LECTURE 5:ANALYSIS METHODS, DECISION TREES, PERFORMANCE OF DIAGNOSTIC TESTS

- This lecture focuses on determining the appropriate clinical strategy for patients

- The analysis method involves building a decision tree, evaluating the tree and performing sensitivity analysis

DECISION ANALYSIS

Decision Analysis is a quantitative, probabilistic method for modelling problems under situations of uncertainty

We make a decision when we irreversibly allocate resources

Typical steps we follow:

- Gather information
- Assess consequences of the alternatives
- Take an action

The goal is to clarify the dynamics and tradeoffs involved in selecting one strategy from a set of alternatives

MEDICAL DECISION ANALYSIS

For DA to be an appropriate tool in medical decision making the following two should be satisfied

There should be **uncertainty** about the appropriate clinical strategy for patients

If there is already a single optimal strategy, no need for an analysis

There should be a **tradeoff**

Two or more alternatives are compared

If one clearly dominates the other (better in terms of all the relevant aspects such as adverse disease outcome, risk of side effects and utilities) no need for an analysis

Unlike the cost effectiveness studies discussed before ,we may only focus on the clinical outcome aspect

MEDICAL DECISION ANALYSIS

Decision Analysts

- Deliberatively seek out new creative alternatives
- Identify possible outcomes
- Identify relevant uncertain factors
- Encode probabilities for the uncertain factors
- Specify the value placed on each outcome (cost, QALYs, LE)
- Formally analyze the decision

Decision tree

Hand-out

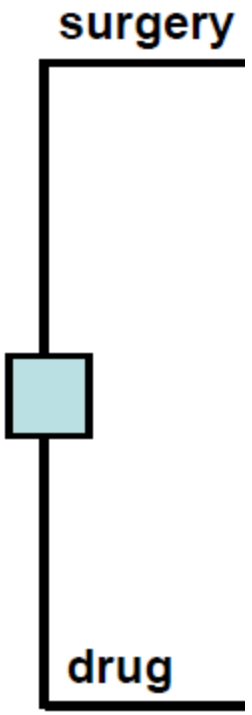
Elements of Decision Analysis

- Structure the Problem
  - Identify decision alternatives
  - List possible clinical outcomes
  - Represent sequence of events
- Assign probabilities to all chance events
- Assign utility or value, to all outcomes
- Evaluate the expected utility of each strategy
- Perform sensitivity analysis

Elements of a Decision Tree

Decision Node: A point in a decision tree at which several choices are possible. The DM controls what is chosen

The choices should be mutually exclusive



Mutually exclusive:

The intersection of the events is empty

One and (only one) of the events must occur

Elements of a Decision Tree

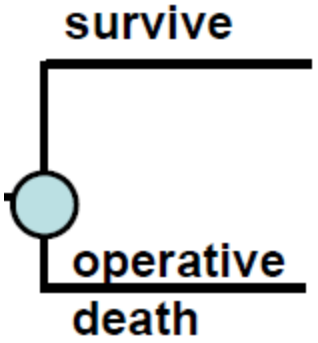
Chance Node: A point in decision tree at which chance determines which outcome will occur

The outcomes should be mutually exclusive and collectively exhaustive

Collectively exhaustive

The set of events take up the entire outcome space

At least one of the events must occur



- Terminal Node: Final outcome state associated with each possible path way

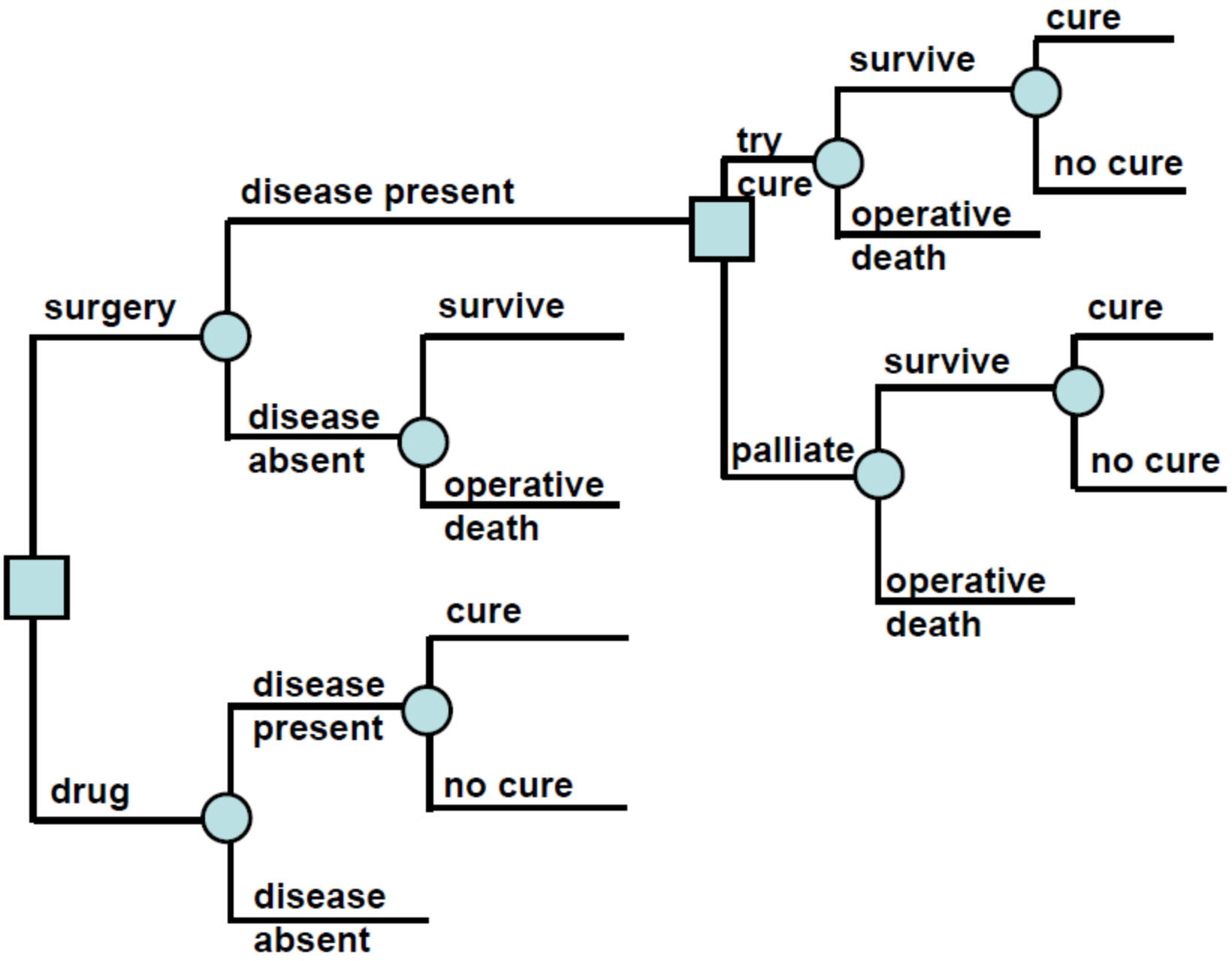
Example

A symptomatic patient

- Operate (risky)
- Medical management (drug therapy)

If disease present at surgery, must decide whether try for cure or palliate (to ease symptoms without curing the underlying disease)

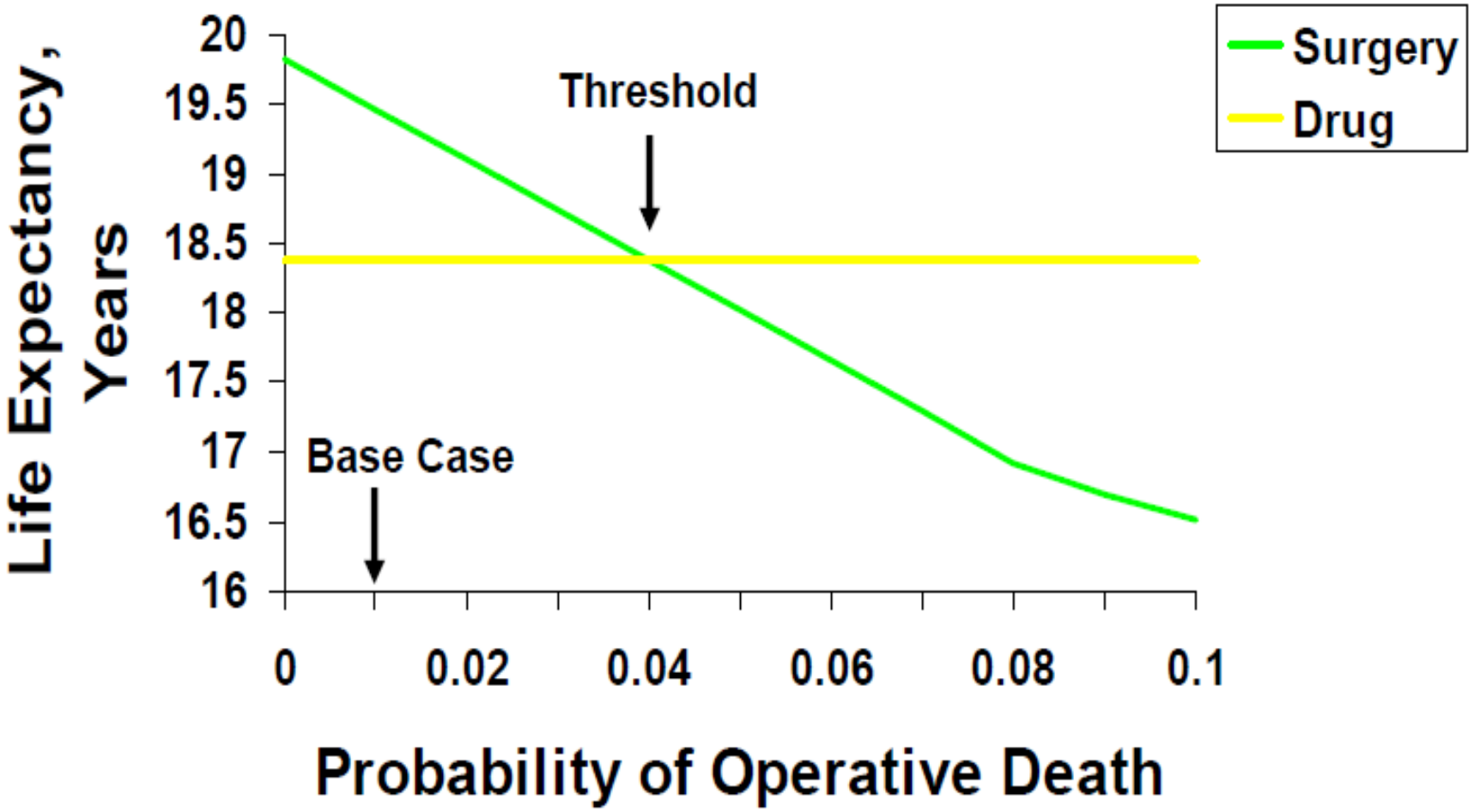
We want to evaluate surgery versus medical management


