Instituto de Salud Carlos III
Madrid, 9 octubre 2019

Análisis de decisiones en medicina con modelos gráficos probabilistas

Francisco Javier Díez
Dept. Artificial Intelligence. UNED
Madrid, Spain

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

OVERVIEW

- The naïve-Bayes method
- Bayesian networks
- Influence diagrams
- Decision analysis networks
- Cost-effectiveness analysis
- Markov models
- Conclusion

Naïve-Bayes method for probabilistic diagnosis

- *n* diagnoses, *m* possible findings
- 1st hypothesis: diagnoses are <u>mutually exclusive</u>

  (i.e., the patient has at most one disease)
- 2nd hypothesis: findings are <u>conditionally independent</u>

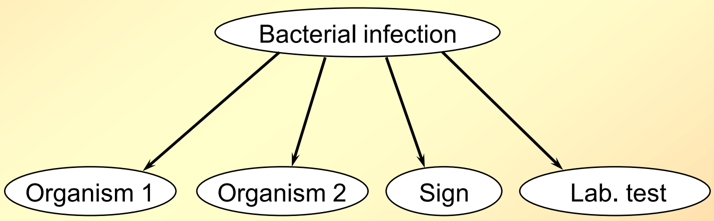
$$P(f_1, \ldots, f_m \mid d_i) = P(f_1 \mid d_i) \cdot \ldots \cdot P(f_m \mid d_i)$$

- Bayes' theorem (naïve method)

$$P(d_i \mid f_1, \ldots, f_m) = \alpha \cdot P(f_1 \mid d_i) \cdot \ldots \cdot P(f_m \mid d_i) \cdot P(d_i)$$

Limitations of the naïve-Bayes method

- In general the diagnoses are not mutually exclusive: how to diagnose <u>multiple disorders</u>.
- In general findings are not conditionally independent.

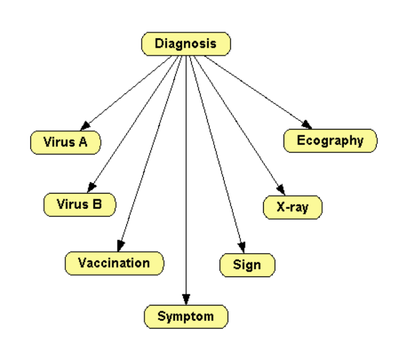


Succesfull applications of the naïve-Bayes

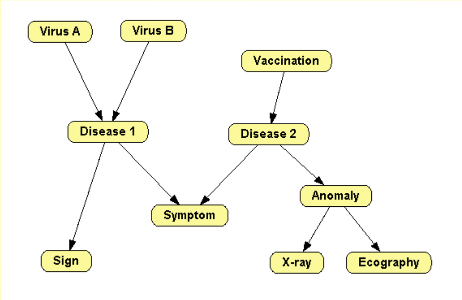
- Lodwick GS, Haun CL, Smith WE, et al. , “Computer diagnosis of primary bone tumors: A preliminary report,” *Radiology* **80** (1963) 273-275.
- Overall JE, Williams CM, “Conditional probability program for diagnosis or thyroid function,” *JAMA* **183** (1963) 307-313.
- Toronto AF, Veasy LG, Warner 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.

Bayesian networks

Naïve Bayes



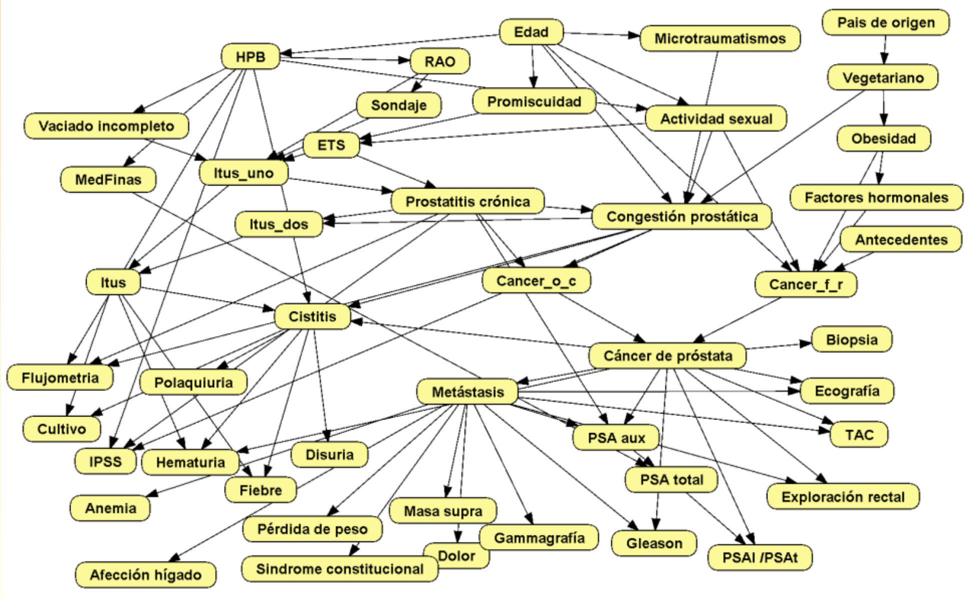
Bayesian network



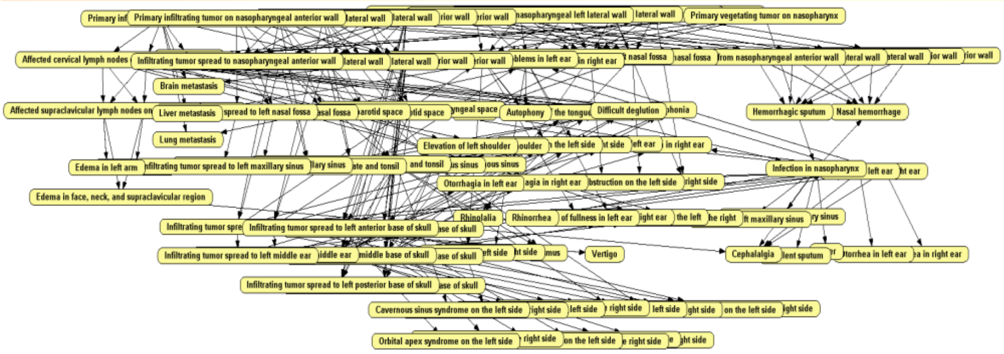
Examples of BNs

- 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 dieseases
    Agnieszka Onisko's thesis, 2003
  - Catarnet: Cataract surgery
    Nuria Alonso's thesis, 2009

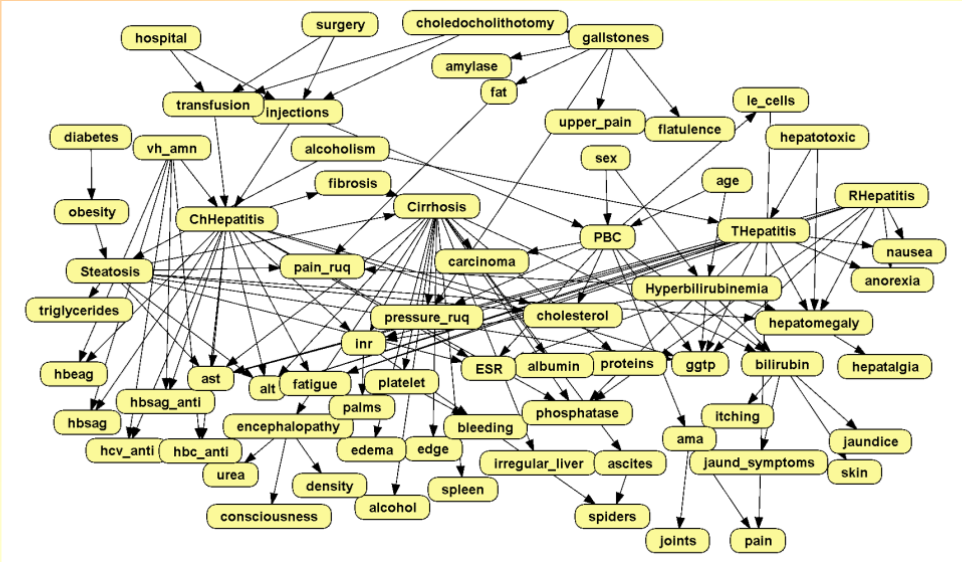
Prostanet (for prostate diseases)



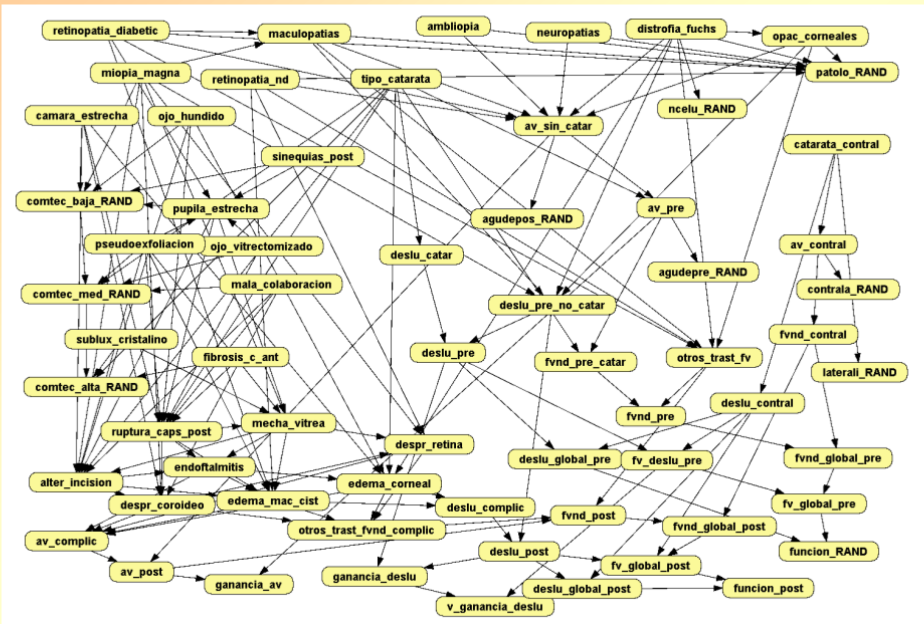
Nasonet (nasopharyngeal cancer spread)



Hepar II (liver diseases)

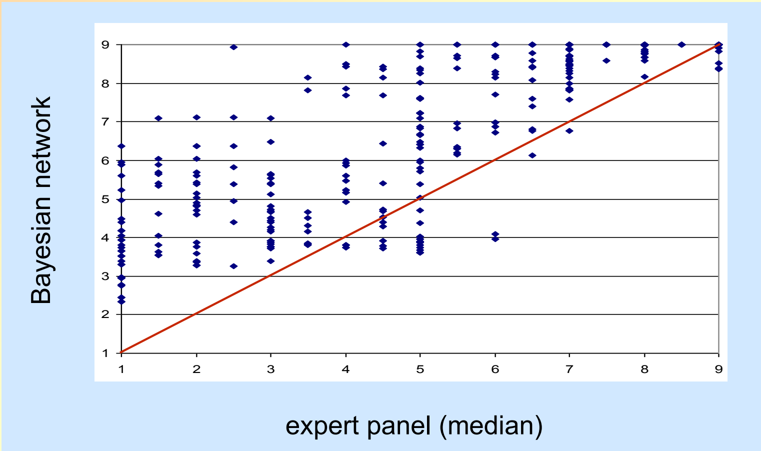


Catarnet (cataract surgery)



