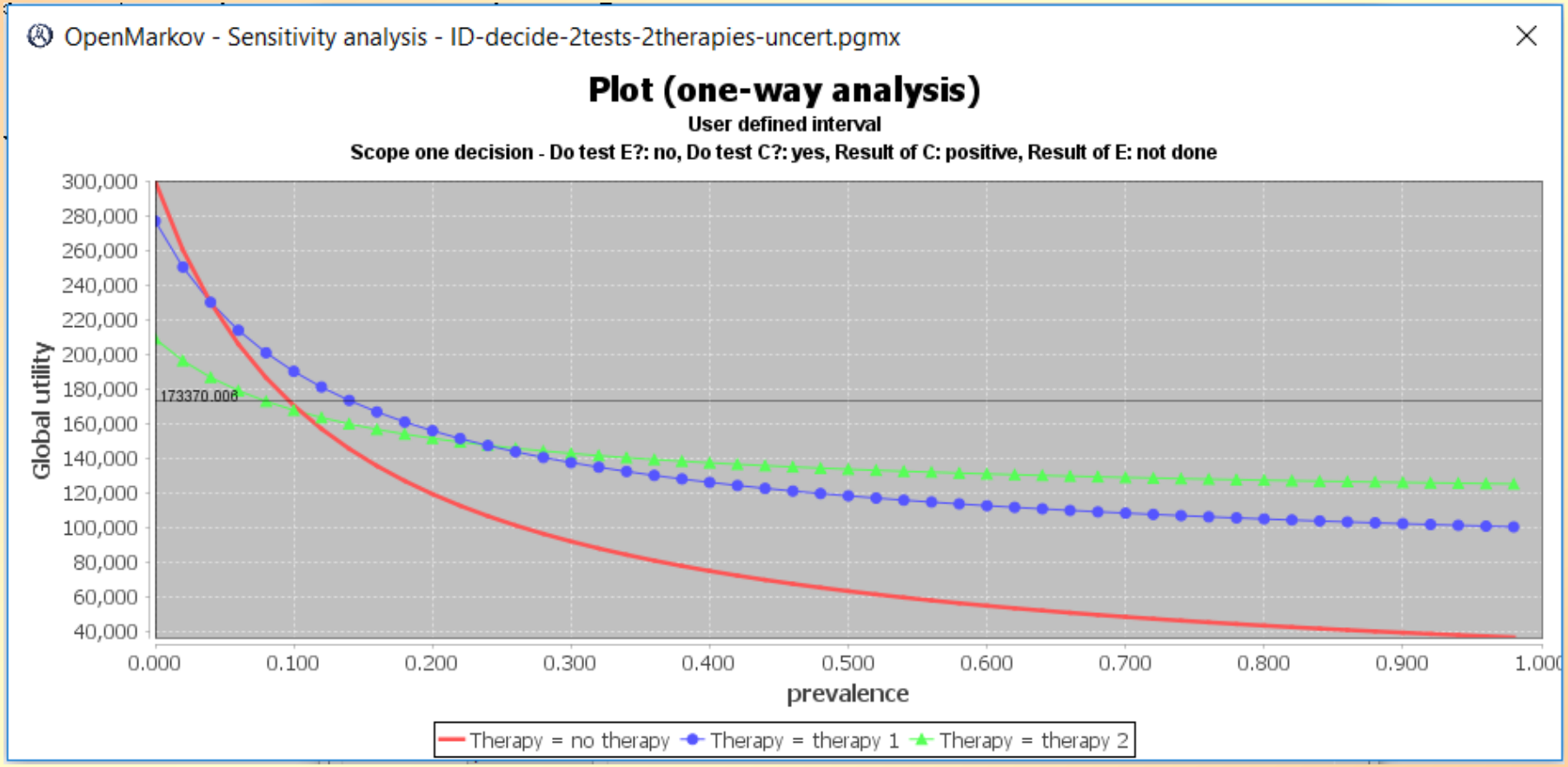
Tornado diagram



Spider diagram



Plot (one-way sensitivity analysis)



7.2. Cost-effectiveness sensitivity analysis

Scatter plot

Acceptability curve

Some sensitivity analysis options

	unicriterion	cost-effectiveness
deterministic	• tornado / spider diagram (global) • plot (global / for a decision) • map (global / for a decision)	• C.E. spider diagram (global)
probabilistic	• acceptability (for a decision) • EVPI (global)	• scatter plot + acceptability curve (for a decision) • EVPI curve (global)

8. Overview of software tools

Software Packages for Graphical Models

Written by Kevin Murphy.
Last updated 16 June 2014.
(Thanks to Alex Gorban for helping me with the switch to Google Sheets.)

Review articles

- List of GM code at MLOSS
- Click here for a short article I wrote for the ISBA (International Society for Bayesian Analysis) Newsletter, December 2007, sumarizing some of the packages below.
- Click here for a more detailed discussion of some of these packages written by Ann Nicholson and Kevin Korb in 2004.
- Click here for a French version of my comparison table (not necessarily up-to-date).

What do the headers in the table mean?