cure
survive
no cure
try
cure
operative
death
disease present
surgery
survive
disease
absent
operative
death
cure
survive
palliate
no cure
operative
death
cure
disease
present
no cure
drug
disease
absent

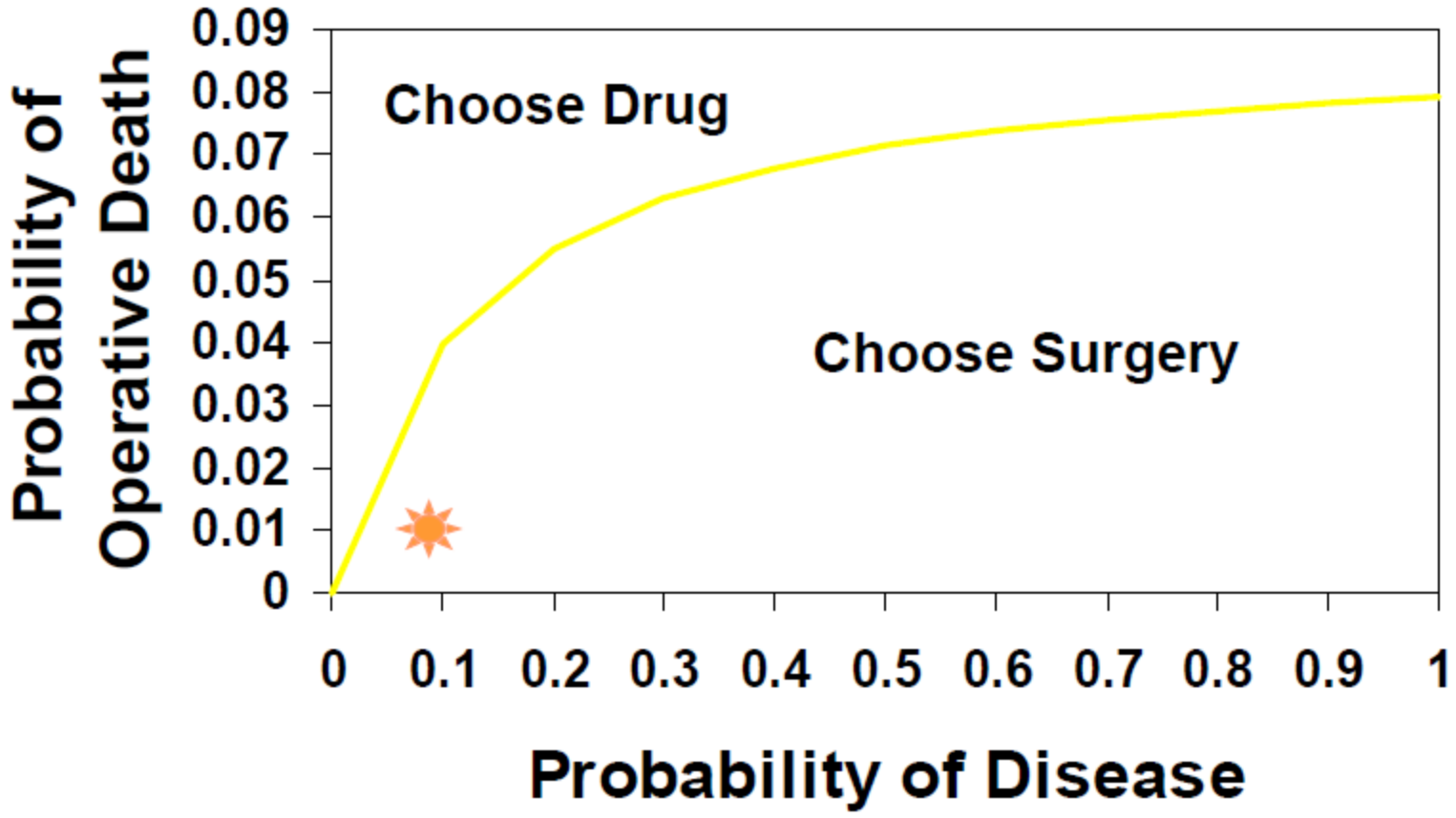
Sensitivity Analysis

- Systematically asking what of question to see how the decision result changes
- Determines how robust the decision is
  - One parameter varied (Threshold analysis)
  - Multiple parameters systematically varied (Multi-way analysis)

One-way Sensitivity Analysis



Two-way Sensitivity Analysis



Examples of Decision Trees

Measles Example

TESTS

Tests

Tests are used in medical diagnosis, screening and research

How well is a subject classified into disease or non-disease group?

- Ideally, all subjects who have the disease should be classified as “having the disease” and vice versa

Practically, the ability to classify individuals into the correct disease status depends on the accuracy of the tests, among other things

Tests

- The primary purpose of screening tests is to detect early disease or risk factors for disease in large numbers of apparently healthy individuals.

- The purpose of a diagnostic test is to establish the presence (or absence) of disease when subject shows signs or symptoms of a disease

Examples of Screening Tests

- Blood pressure for hypertension
- Mammography for breast cancer
- Fasting blood sugar for diabetes

Examples of Diagnostic Tests

- X-rays, biopsies, pregnancy tests, medical histories

BAYES THEOREM ON DIAGNOSTIC TEST

Hand-written notes

Diagnostic Tests

Prevalence or pretest probability = probability of disease in the relevant population (i.e. 5% of patients presenting with cough have pneumonia)

Post-test probability = probability of disease given a positive or negative test. Also called the "conditional probability", as in “Probability of disease conditional on the patient having a positive test.“

Gold Standard

We must know the correct disease status of the patient prior to calculation to assess performance of a test

Gold standard test: is the best test available to obtain a definitive diagnosis for a particular disease

It may be biopsy, surgery, autopsy or an acknowledged standard.

Gold Standards are used to define true disease status against which the results of a new diagnostic test are compared.

Gold Standard

Target Disorder	Gold Standard
breast cancer	excisional biopsy
prostate cancer	transrectal biopsy
coronary stenosis	coronary angiography
myocardial infarction	catheterization
strep throat	throat culture

- A new test: a new screening test or a less expensive diagnostic test

- Use a 2 x 2 table to compare the performance of the new test

Contingency table

True positives (TP): those that have the disease and test positive
False positives (FP): those that do not have the disease and test positive
False negatives (FN): those that do have the disease and test negative
True negatives (TN): those that do not have the disease and test negative)

A 2x2 contingency table

		Gold Standard (Reference Standard) (True Condition)		
		Diseased	Healthy	TOTAL
Diagnosis or Screening Test	Positive (+)	TP	FP	
	Negative (-)	FN	TN	
	TOTAL			

Performance of diagnostic tests

		Gold Standard (Reference Standard) (True Condition)		
		Diseased	Healthy	TOTAL
Diagnosis or Screening Test	Positive (+)	TP	FP	
	Negative (-)	FN	TN	
	TOTAL			

Sensitivity : The ability of the test to identify correctly those who have the disease among patients with disease, the probability of a positive test

The probability that an individual with the disease of interest has a positive test

TP/(TP+FN)

Specificity : The ability of the test to identify correctly those who do not have the disease among patients without disease (i.e. healthy patients),

The probability that an individual without the disease has of a negative test

TN/FP+TN

Performance of diagnostic tests

		Gold Standard (Reference Standard) (True Condition)		
		Diseased	Healthy	TOTAL
Test	Positive (+)	TP	FP	
	Negative (-)	FN	TN	
	TOTAL			

Positive predictive value (PPV) is the proportion of patients with a positive test that actually have disease

$$TP/(TP+FP)$$

The PPV describes how likely it is that a positive test result in a given patient population represents a true positive.