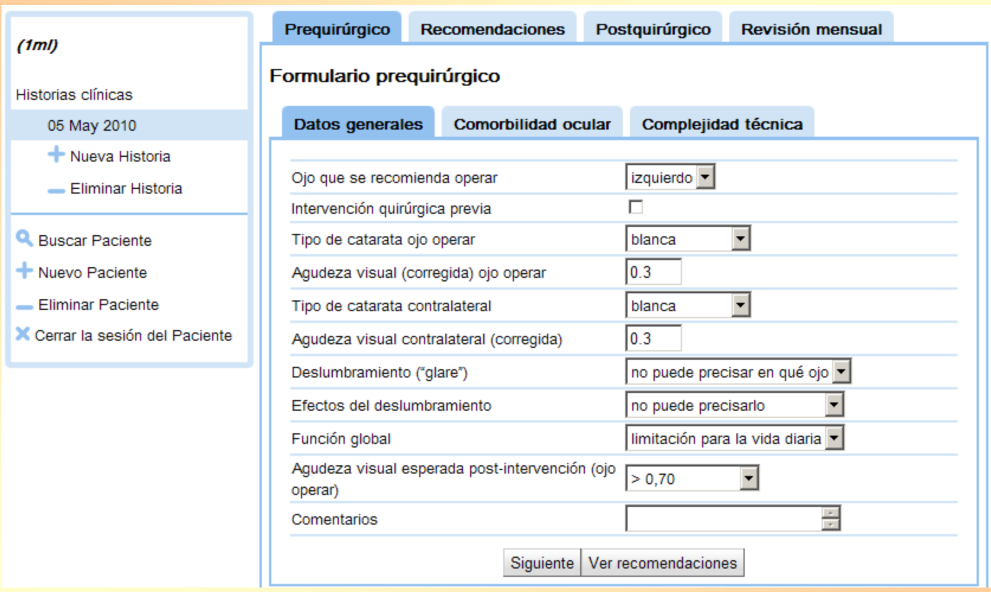
BN vs. a panel of experts (Delphi)

- Comparison in 429 clinical scenarios

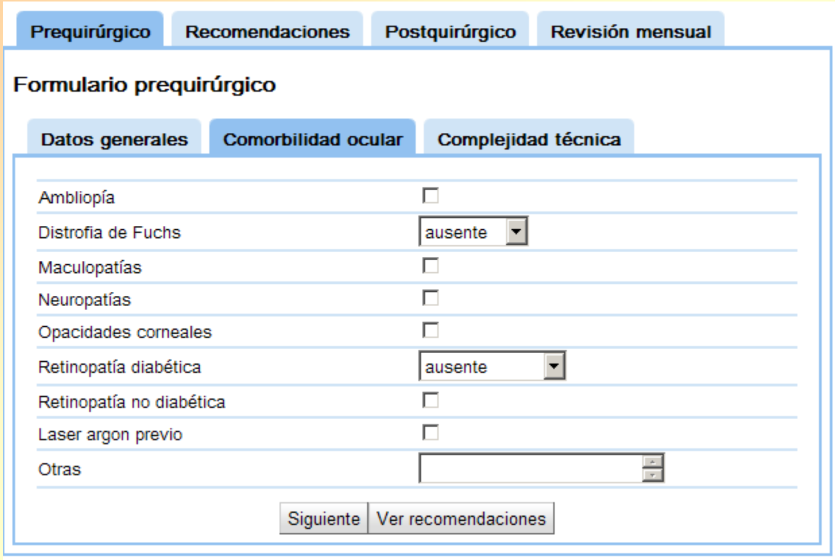


- Result: ICC=0.83 [IC95%: 0.80 – 0.86] ($p<0.001$)

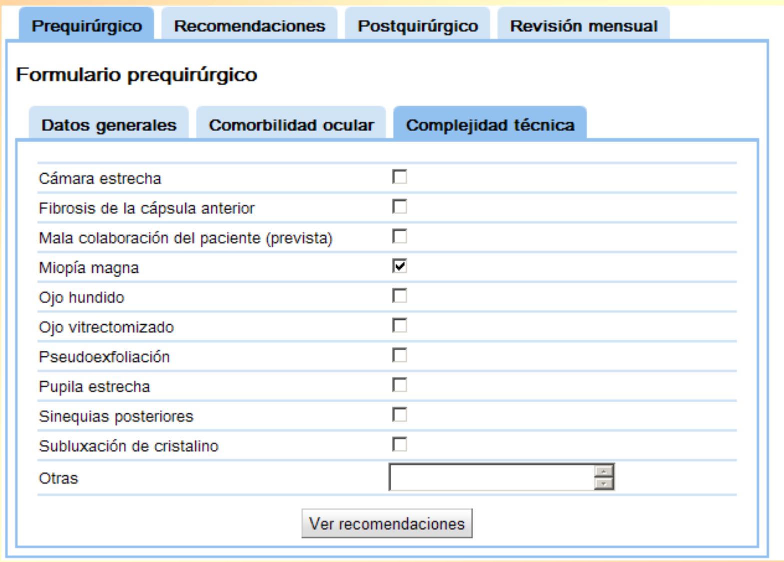
Input: 1. General data



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

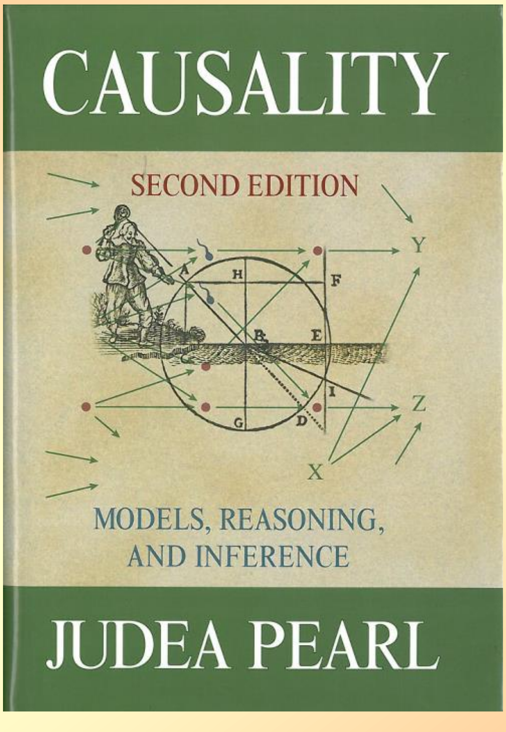
Red bayesiana CatarNet

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



HARVARD T.H. CHAN | SCHOOL OF PUBLIC HEALTH

Email People Departments Calendar Careers Give my.harvard

ABOUT | FACULTY & RESEARCH | ADMISSIONS & AID | ACADEMICS | EXECUTIVE/CONTINUING ED | NEWS

Miguel Hernan

> Miguel Hernan > Causal Inference Book

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

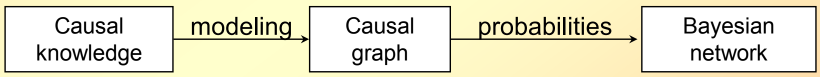
Causal Inference

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

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.

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!

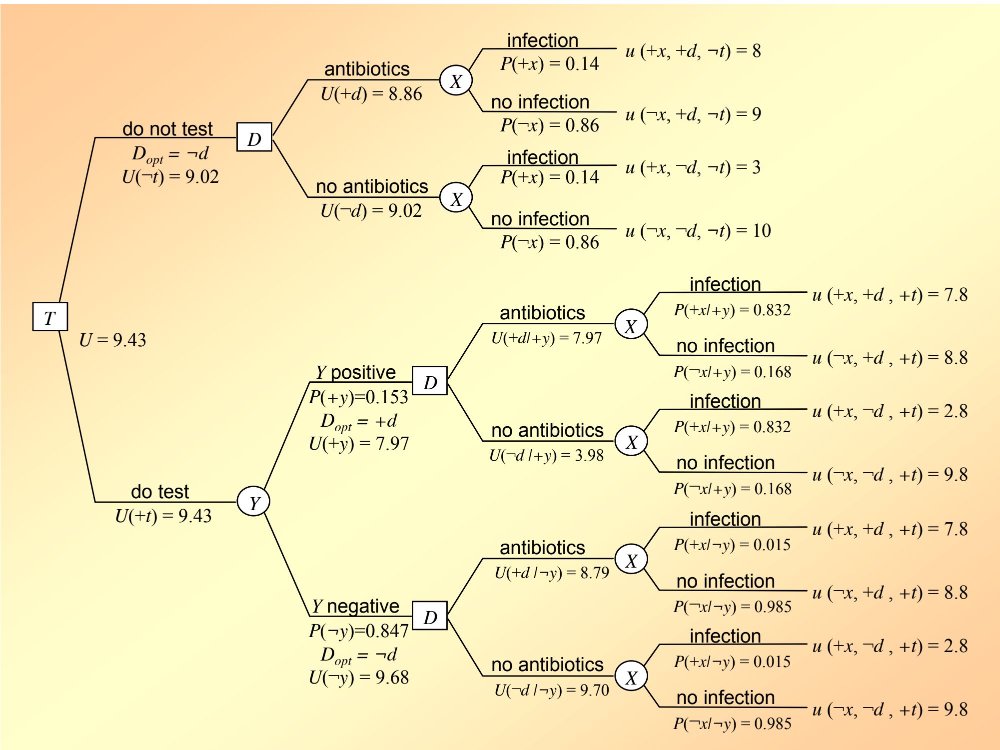
Influence diagrams

A medical problem

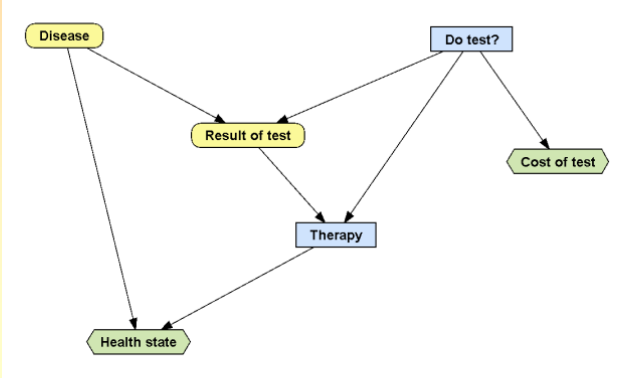
- **Disease** X
  - Prevalence: $P(+x) = 0'14$
- **Therapy** D
  - Utility:

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

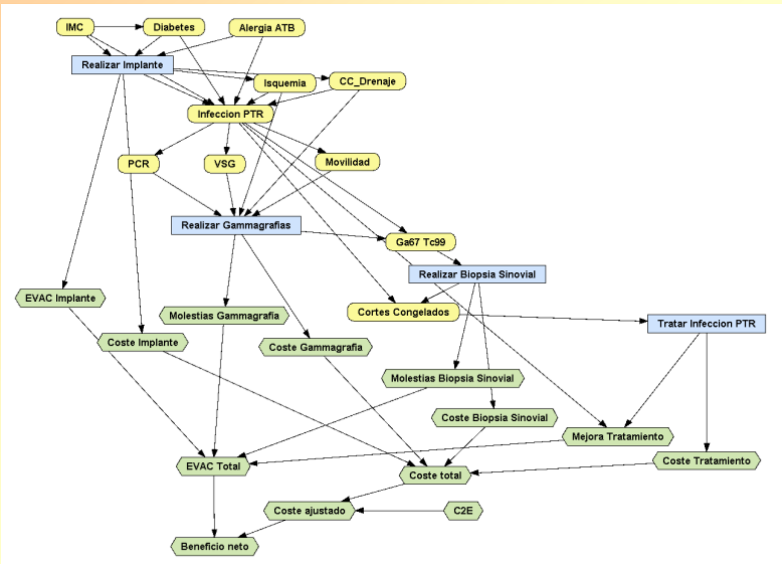
- **Test** Y
  - Sensitivity: $P(+y/+x) = 0'91$
  - Specificity: $P(\neg y/\neg x) = 0'97$
  - Cost: $u_{\text{test}}(x, d) = u_{\text{not-test}}(x, d) - 0'2$
- **Decisions:**
  - Is it worth doing the test?
  - In what cases should we apply the therapy?