- Src = source code included? (N=no) If so, what language?
- Cts = are continuous (latent) nodes supported? G = (conditionally) Gaussians nodes supported analytically, Cs = continuous nodes supported by sampling, Cd = continuous nodes supported by discretization, Cx = continuous nodes supported by some unspecified method, D = only discrete nodes supported.
- GUI = Graphical User Interface included?
- Learns parameters?
- Learns structure? CI = means uses conditional independency tests
- Utility = utility and decision nodes (i.e., influence diagrams) supported?
- Free? 0 = free (although possibly only for academic use). $ = commercial software (although most have free versions which are restricted in various ways, e.g., the model size is limite or models cannot be saved, or there is no API.)
- Undir? What kind of graphs are supported? U = only undirected graphs, D = only directed graphs, UD = both undirected and directed, CG = chain graphs (mixed directed/undirected).
- Inference = which inference algorithm is used? jtree = junction tree, varelim = variable (bucket) elimination, MH = Metropols Hastings, G = Gibbs sampling, IS = importance sampling sampling = some other Monte Carlo method, polytree = Pearl's algorithm restricted to a graph with no cycles, VMP = variational message passing, EP = expectation propagation, SL = the program is designed for structure learning from completely observed data, not state estimation
- Comments. If in "quotes", I am quoting the authors at their request.

If you want your package to be listed, please fill out this form.

Name	Authors	Src	Cts	GUI	Params	Struct	Utility	Free	Undir	Inference	Comments
AgenaRisk	Agena	N	Cx	Y	Y	N	N	$	D	JTree	Simulation by Dynamic discretisatio
Analytica	Lumina	N	G	Y	N	N	Y	$	D	sampling	spread sheet compatible
B-course	U. Helsinki	N	Cd	Y	Y	Y	N	0	D	?	Runs on their server: view results us
Banjo	Hartemink	Java	Cd	N	N	Y	N	0	D	none	structure learning of static or dynam
Bassist	U. Helsinki	C++	G	N	Y	N	N	0	D	MH	Generates C++ for MCMC. (No long

69 packages!

Open-source tools for PGMs

	Weka	JavaBayes	Elvira	BNT	Riso	UnBBayes	OpenMarkov	BayesLine	PNL	BNJ	OBP
Start	1993	1996	1997	1999	2000	2000	2002	2003	2003	2004	2006
Stopped	—	2001	2010	2007	2004	2014	—	2003	2005	2004	2007
Programming language	Java	Java	Java	Matlab	Java	Java	Java	Java	C++	Java	Python
License	GPL	GPL	?	GPL	GPL	GPL	GPL	LGPL	IOSL	GPL	GPL
Bayesian networks	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Influence diagrams	no	no	yes	yes	no	yes	yes	no	no	no	no
Dynamic/Markov models	no	no	no	yes	no	no	yes	no	no	no	no
User manuals	yes	yes	yes	yes	yes	yes	yes	no	no	yes	yes
Developer manuals	yes	no	no	no	no	yes	yes	no	no	no	no
Users list/forum	yes	no	no	yes	yes	yes	yes	yes	no	yes	yes
Developers list/forum	yes	no	yes	yes	yes	yes	yes	yes	no	yes	yes
Source HTML docs	yes	yes	yes	no	yes	yes	yes	yes	no	no	no
Version control	yes	no	yes	no	yes	yes	yes	yes	no	yes	yes
Bug tracker	yes	yes	no	no	yes	yes	yes	yes	no	yes	yes

- Only BNT and OpenMarkov can represent Markov models.
- Among the tools having a GUI for editing PGMs, only Weka and OpenMarkov are still under active development.

OpenMarkov. Main features

- Main advantage: open source
  - Free
  - Users can adapt it to their needs
  - Software engineering tools:
    JUnit, maven, mercurial (bitbucket), nexus, bugtracker, etc.
- Strengths
  - Written in Java: portability (Windows, linux, MacOS…)
  - Many types of models, potentials, etc.
  - Algorithms not available in any other package
    - CEA with IDs
    - interactive learning
  - Very active: new features are continuously added
  - Support for users and developers: wiki, lists, mail…
  - Well-documented format for encoding networks: ProbModelXML.

OpenMarkov. Limitations

- Main weakness
  - Still a prototype: needs debugging
- Other weaknesses
  - Written in Java: relatively slow (in some cases)
  - No on-line help, documentation still poor
  - Support is limited, due to scarcity of human resources.

8. Conclusions

Conclusions

- BNs overcame the limitations of the naïve Bayes method.
- IDs have several advantages over decision trees, but also have serious limitations for medical decision making.
- DANs are similar to IDs, but more suitable for asymmetric decision problems.
- It is possible to do cost-effectiveness analysis with IDs.
- and also with Markov IDs (MIDs) if all decisions are atemporal.
- There are other types of Markov PGMs having one or more decisions per cycle: MDPs, POMDPs, DLIMIDs…

How to bring PGMs from artificial intelligence into medical decision making

- Dissemination
  - Seminars, short courses…
  - Tutorials and textbooks written in the language of clinicians, epidemiologists and health economists
- Research
  - New methods for the representation of knowledge
  - New algorithms for CEA, sensitivity analysis…
- User-friendly software tools
  - for building, debugging and maintaining the models
  - for displaying the results using charts, tables, etc.

Future work

- New models and algorithms
  - Markov DANs
  - CEA with models having one or several decisions per cycle
  - new methods for CEA, sensitivity analysis, explanation of "reasoning"…
- Integration of PGMs, cost-effectiveness analysis, and Bayesian inference
  - integration of OpenMarkov with OpenBUGS and/or STAN.

Thank you very much for your attention!

- Links
  - www.cisiad.uned.es
  - www.OpenMarkov.org
  - www.ProbModelXML.org/networks
- Contact: fjdiez@dia.uned.es