Negative predictive value (NPV) is the proportion of patients with a negative test result that are actually disease free

$$TN/(TN+FN)$$

The NPV describes how likely it is that a negative test result in a given patient population represents a true negative.

Performance of diagnostic tests

LR+: The ratio of the likelihood of a positive test result occurring in patients with disease (true positive) to the likelihood of a positive test result in patients without disease (false positive)

$$\frac{TP/(TP+FN)}{FP/(TN+FP)}$$

OR

$$Sens/(1-Spec)$$

LR-: The ratio of the likelihood of a negative test result in patients with disease (false negative) to the likelihood of a negative test result in patients without disease (true negative)

$$\frac{FN/(TP+FN)}{TN/(TN+FP)}$$

OR

$$(1-Sens)/Spec$$

		Gold Standard (Reference Standard) (True Condition)		
		Diseased	Healthy	TOTAL
Diagnosis or Screening Test	Positive (+)	TP	FP	FP+TP
	Negative (-)	FN	TN	TN+FN
	TOTAL	TP+FN	FP+TN	

		Gold Standard (Reference Standard) (True Condition)		
		Diseased	Healthy	TOTAL
Diagnosis or Screening Test	Positive (+)	TP	FP	FP+TP
	Negative (-)	FN	TN	TN+FN
	TOTAL	TP+FN	FP+TN	

Post-test probability = probability of disease given a positive or negative test. Also called the "conditional probability", as in "Probability of disease conditional on the patient having a positive test.“

P(Disease|Positive test (+))=P(Disease and Positive)/P(Positive) =TP/FP+TP

		Gold Standard (Reference Standard) (True Condition)		
		Diseased	Healthy	TOTAL
Diagnosis or Screening Test	Positive (+)	TP	FP	FP+TP
	Negative (-)	FN	TN	TN+FN
	TOTAL	TP+FN	FP+TN	

Note that one can also construct a contingency table in terms of percentages

	Diseased	**Healthy**	**TOTAL**
Positive (+)	8	400	408
Negative (-)	2	9590	9592
TOTAL	10	9990	10000

	Diseased	**Healthy**	**TOTAL**
Positive (+)	0.0008(=8/10000)	0.040	0.0408
Negative (-)	0.0002	0.959	0.9592
TOTAL	0.001	0.999	1

Example

	Diseased	Healthy	TOTAL
Positive (+)	8	400	408
Negative (-)	2	9590	9592
TOTAL	10	9990	10000

P(Positive)=

P(Negative)=

P(Diseased)=

P(Healthy)=

Example

	Diseased	Healthy	TOTAL
Positive (+)	8	400	408
Negative (-)	2	9590	9592
TOTAL	10	9990	10000

P(+ and Diseased)=

P(- and Diseased)=

P(+ and Healthy)=

P(- and Healthy)

Example

	Diseased	Healthy	TOTAL
Positive (+)	8	400	408
Negative (-)	2	9590	9592
TOTAL	10	9990	10000

P(+|Diseased)=

P(-|Diseased)=

P(+|Healthy)=

P(-|Healthy)=

Example

	Diseased	Healthy	TOTAL
Positive (+)	8	400	408
Negative (-)	2	9590	9592
TOTAL	10	9990	10000

P(Diseased|+)=

P(Diseased|-)=

P(Healthy|+)=

P(Healthy|-)=

Example

	Diseased	Healthy	TOTAL
Positive (+)	8	400	408
Negative (-)	2	9590	9592
TOTAL	10	9990	10000

Positive predictive value

Negative predictive value

Sensitivity=

Specificity=

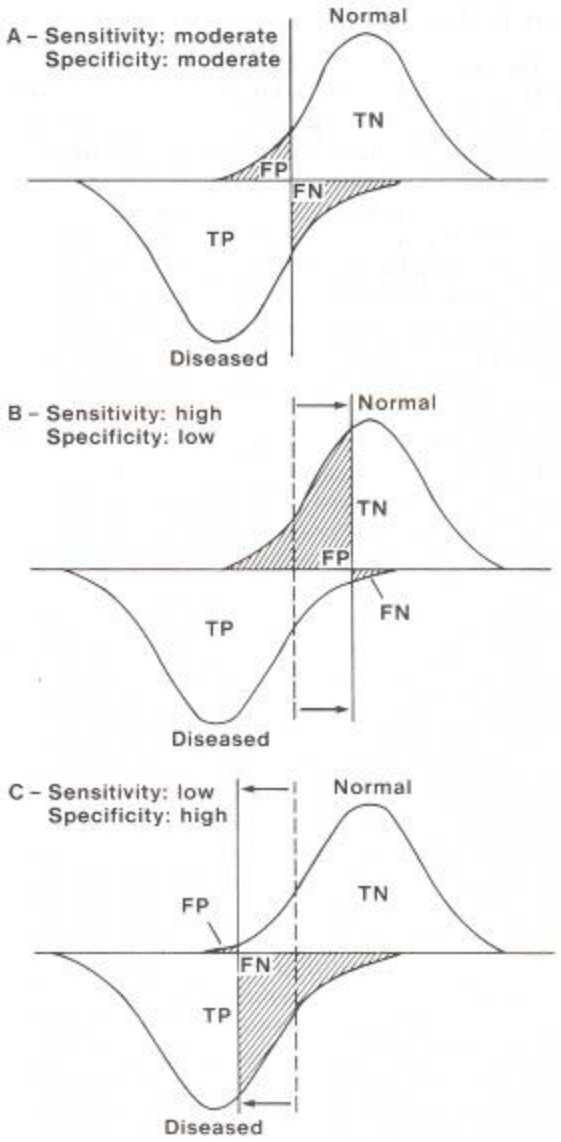
The Cutoff point

If we define the "cutoff point": The value below which the test result is said to be abnormal

Then a certain percentage of normal individuals will be declared abnormal (the false positives) and a certain percentage of abnormal individuals will be declared normal (the false negatives).

The sensitivity and specificity of a quantitative test are dependent on the cut-off value above or below which the test is positive.

In general, the higher the sensitivity, the lower the specificity, and vice versa.



Kaplan C. Use of the Laboratory. In: Walker HK, Hall WD, Hurst JW, editors. Clinical Methods: The History, Physical, and Laboratory Examinations. 3rd edition. Boston: Butterworths; 1990. Chapter 5. Available from: http://www.ncbi.nlm.nih.gov/books/NBK372/

A special type of diabetes disease can be diagnosed using oral glucose tolerance test

The subject drinks a glucose solution and blood is drawn at intervals for glucose measurement

Screening test is fasting plasma glucose

- Easier, faster, more convenient and less expensive

ROC curves

- **Receiver operating characteristic (ROC) curves:**

A graphic method depicting then trade-off between the true positive rate and the false positive rate of a test

True positive rate is on the vertical axis (y-axis) and the false positive rate on the horizontal axis (x-axis)

The greater the area under the ROC curve, the better the test discriminates between patients with or without disease.

ROC curves allow tests to be compared over a variety of cutoff points.

Table 1: Pooled estimates of the sensitivity, specificity, positive and negative likelihood ratios and diagnostic odds ratios of the brief Patient Health Questionnaire (PHQ-9) for diagnosing major depressive disorder, by cut-off score

Cut-off score	No. of studies	No. of patients	Sensitivity (95% CI)	Specificity (95% CI)	Positive likelihood ratio (95% CI)	Negative likelihood ratio (95% CI)	Diagnostic odds ratio (95% CI)
7	5	2794	0.83 (0.70–0.91)	0.73 (0.63–0.82)	3.20 (2.21–4.63)	0.22 (0.12–0.41)	14.18 (6.21–32. 33)
8	6	3306	0.82 (0.66–0.92)	0.83 (0.69–0.92)	5.17 (2.30–11.58)	0.20 (0.08–0.47)	25.46 (5.34–121.38)
9	7	3269	0.83 (0.68–0.92)	0.86 (0.76–0.92)	6.07 (3.35–11.01)	0.18 (0.08–0.39)	32.23 (10.06–103. 25)
10	16	5782	0.85 (0.75–0.91)	0.89 (0.83–0.92)	7.83 (5.22–11.74)	0.16 (0.09–0.27)	47.50 (22.94–98.35)
11	10	3451	0.89 (0.75–0.96)	0.89 (0.79–0.94)	8.43 (4.29–16.58)	0.11 (0.04–0.29)	75.03 (25.29–222.53)
12	10	3451	0.77 (0.60–0.88)	0.91 (0.84–0.95)	8.86 (5.65–13.90)	0.24 (0.13 - 0.44)	36.15 (20.16–64.85)
13	8	2759	0.81 (0.75–0.85)	0.92 (0.86–0.95)	10.19 (6.19 - 16.77)	0.20 (0.15–0.26)	50.03 (29.81–83.97)
14	6	2162	0.67 (0.57–0.76)	0.95 (0.90–0.97)	14.23 (7.10 - 28.52)	0.33 (0.24–0.45)	42.13 (18.73–94.74)
15	6	2482	0.62 (0.48–0.74)	0.96 (0.94–0.97)	18.57 (11.37-30.33)	0.39 (0.27–0.55)	47.60 (23.09–98.14)

Note: CI = confidence interval.