An ID for this example

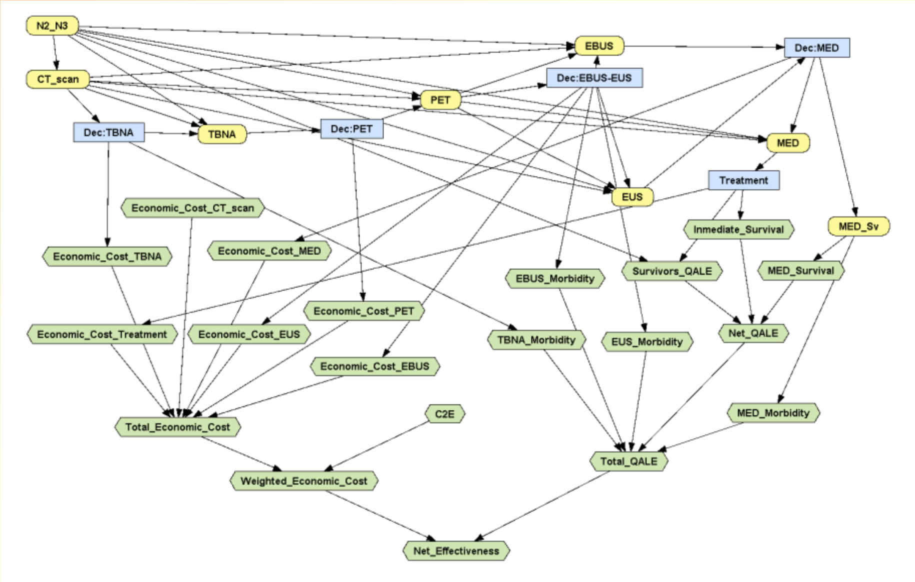


Arthronet (total knee arthroplasty)



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

Mediastinet (lung cancer)



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

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

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

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

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.

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?

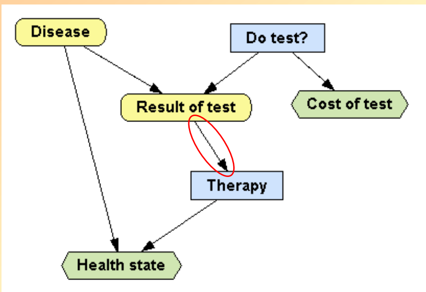
Limitations of IDs

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

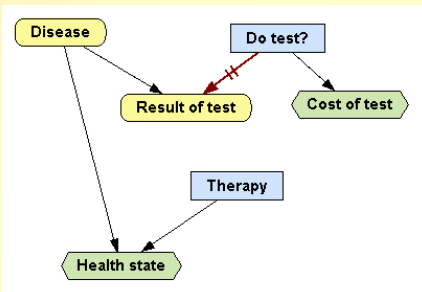
1. Explanation in influence diagrams
   - showing the posterior probabilities and expected values
   - introduction of evidence
   - hypothetical reasoning (what if) by means of imposed policies
2. Synthesizing the optimal intervention
   - in the form of a compact tree
3. Decision analysis networks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

Influence diagram



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

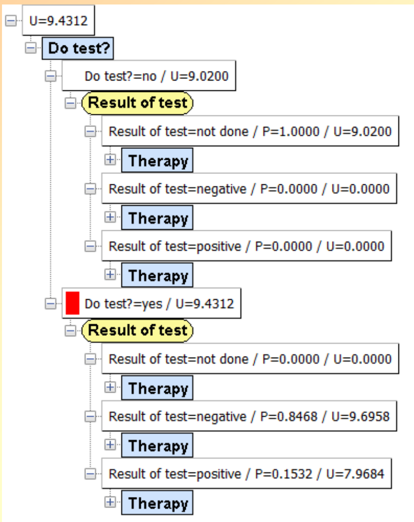
DAN



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

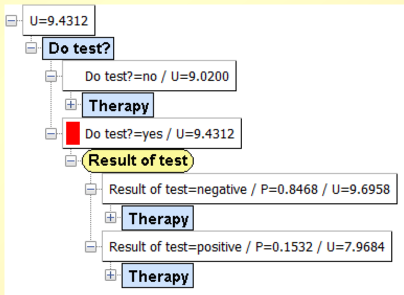
- Different ways of indicating the flow of information.
- The decision trees are different but equivalent:
  the same probabilities, utilities, and policies.

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

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

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

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.

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

A R T I C L E I N F O

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

A B S T R A C T

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

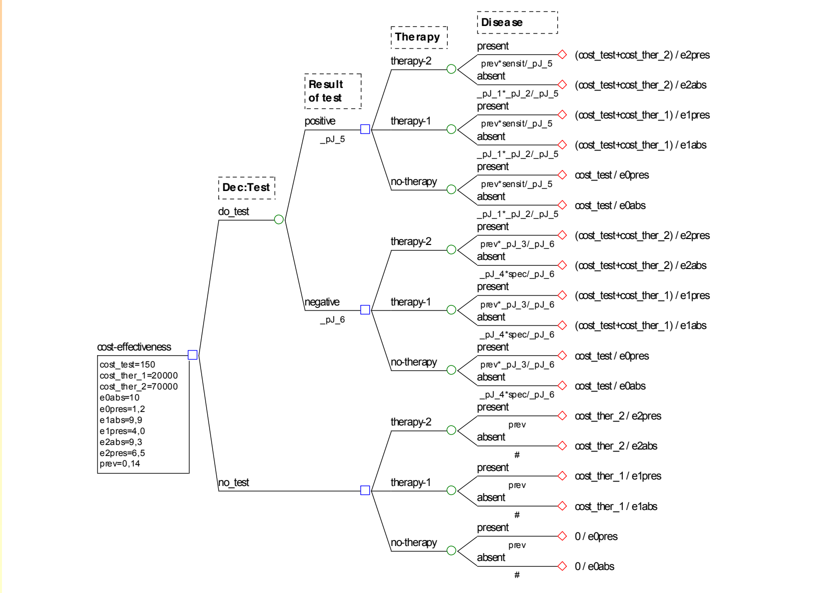
The two formalisms most widely used for the representation and analysis of decision problems are decision trees (DTs) [31] and influence diagrams (IDs) [15]. DTs have the advantage of almost absolute flexibility, but also have three drawbacks: their size grows exponentially with the number of variables,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

5. Multicriteria decision making

Example with uncertain outcomes: cost-effectiveness of a test

- Two therapies, with different cost and effectiveness
- A test, which also has a cost
- There is uncertainty (probabilities):
  - prevalence of the disease: 0.14
  - test: sensitivity 0.90
specificity 0.93
- Questions:
  - When is the test cost-effective? = What is its ICER?
  - What is the most beneficial therapy for each value of λ?

A decision tree for this example



Problem: the standard algorithm only works for the unicriterion case

A warning and a (rudimentary) solution

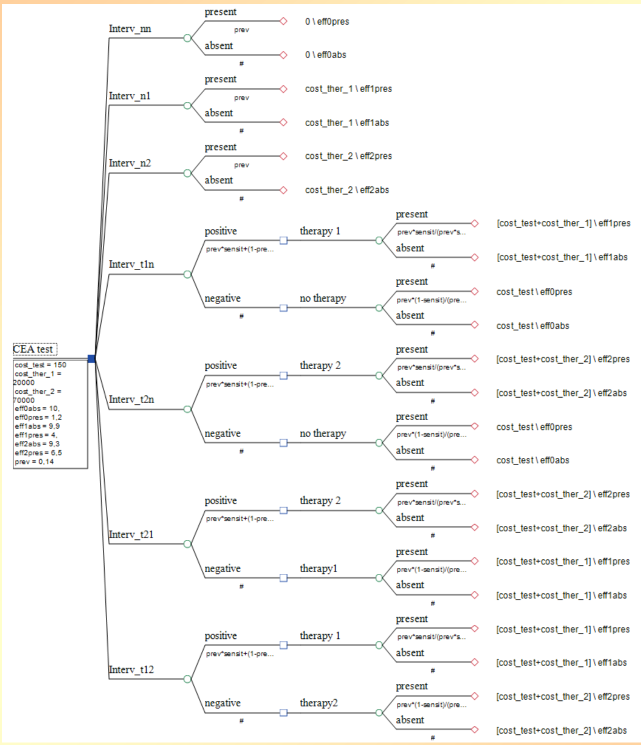
"Embedded, or downstream, decision nodes are not useful in cost-effectiveness analysis because the optimal branch cannot be determined when folding back the tree without an explicit decision rule for comparing costs and consequences.

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.~~



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

CISIAD

Español | English

Home

Members

Research

- Areas
- Projects
- Contracts
- Seminars
- Colleagues

Publications

- Books
- Papers
- Conferences and workshops
- Technical reports
- Theses

Postgraduate courses (in Spanish)

- Medicina (modular)
- Master AI
- Doctorado

News

Links

Contact

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)

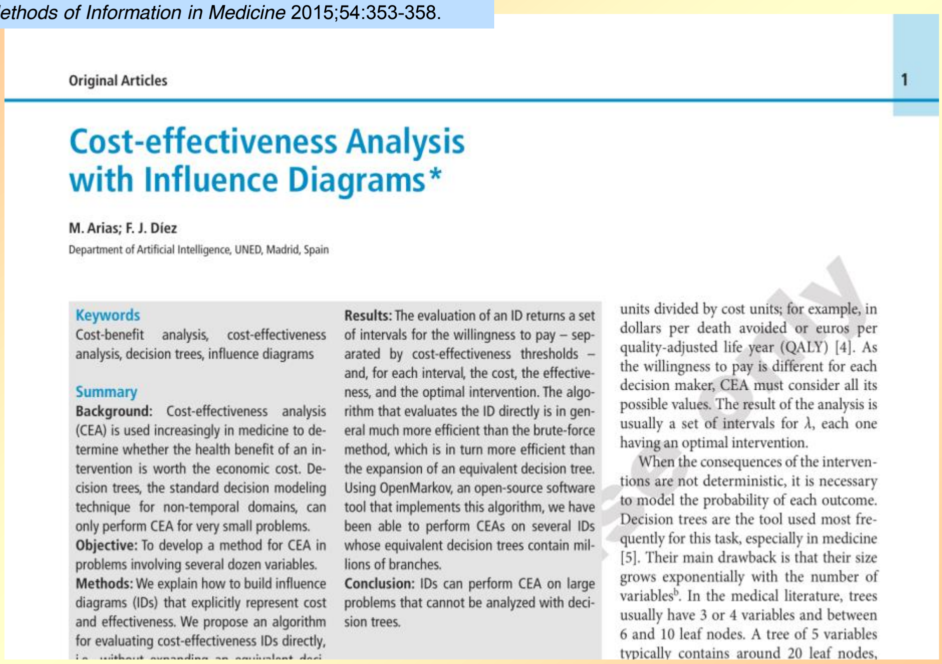
Additional information

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

CISIAD. Research Center on Intelligent Decision-Support Systems. UNED. Madrid, Spain.

5.3.3. CEA with IDs and DANs

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


Original Articles 1

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,
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,

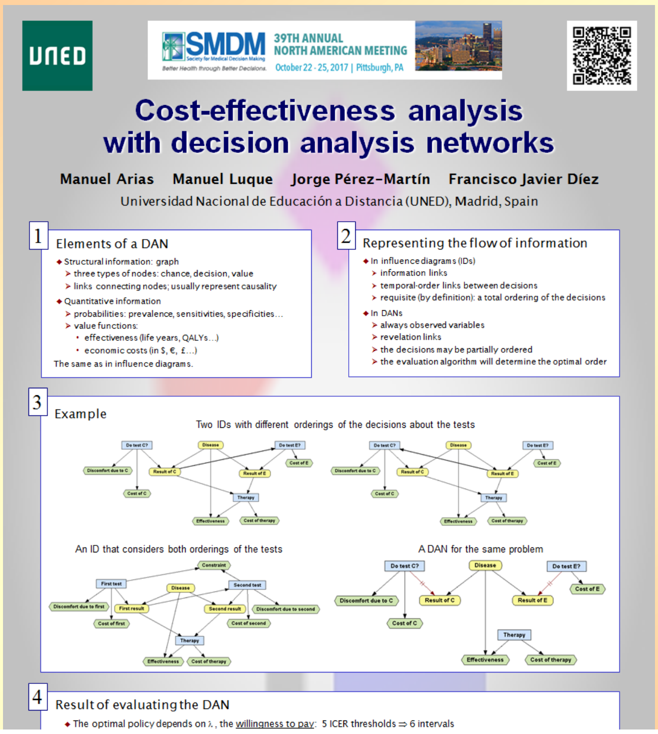
Example: Optimal stratety for two tests

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

- Question: What is the most beneficial strategy?
- This problem cannot be solved with IDs because the order of the tests is not specified.



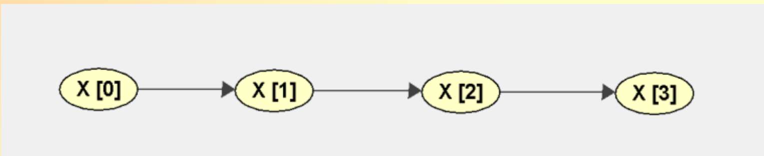
6. Temporal models

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.

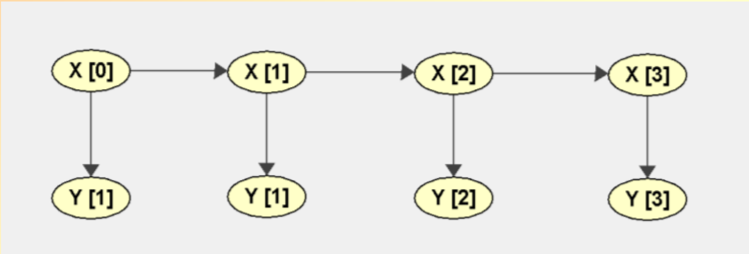
6.1. Types of Markov models

Markov chain



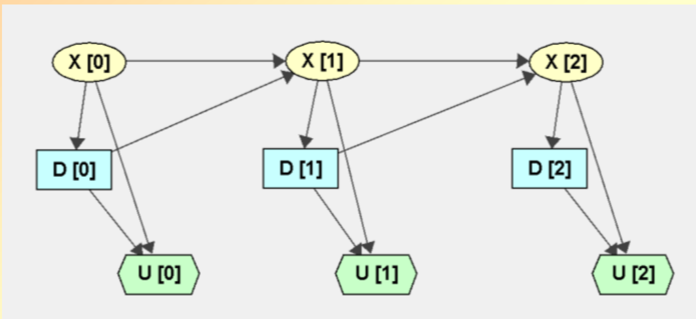
- One variable that evolves over time
- Transition probabilities: $P(x_{i+1}|x_i)$

Hidden Markov model (HMM)



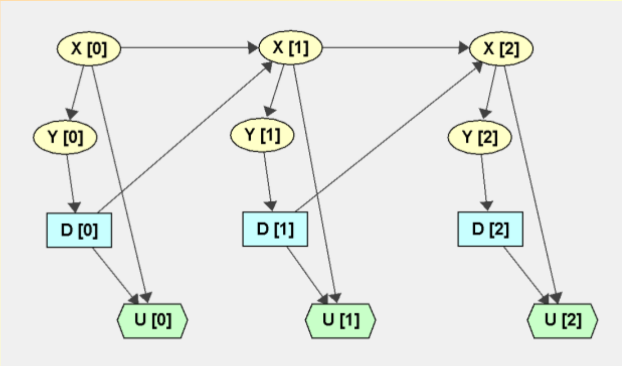
- Observed variable: Y
- Non-observed (hidden) variable: X
- Transition probabilities: $P(x_{i+1}|x_i)$
- Probability of each observation: $P(y_i|x_i)$

Markov decision process (MDP)



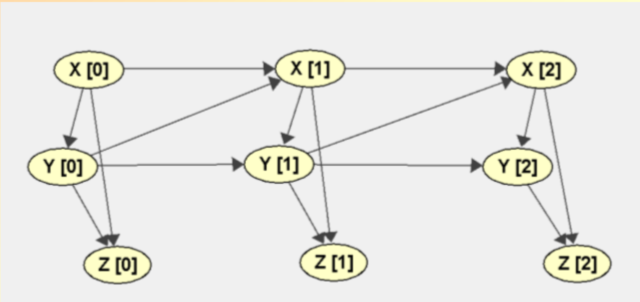
- Observed variable: X
- Decision: D
- Transition probabilities: $P(x_{i+1}|x_i)$
- Reward: $U(x_i, d_i)$

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)

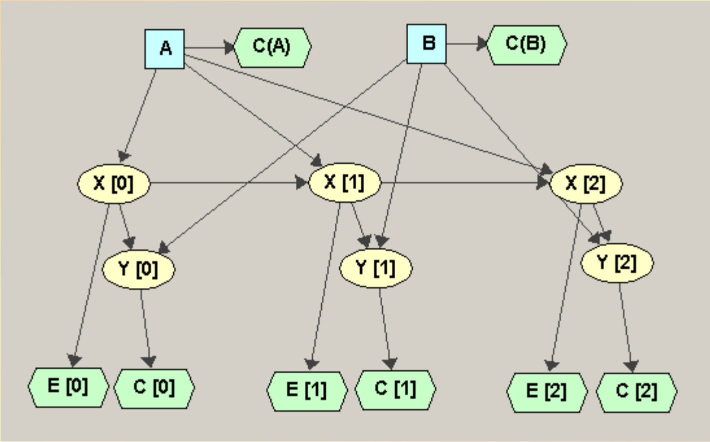


- Markov chain or hidden Markov model:
  - one variable, X
  - one conditional probability: $P(x_{i+1}|x_i)$
- Dynamic Bayesian network:
  - several variables, $\{X, Y, Z\ldots\}$
  - factored probability: $P(y_i|x_i)$, $P(z_i|x_i, y_i)$, $P(x_{i+1}|x_i, y_i)\ldots$

Factored extensions of Markov models

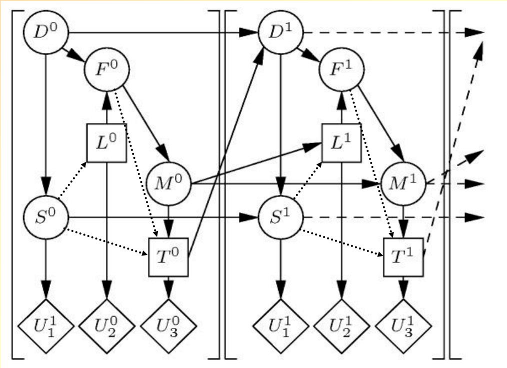
Flat model	Factored model
Markov chain	Dynamic Bayesian network [Dean and Kanazawa, 1989]
Hidden Markov model	
Markov decision process (MDP)	Factored MDP [Boutilier et al., 1995, 2000]
Partially-observable MDP (POMDP)	Factored POMDP [Boutilier and Poole, 1996]

Markov influence diagrams



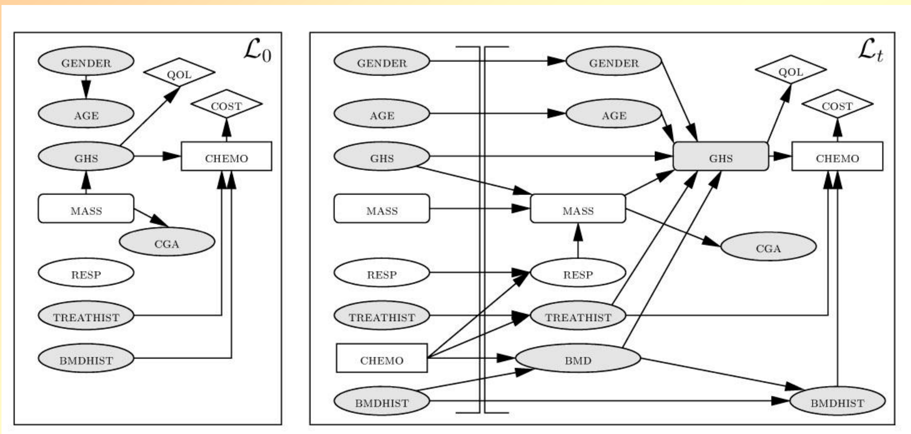
- Can be used for cost-effectiveness analysis

Dynamic limited-memory IDs (DLIMIDs)



- Differences wrt POMDPs
  - Several decisions in each time slice.
  - Limited memory: the decision maker only knows the observations made in the current time slice and in the previous one
  - 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, Open-Markov, 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