In the example, Test A performs better than Test B over all ranges.

ROC curves also assist in the selection of the cutoff point designed to maximize a test's utility.

If a test is designed to confirm a disease, a cutoff point with greater specificity and lower sensitivity is selected.

If a test is designed to screen for occult disease, a cutoff point with greater sensitivity and lower specificity is selected.

ROC curve

To sum up,

- The ROC curve shows the trade-off between sensitivity and specificity (any increase in sensitivity will be accompanied by a decrease in specificity and vice-versa)

- The closer the curve follows the left-hand border and then the top border of the ROC space, the more accurate the test.

- The closer the curve comes to the 45-degree diagonal of the ROC space, the less accurate the test.

- The area under the curve (AUROC) can be used to assess test accuracy, and to compare the performance of different tests.

Origin of ROC

'ROC analysis is part of a field called "Signal Detection Theory" developed during World War II for the analysis of radar images. Radar operators had to decide whether a blip on the screen represented an enemy target, a friendly ship, or just noise. Signal detection theory measures the ability of radar receiver operators to make these important distinctions.

http://www.healthknowledge.org.uk/public-health-textbook/disease-causation-diagnostic/2c-diagnosis-screening/statistical-aspects-screening called the 'Receiver Operating Characteristics'.

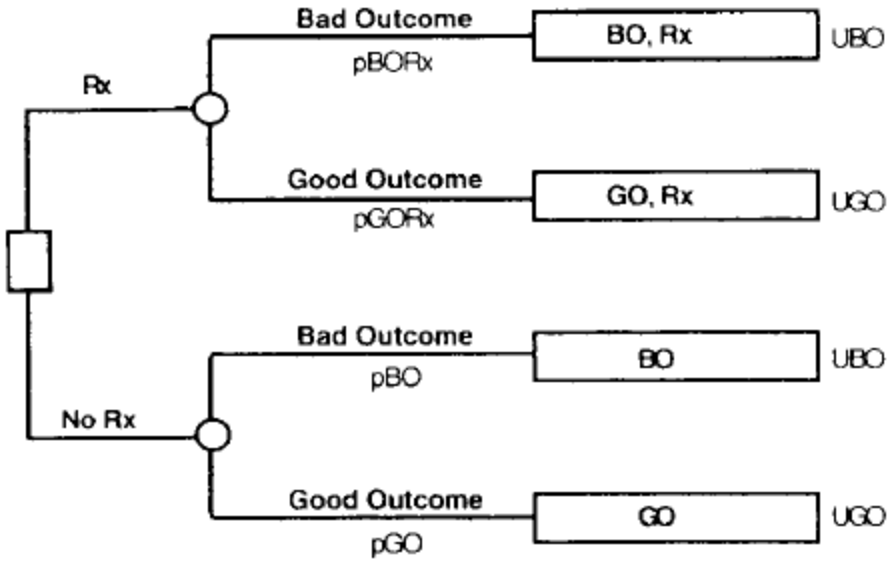
KEY POINTS IN BUILDING TREES

KEY POINTS IN BUILDING TREES

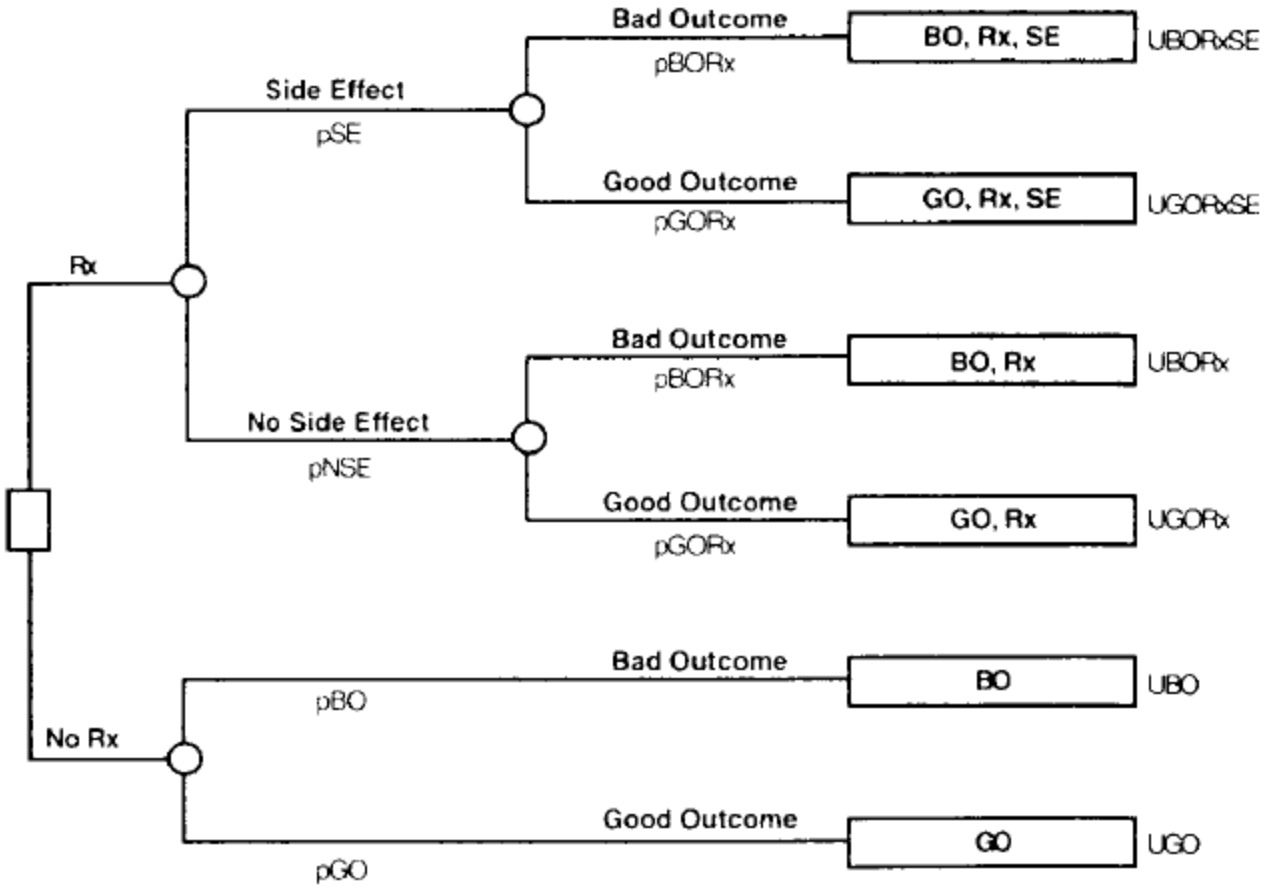
- The choice of management strategies for patients with giant cell arteritis (GCA)
- GCA can lead to a very severe complication of blindness
- Steroids decrease the risk of blindness but there is a risk of side-effects (hypertension, fluid retention, avascular necrosis of bone)
- A test for GCA, biopsy, however the sensitivity of the test is not ideal

1. The tree must have balance

- There should be a tradeoff (between risks and benefits)
- If one strategy is dominating the other, the tree is not balanced
- Two strategies: Treating GCA vs not treating
- Suppose that pBORx<pBO and the treatment has no down side (e.g. Risk of a side effect)
- Then treatment (Rx) dominates no treatment
- No tradeoff, no balance

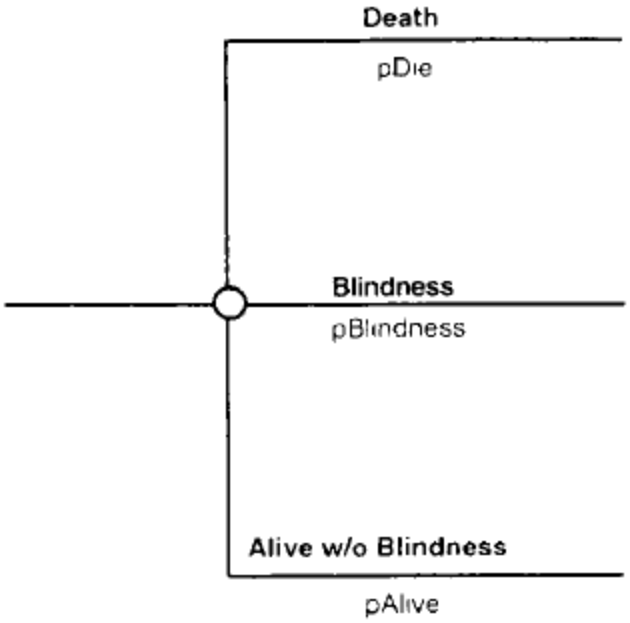


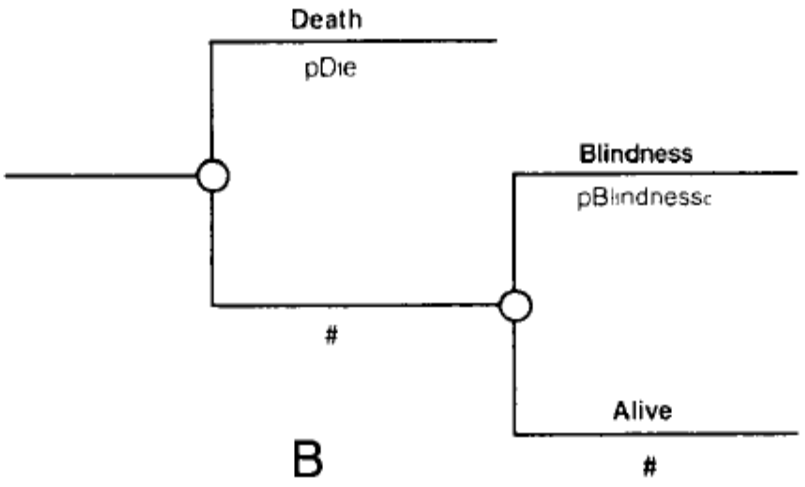
1. The tree must have balance



There is a side effect associated with the treatment option
Each branch has advantages and disadvantages