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

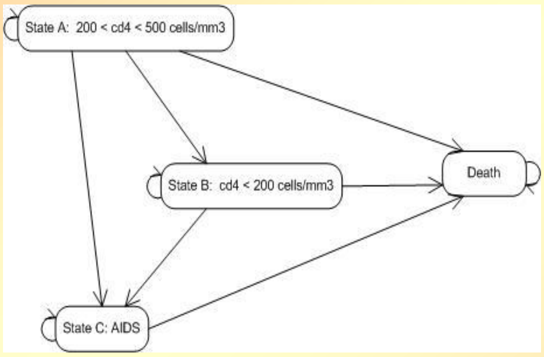
Pharmacoeconomics 1997 Jul 12 (1) 54-66
1170-7690/97/0007-0054/$06.50/0

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

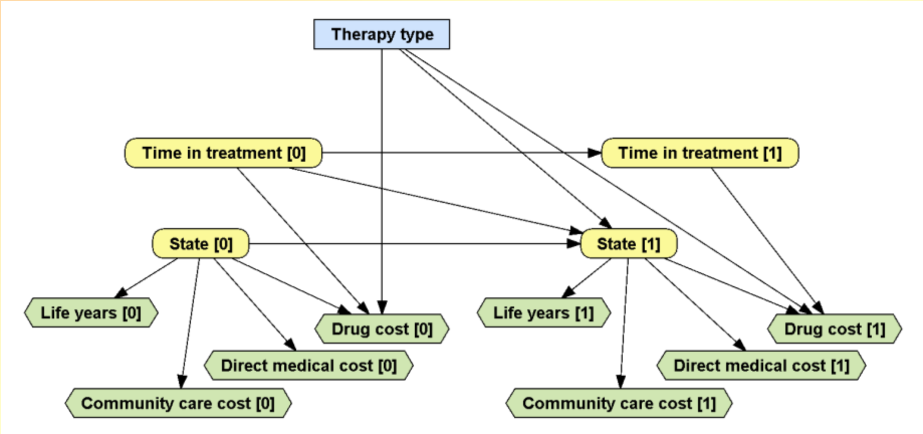
- State-transition diagram: 4 states



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

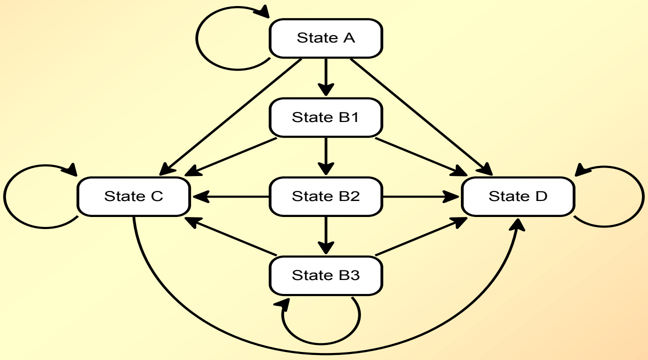
A MID version of the HIV model

[Chancellor et al., 1997]



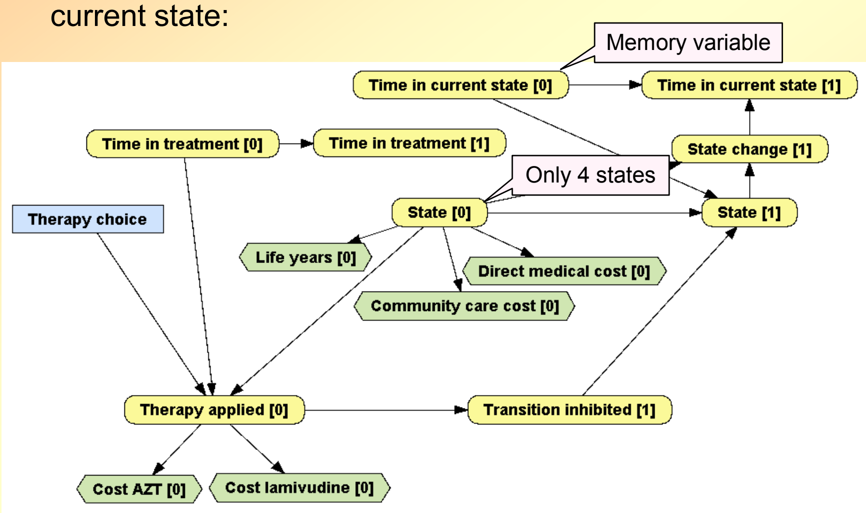
Representing the patient history (1)

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



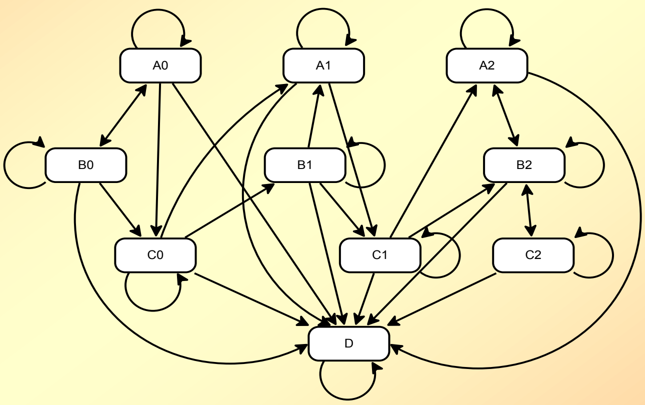
Representing the patient history (1)

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



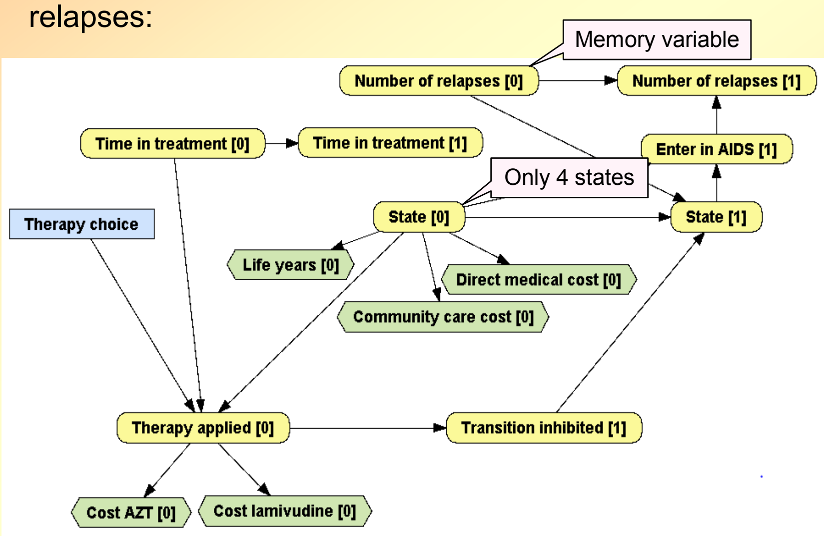
Representing the patient history (2)

- 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

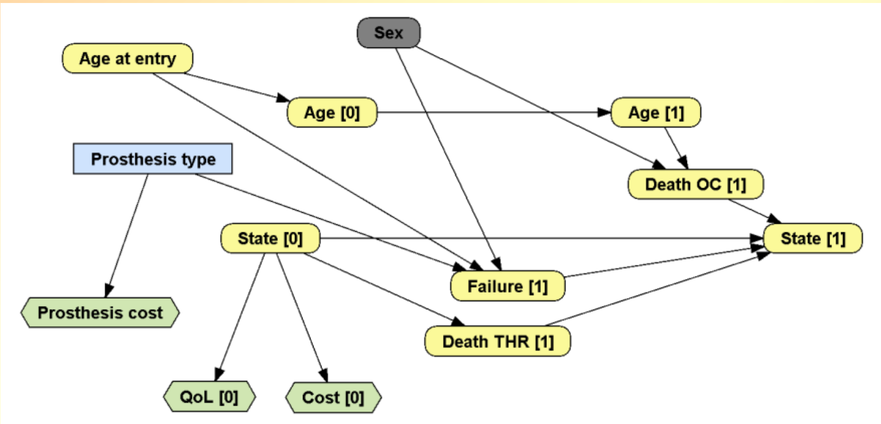
Appl Health Econ Health Policy 2004; 3 (2): 79-89
1175-5652/04/0002-0079/$31.00/0

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

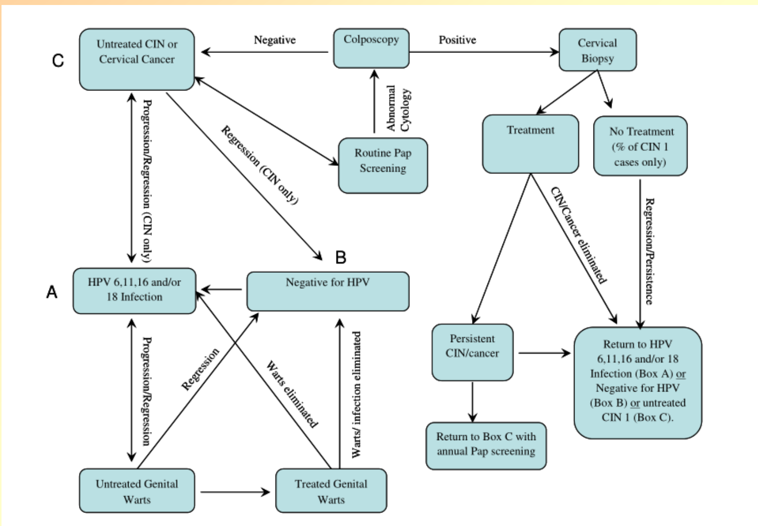
* Corresponding author

Published: 29 July 2009

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

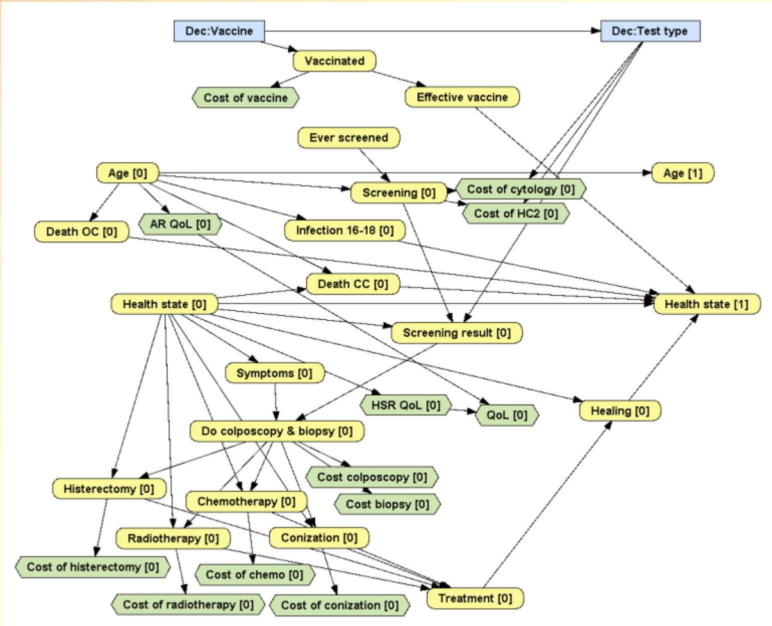
Received: 13 May 2008
Accepted: 29 July 2009

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



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

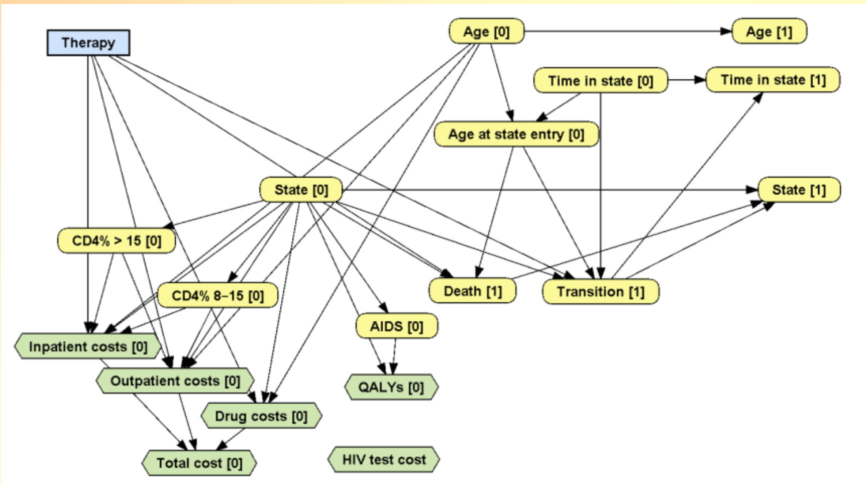
Máirín Ryan[a], Susan Griffin[b], Bona Chitah[c], A. Sarah Walker[d], Veronica Mulenga[e], Donald Kalolo[e], Neil Hawkins[b], Concepta Merry[a], Michael G. Barry[a], Chifumbe Chintu[e], Mark J. Sculpher[b] and Diana M. Gibb[d]

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.

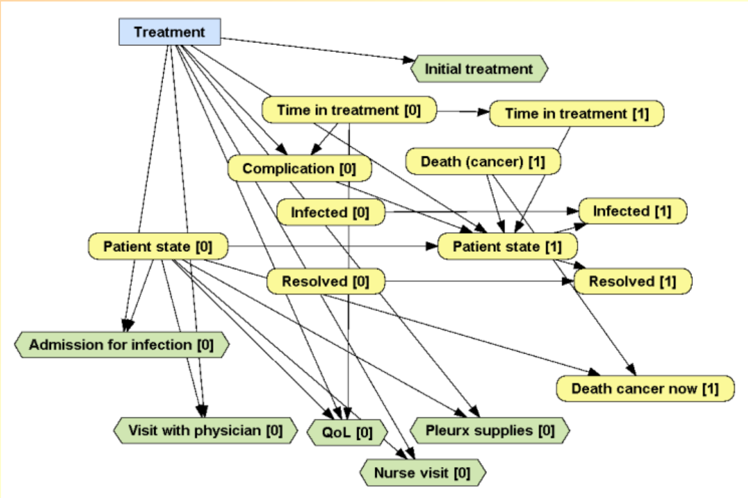
Design: A probabilistic decision analytical model of HIV/AIDS progression in children based on the CD4 cell percentage (CD4%) was populated with data from the placebo-controlled Children with HIV Antibiotic Prophylaxis trial that had reported a 43% reduction in mortality with cotrimoxazole prophylaxis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

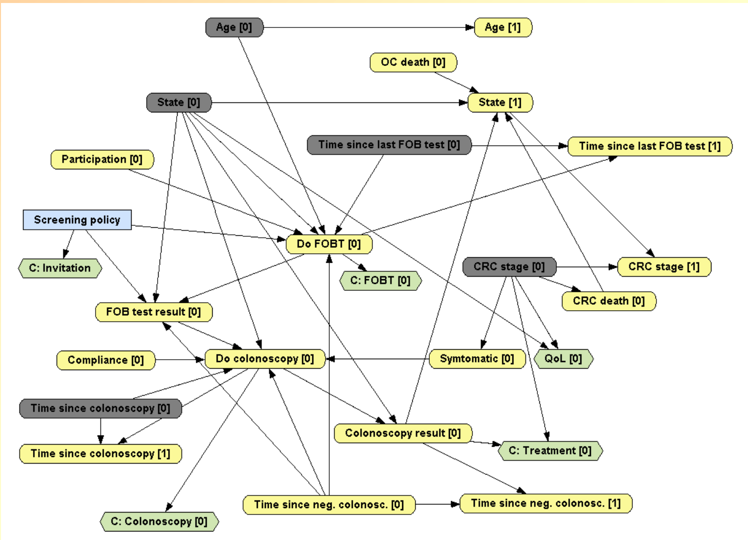


Our model for malignant pleural effusion



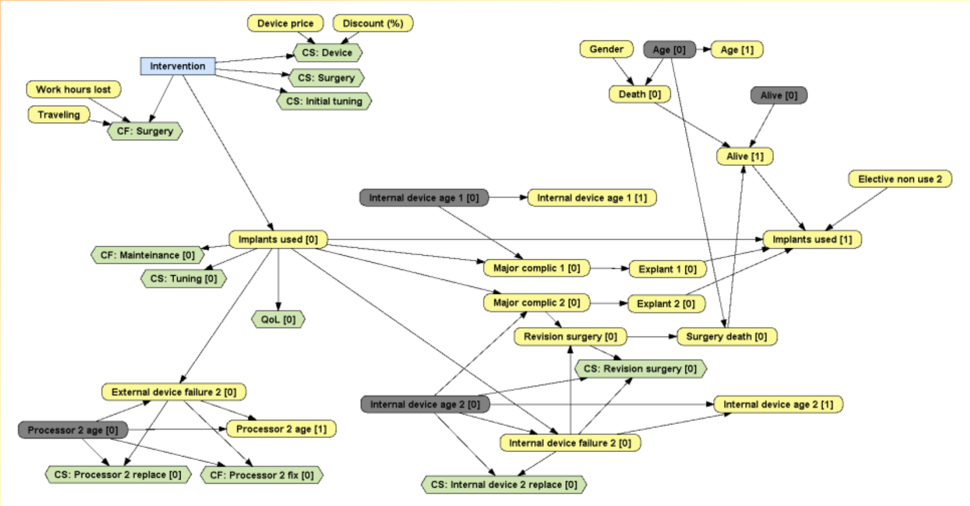
➢ *Meeting of the Society for Medical Decision Making (SMDM 2015)*, St. Louis, October 2015.

Our model for colorectal cancer screening



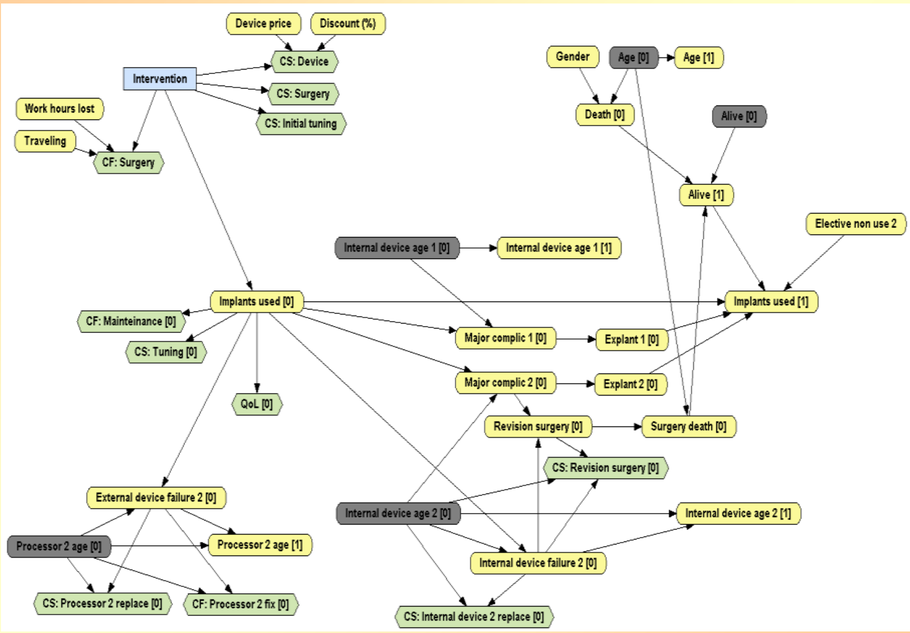
➢ *European Conference of the Society for Medical Decision Making*, London, UK, June 2015.

Our model for bilateral cochlear implantation

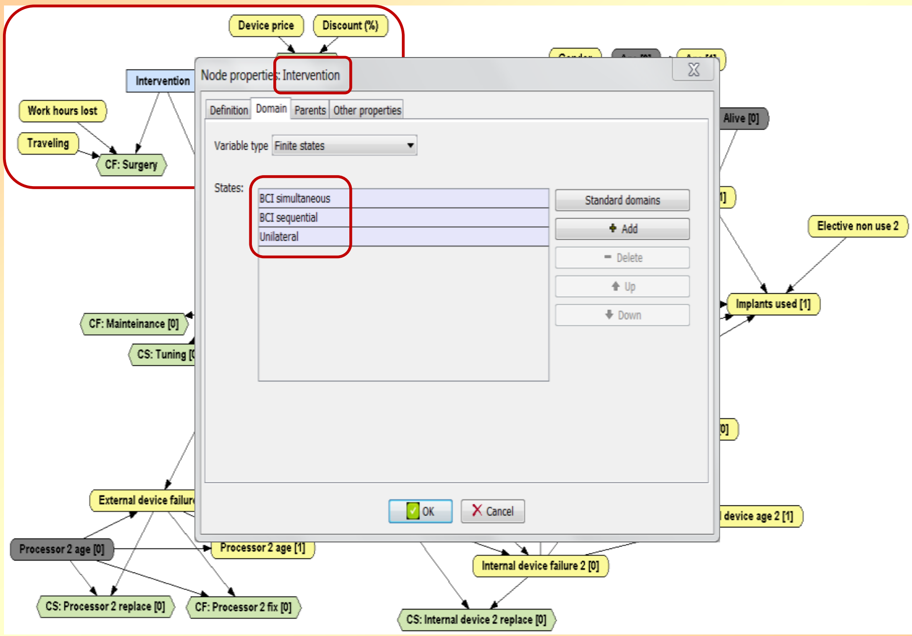


- *Cochlear Implant Symposium*, Washington DC, October 2015.