2.Only two branches after each chance node





Sensitivity analysis (SA)

	Baseline values
pDie	5%
pBlindness	80%
pAlive (=1-pDie-pBlindness)	15%

Each chance node followed by two branches

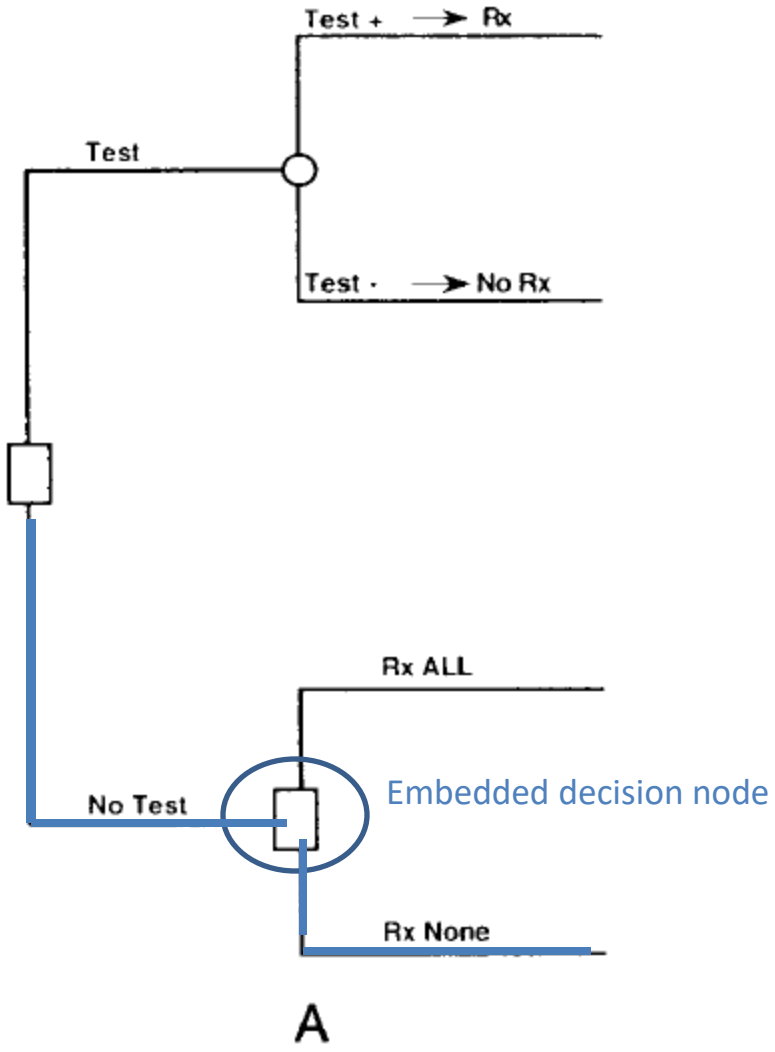
PBlindnessc≠80% (now conditional on being alive)

PBlindnessc=P(Blind|Alive)=0.8/0.95

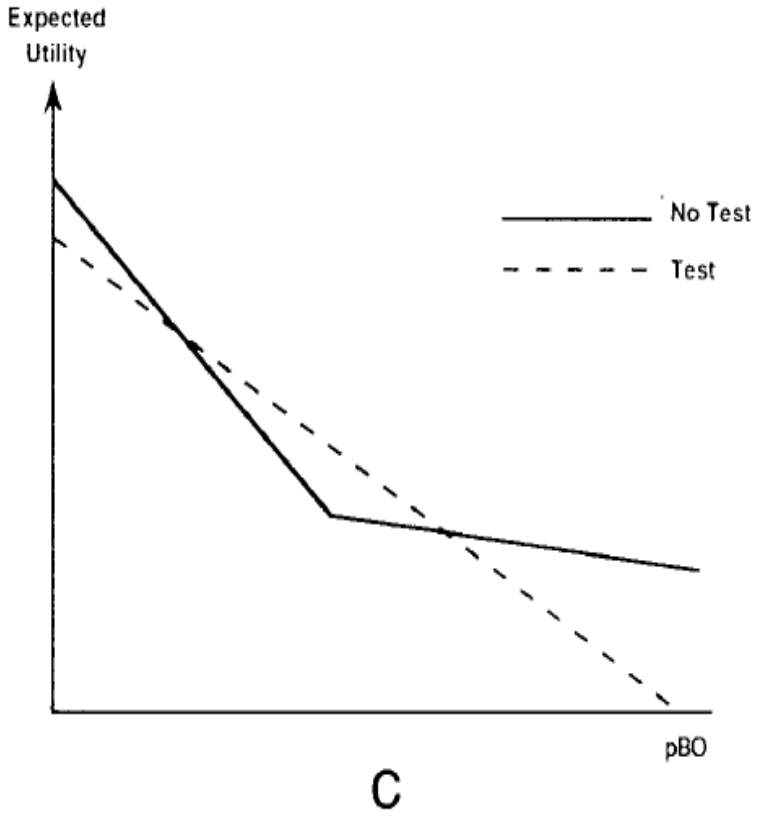
Suppose pDie varies between 0% and 40%

Only 0% and 20% will be considered by the program!