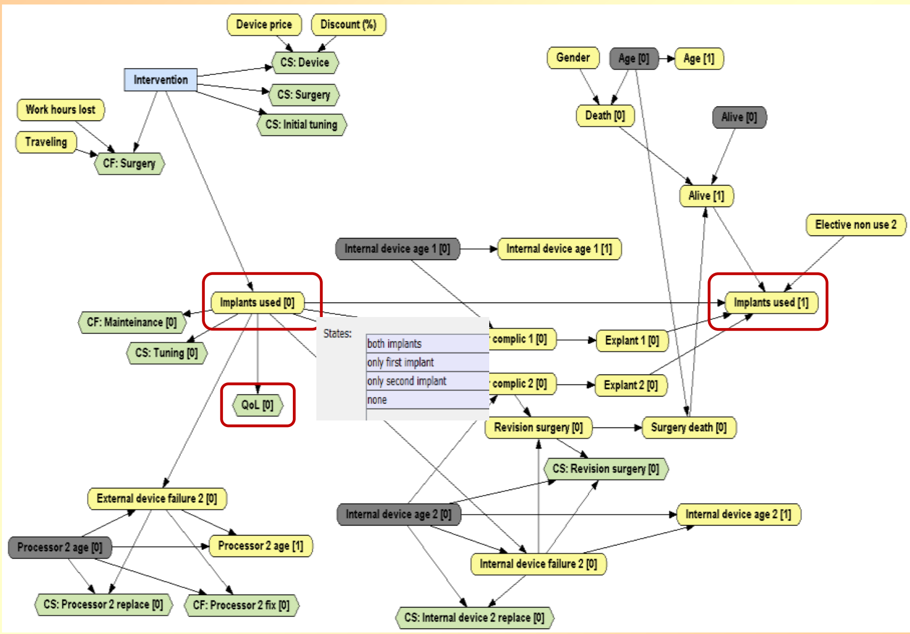
Our model for bilateral cochlear implantation



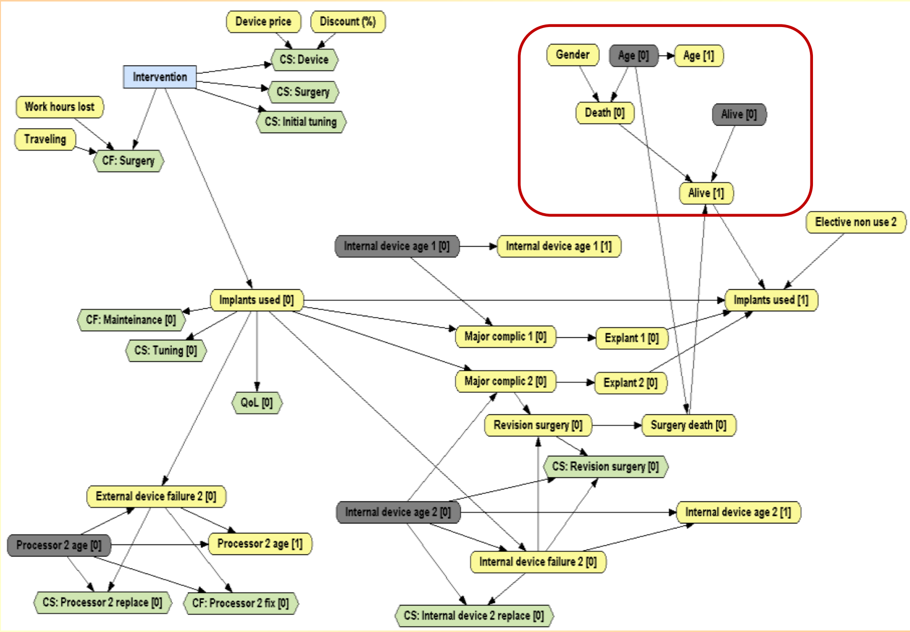
Our model for bilateral cochlear implantation



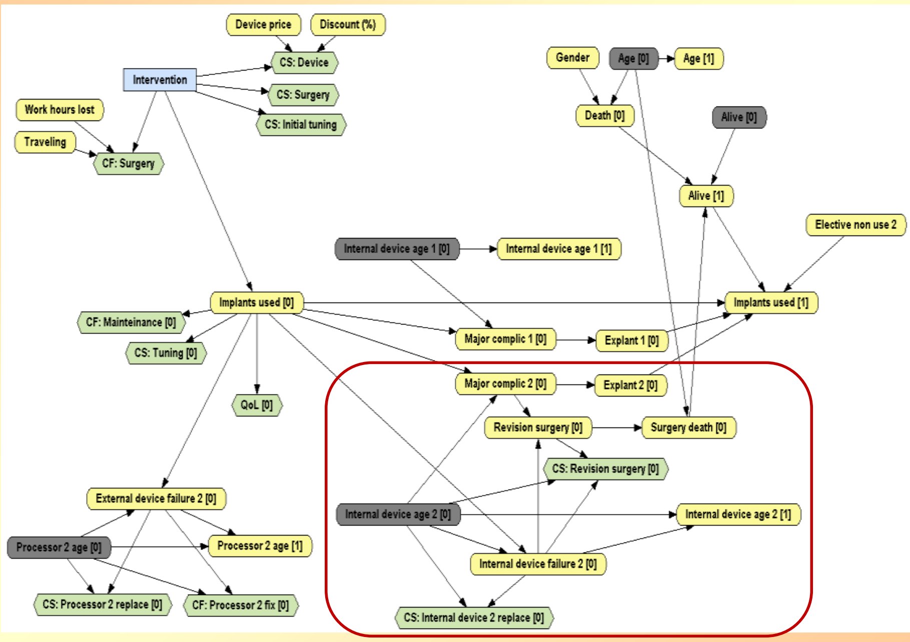
Our model for bilateral cochlear implantation



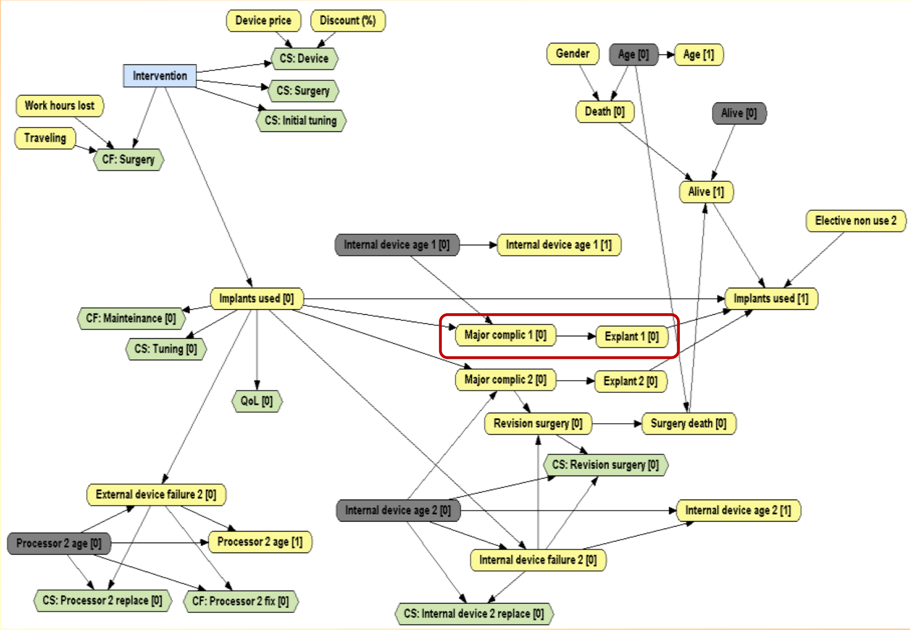
Our model for bilateral cochlear implantation



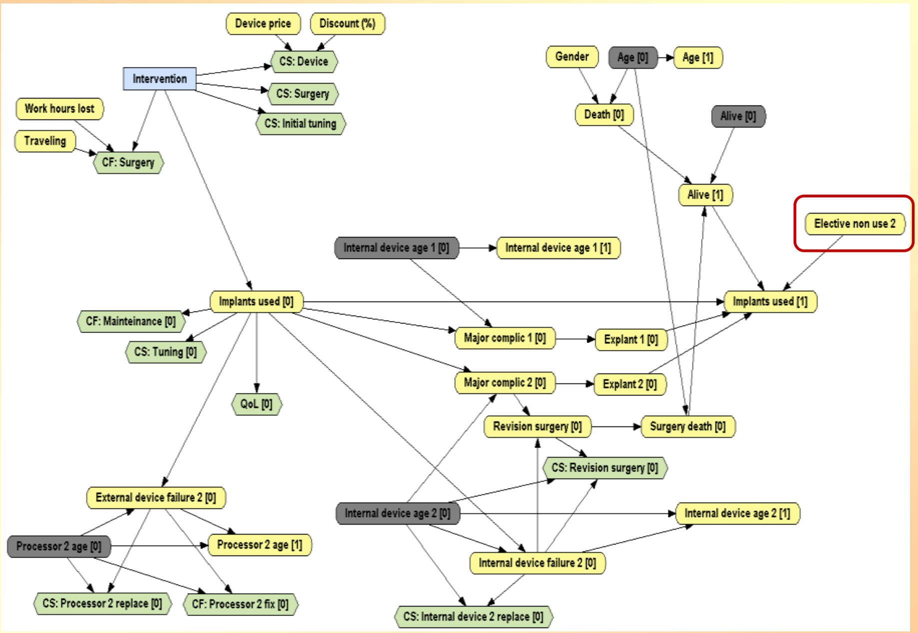
Our model for bilateral cochlear implantation



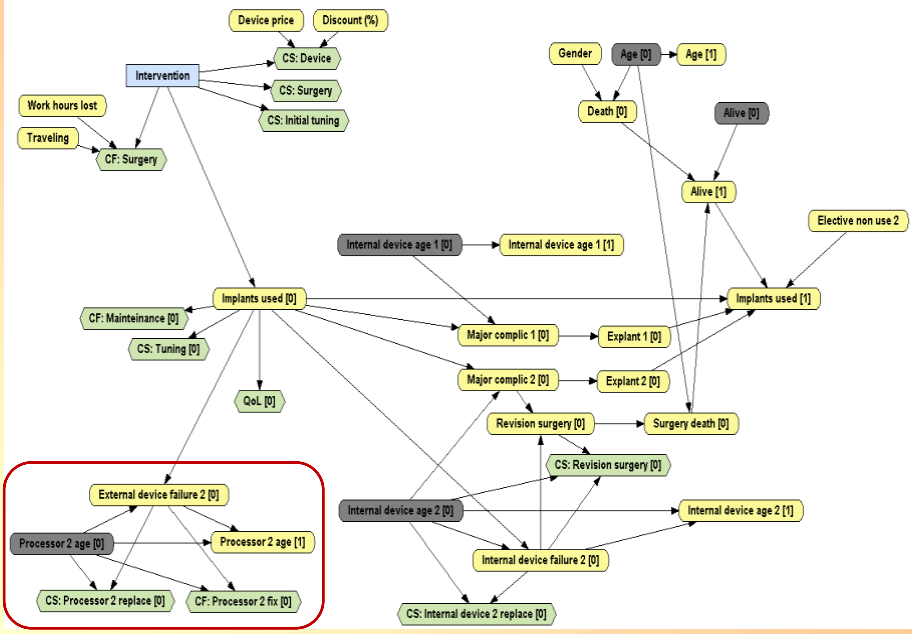
Our model for bilateral cochlear implantation



Our model for bilateral cochlear implantation

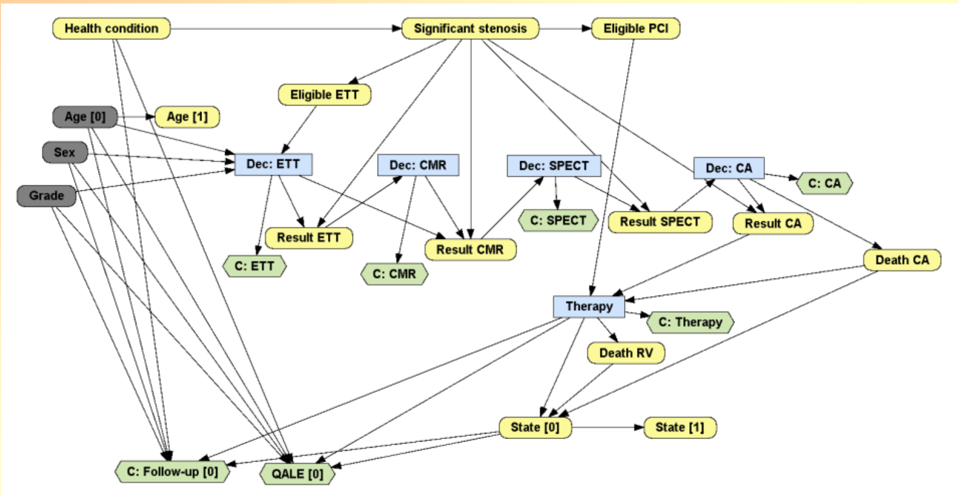


Our model for bilateral cochlear implantation



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.