3. Avoid embedded decision nodes

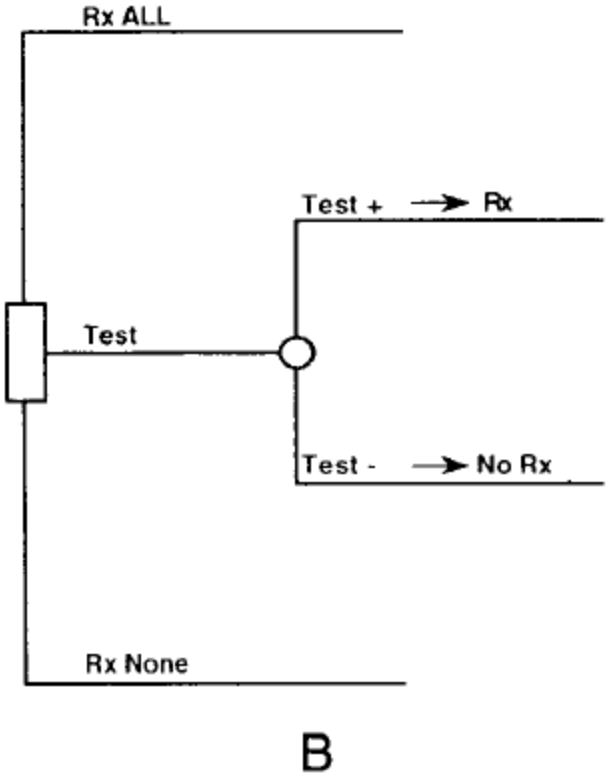


Embedded decision node

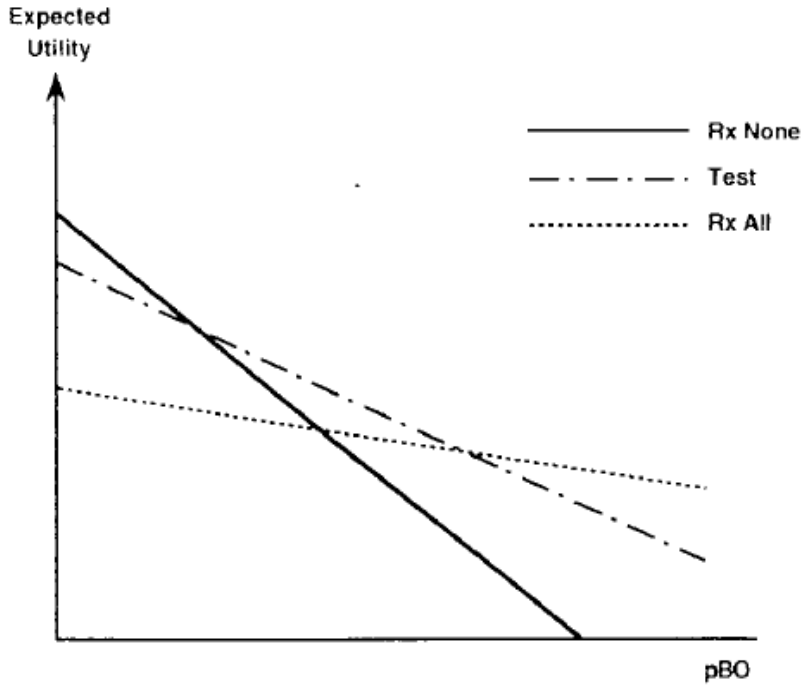


Difficulty in SA

3. Avoid embedded decision node

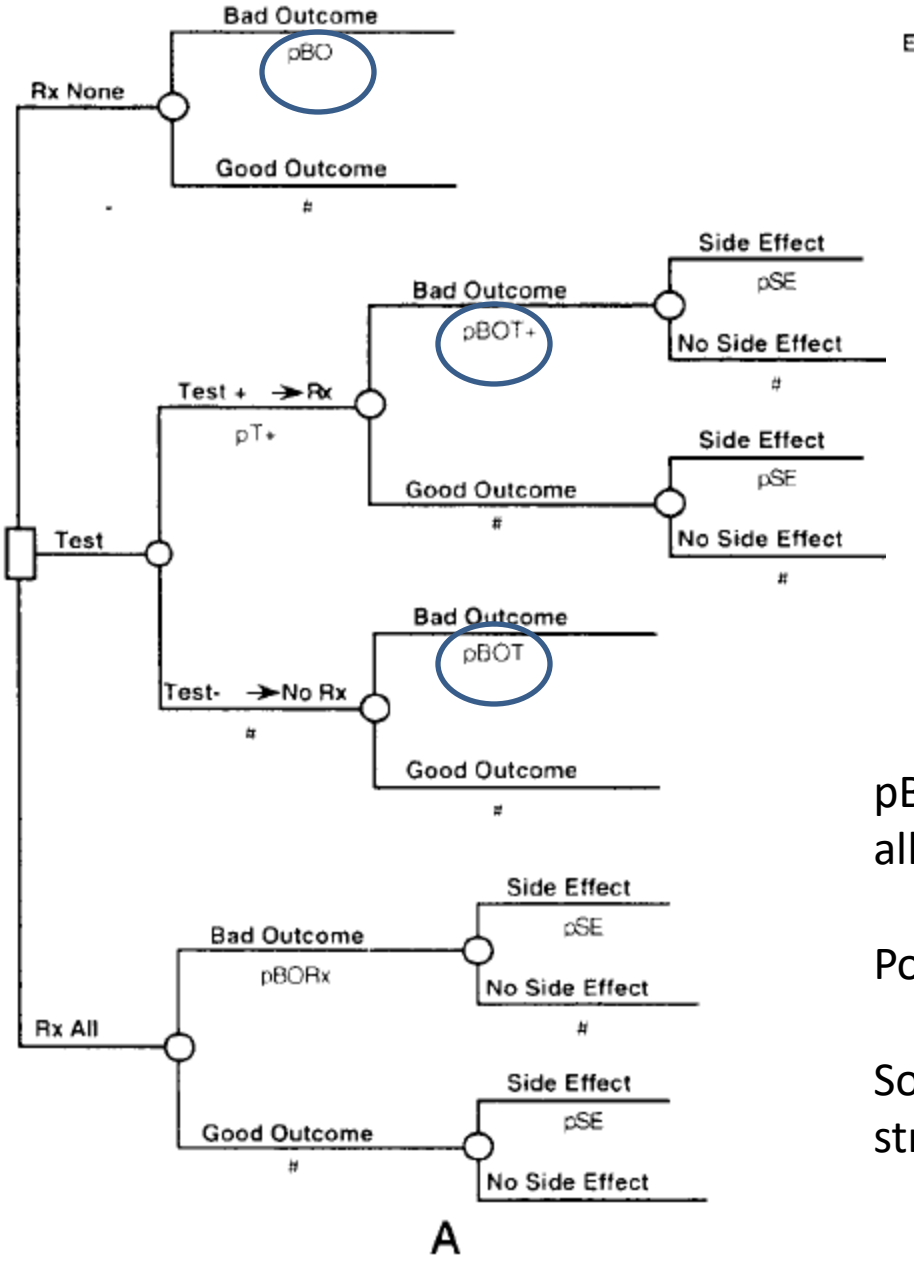


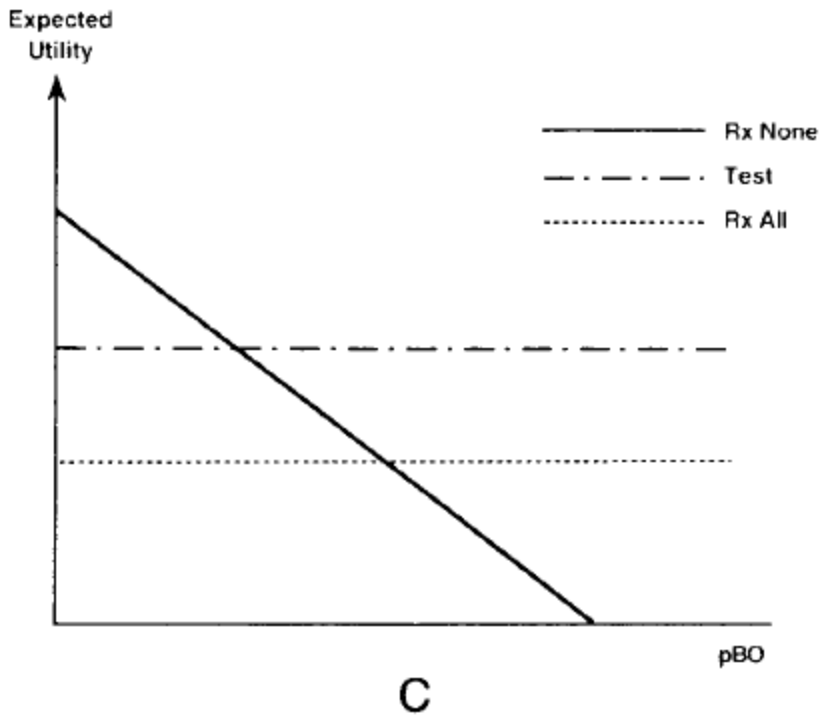
B



D

4. Branches must be linked

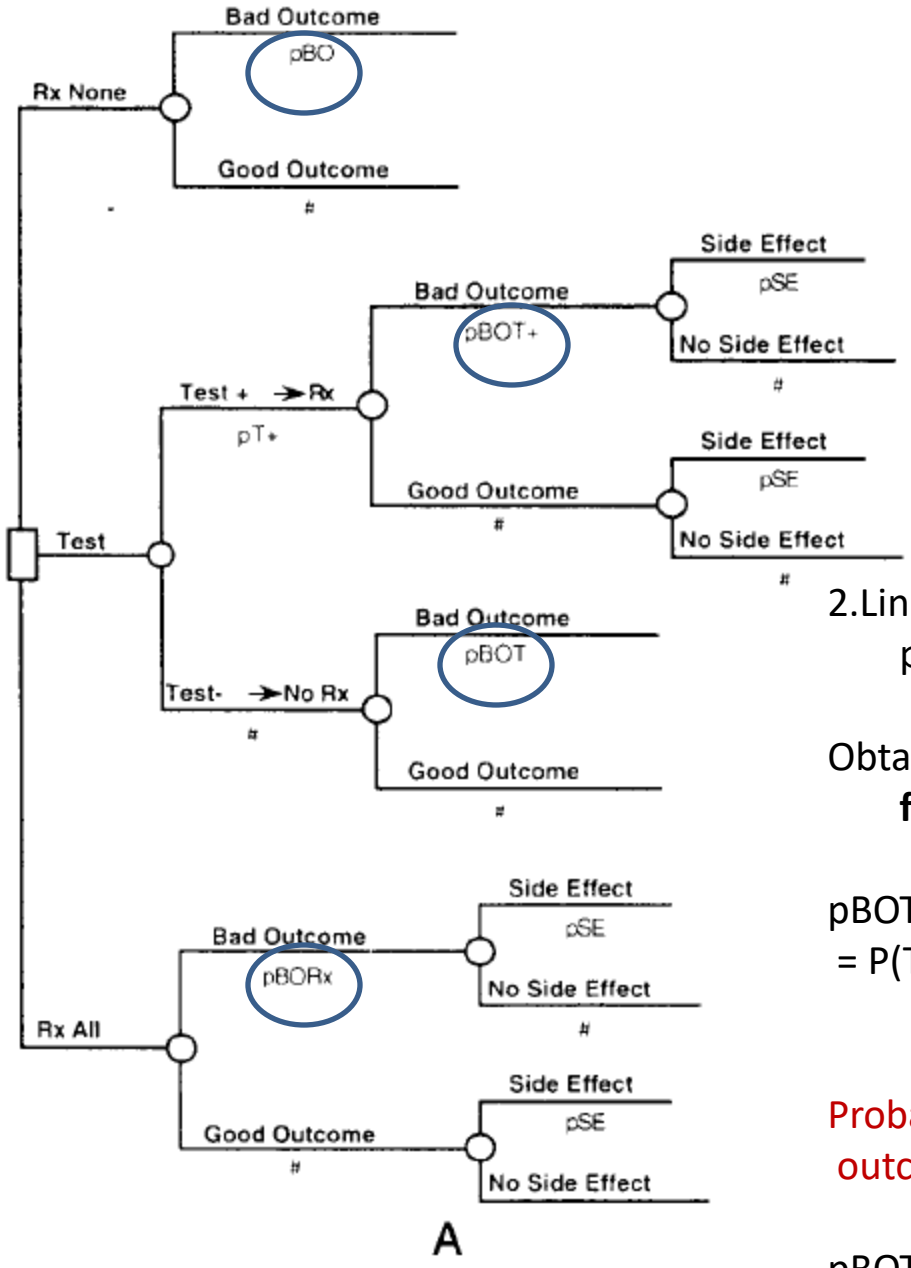




pBO has no effect on EU of Test and Treat all strategies (?)!!!

Possible for Treat All

Something wrong with the EU of the Test strategy !



1. Linking the probabilities of outcomes by effectiveness of a treatment:

Effectiveness: ***proportionate reduction*** in the probability of an adverse outcome resulting from treatment

e=(pBO-pBORx)/pBO

pBORx=PBO-e*pBO=pBO*(1-e)

2.Linking the outcomes by test characteristics and prevalence

Obtain posttest probabilities (pBOT+, pBOT-) **as a function of the baseline disease probability** (pBO)

pBOT+=P(Bad Outcome if Test+)

= P(Test+ and Bad Outcome) /P(Test+)

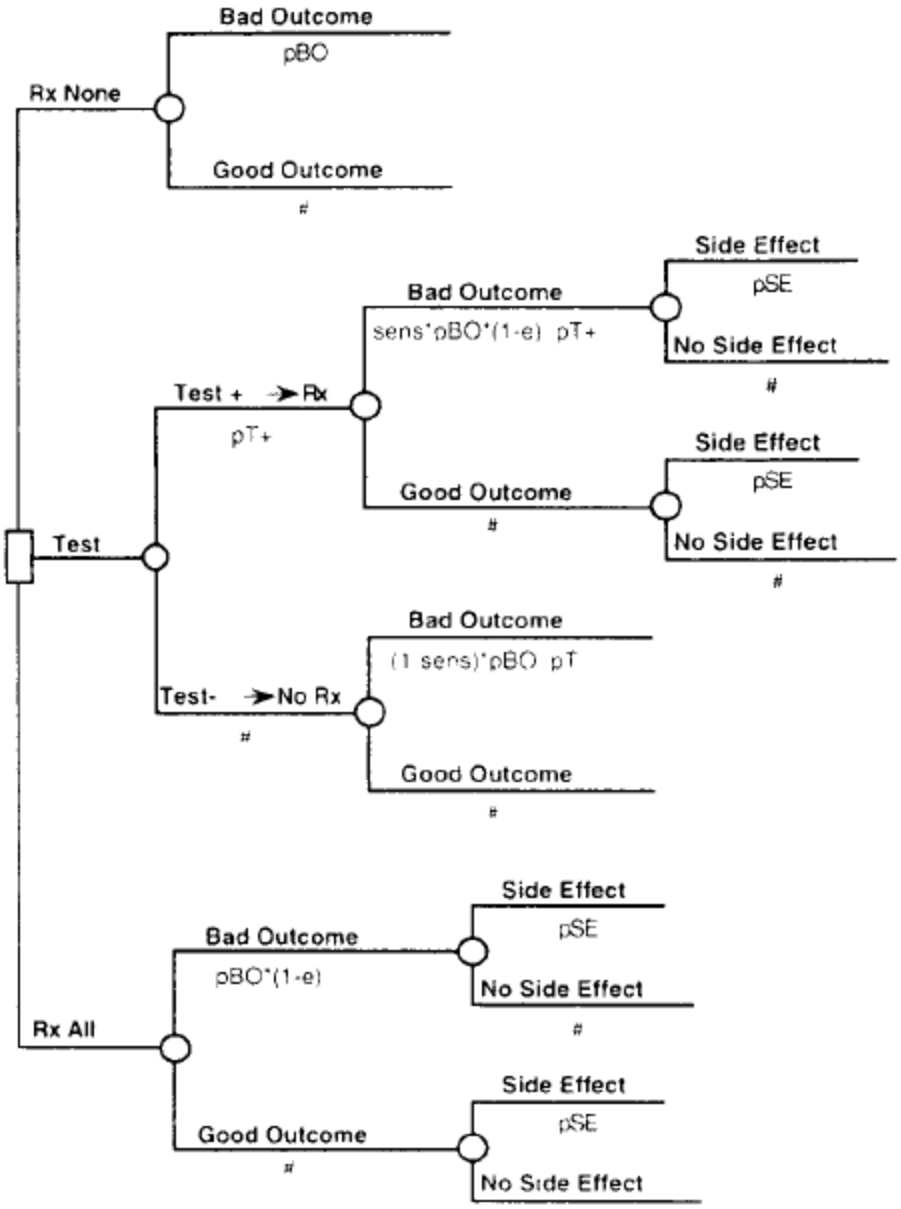
=sens* PBO* (1-e)/PT+

Probability of a bad outcome without treatment

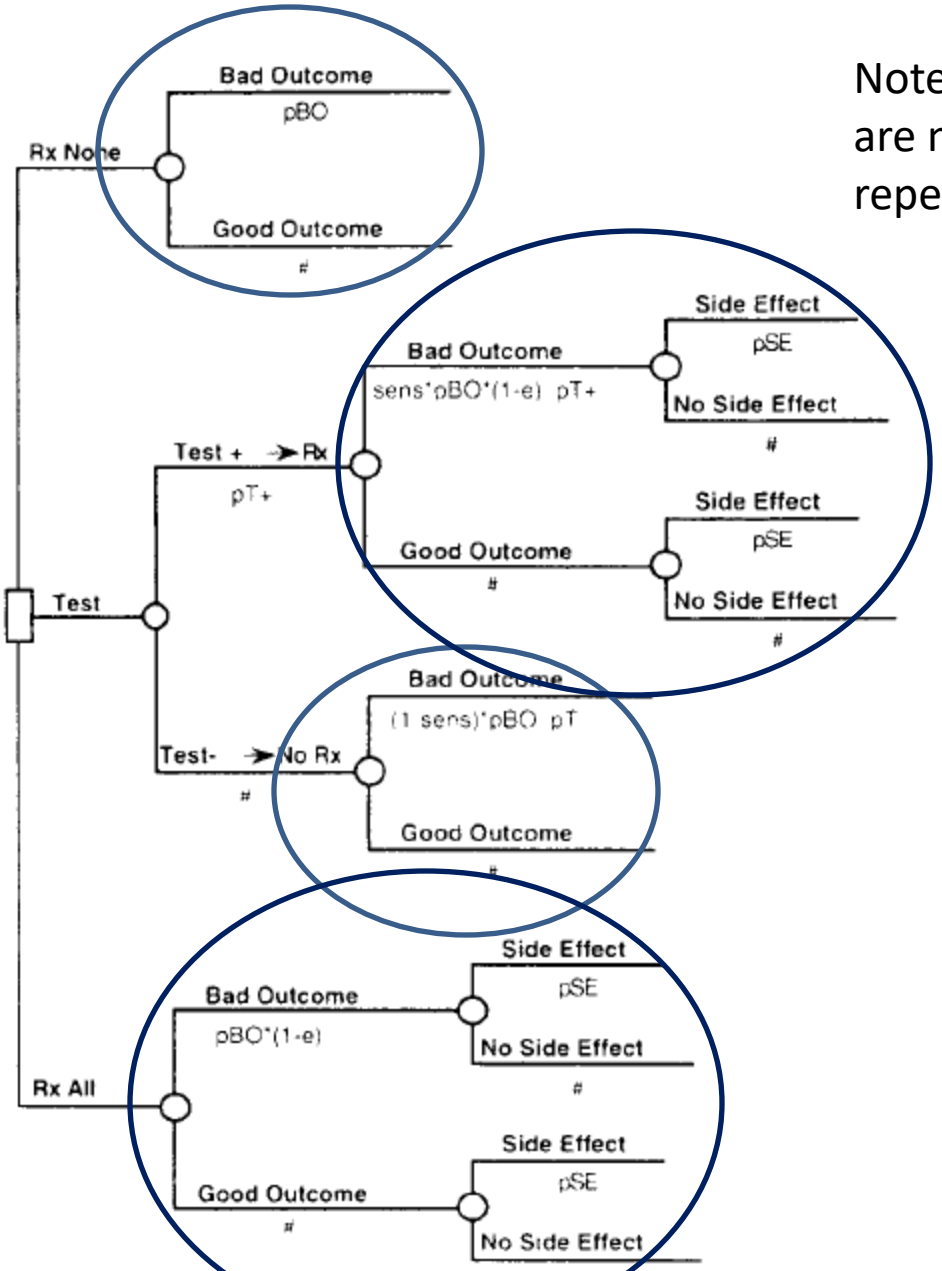
probability of failure in treatment

pBOT-=P(Bad outcome without treatment|Test)-

=pBO*(1-sens)/PT-


Bad Outcome
pBO
Rx None
Good Outcome
#
Side Effect
pSE
Bad Outcome
sens*pBO*(1-e) pT+
No Side Effect
#
Test + ➔Rx
pT+
Side Effect
pSE
Good Outcome
#
No Side Effect
#
Test
Bad Outcome
(1 sens)*pBO pT
Test- ➔No Rx
#
Good Outcome
#
Side Effect
pSE
Bad Outcome
pBO*(1-e)
No Side Effect
#
Rx All
Side Effect
pSE
Good Outcome
#
No Side Effect