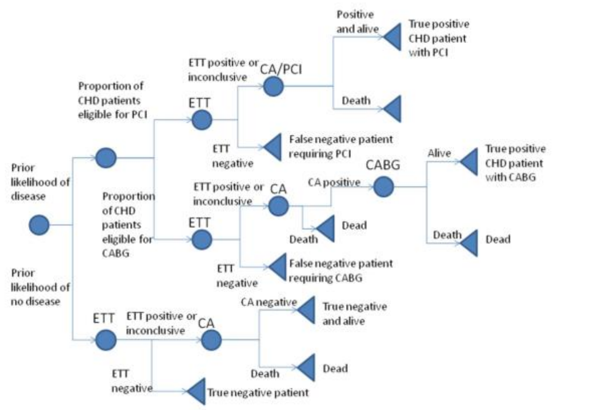


Figure 1 Structure of decision tree using Strategy 2 as an example. CA, coronary angiography; CABG, coronary artery bypass grafting; CHD, coronary heart disease; ETT, exercise treadmill testing; PCI, percutaneous coronary intervention.

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.

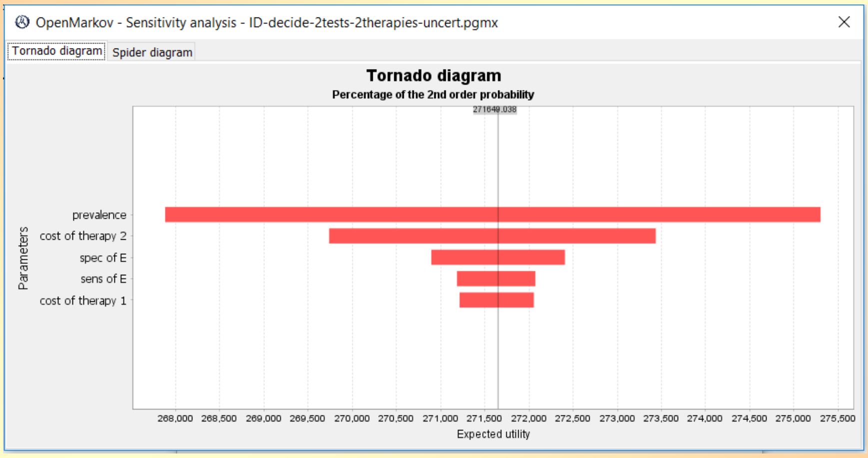
7. Sensitivity analysis

Types of sensitivity analysis

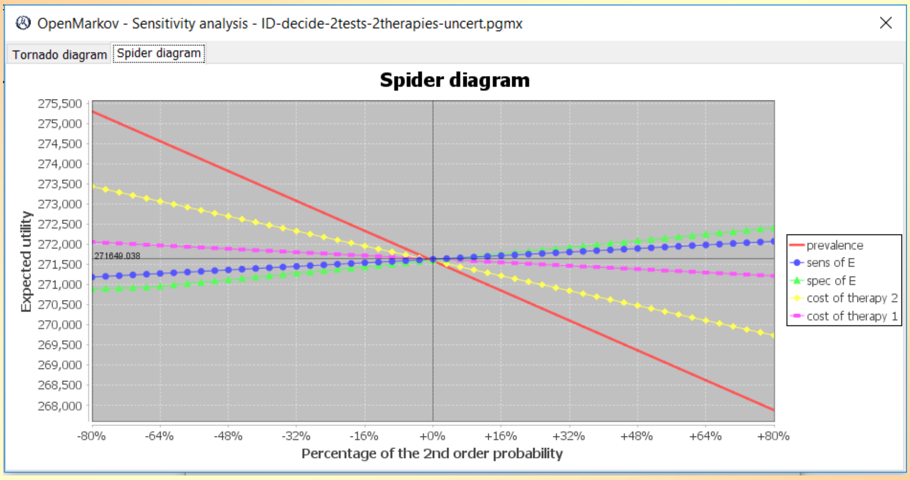
- Two main types
  - structural (qualitative)
  - parametric (quantitative)
- Depending on the effect analyzed
  - analysis of utility
  - analysis of decisions / policies
- Depending on how many parameters are varied
  - one-way analysis
  - *n*-way analysis (independent or join analysis)
- Depending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

7.1. Unicriterion sensitivity analysis

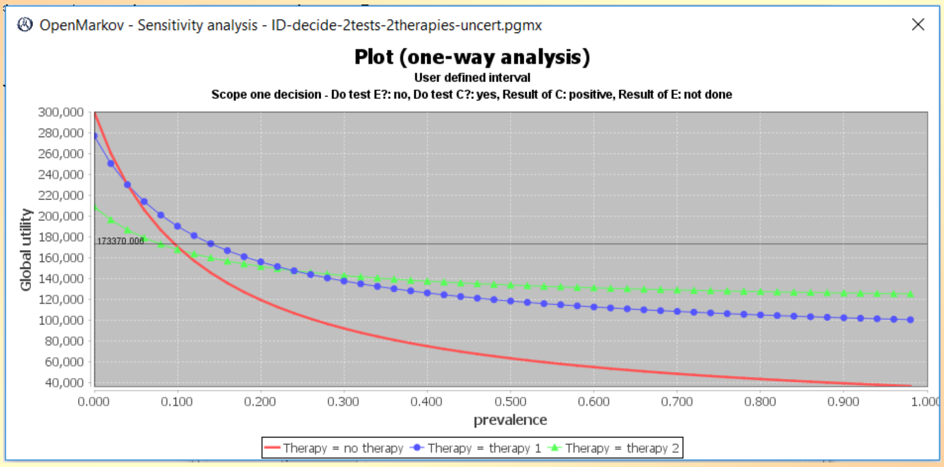
Tornado diagram



Spider diagram

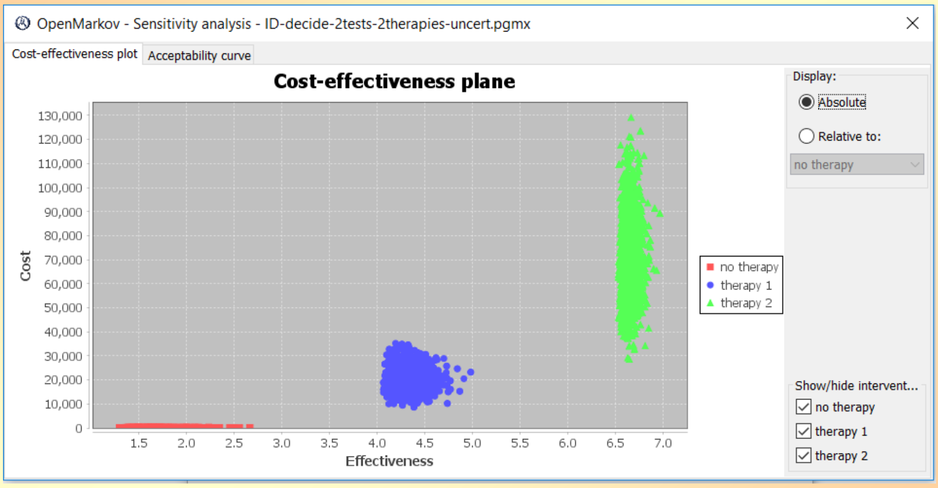


Plot (one-way sensitivity analysis)



7.2. Cost-effectiveness sensitivity analysis

Scatter plot

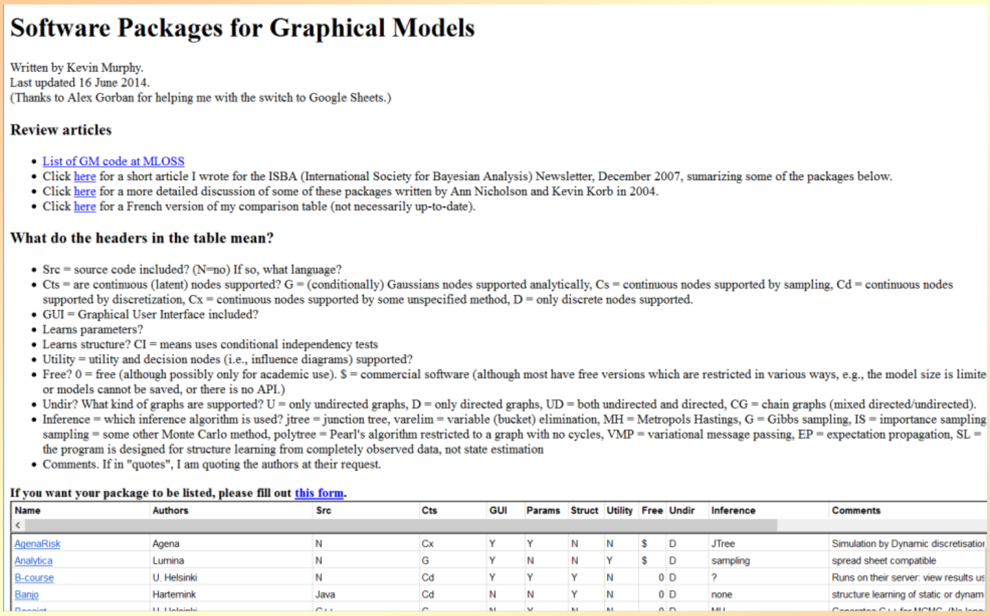


Acceptability curve

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 discretisatior
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

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, especially partially ordered decisions.
- 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…

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.

How to bring PGMs from artificial intelligence into medical decision making

- Develop powerful user-friendly software tools
- 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…
  - Discrete event simulation with PGMs

Thank you very much for your attention!

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