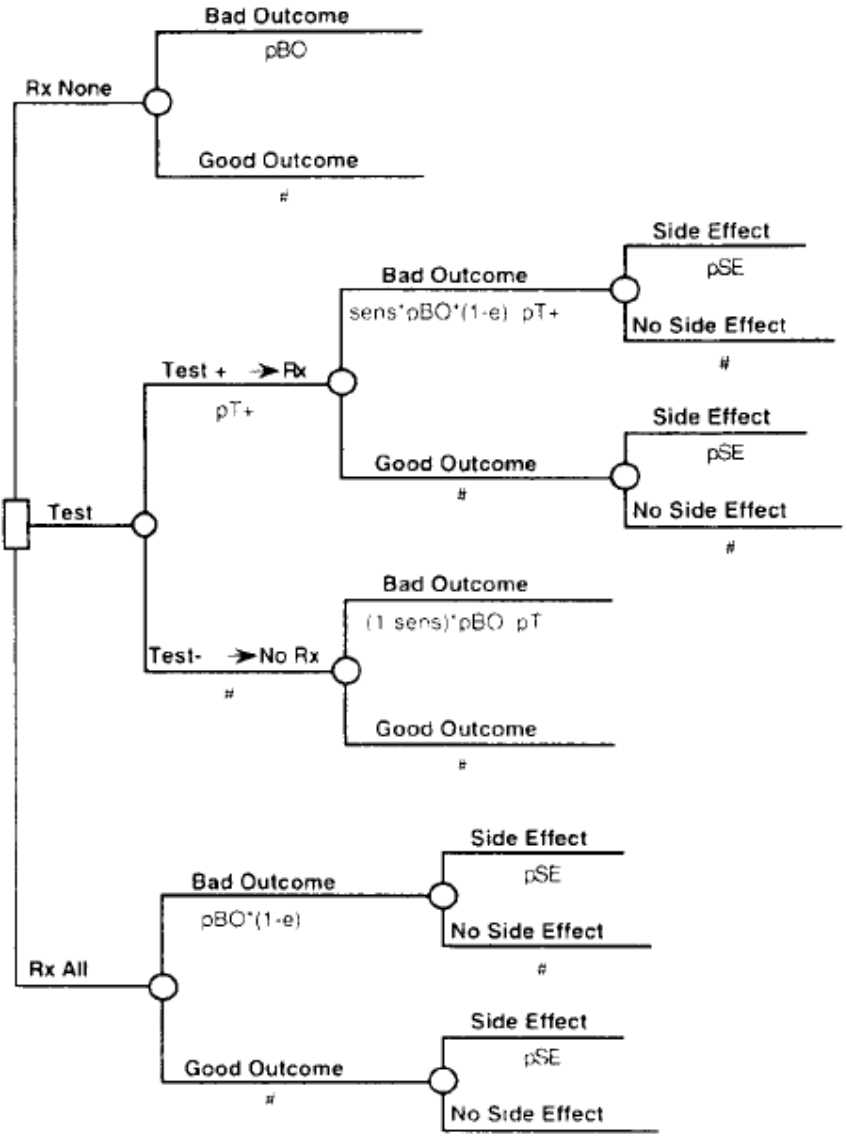
5. The tree must have symmetry

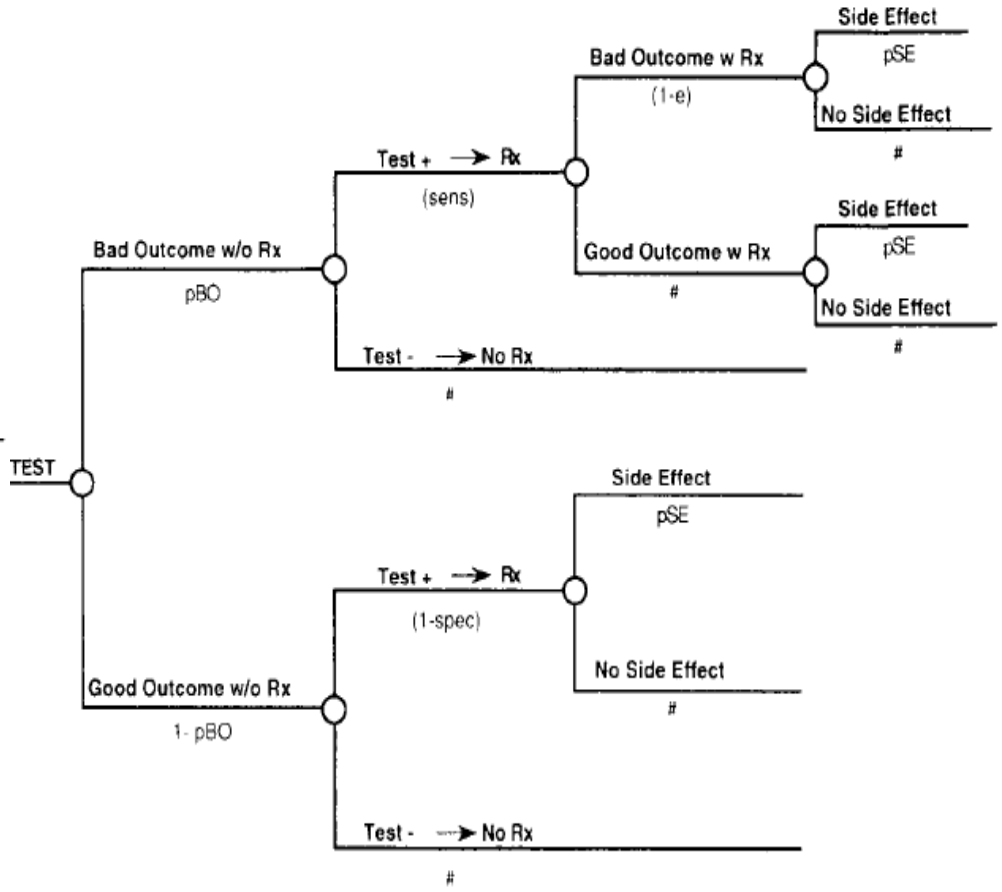


Note that the probabilities are not the same in the repeated subtrees!

Order is not important

- The order of the outcomes is not important





Underlying probability of a bad outcome for the patient is modelled first

Test results are modelled before the adverse outcomes

Estimating probabilities

- The analysis that uses the best estimates of the probabilities is called a baseline analysis
- The range of reasonable estimates should be specified
  - From different studies
  - 95% confidence interval

- Systematic search of the literature
  - Computerized search
  - Personal files
  - Files of content experts
  - A review of reference lists from the articles

- Evaluate the validity of their results (critical appraisal)
  - Quality poor
  - Quality high but study population or treatment intervention different
  - Beware of over estimations of effectiveness

Estimating probabilities

- Eliminate the results that are of poor methodological quality
- Use the average of the remaining results (meta-analysis may be already available)

Table 1 • Example of Probability Table

Variable	Probability*	
	Baseline	Range
Probability of major complication of giant cell arteritis	0.12	0.05–0.40
Temporal artery biopsy		
Sensitivity	0.80	0.58–0.97
Specificity	1.00	0.90–1.00
Effectiveness of prednisone	0.89	0.69–1.00
Probability of major complication of prednisone use	0.19	0.05–0.40

*Baseline probabilities are the averages of estimates from published studies; ranges are based on the highest and lowest estimates from published studies. The specific references for the probability estimates can be found in Buchbinder and Detsky.[10]

If no such information use alternative sources:

Expert judgement

Existing databases

Primary data collection

Sensitivity analysis with a wide range of possible values!

Estimating utilities

Outcomes

- Life years
- QALYs
- Cases of disease
- Complications prevented
- Utilities

- If only two possible outcomes: use 1 for the better outcome 0 for the worse
- Utilities can be used either as the actual outcome values or can be used to as weights to calculate QALYs
- Estimating utilities
  - Based on your judgement
  - Expert judgement (a group of experts)
  - Published values from the literature
  - Direct measurement in appropriate subjects (time trade-off, VAS, Standard gamble etc)

Estimating utilities

Table 2 • Outcomes for Decision Options in the Management of Giant Cell Arteritis (GCA)

Treat all
No GCA, prednisone treatment
No GCA, prednisone treatment, major prednisone complication
GCA, prednisone treatment
GCA, prednisone treatment, major prednisone complication
GCA, prednisone treatment, major GCA complication
GCA, prednisone treatment, major prednisone complication, major GCA complication
Treat none
No GCA
GCA
GCA, major GCA complication
Biopsy and treat positives
Biopsy positive
No GCA, prednisone treatment, TA* biopsy
No GCA, prednisone treatment, major prednisone complication, TA biopsy
GCA, prednisone treatment, TA biopsy
GCA, prednisone treatment, major prednisone complication, TA biopsy
GCA, prednisone treatment, major GCA complication, TA biopsy
GCA, prednisone treatment, major prednisone complication, major GCA complication, TA biopsy
Biopsy negative
No GCA, TA biopsy
GCA, TA biopsy
GCA, major GCA complication, TA biopsy

*TA = temporal artery.

One possible outcome

When outcomes are a combination of health states:

Holistic method (Assess utility as a whole)

U(GCA, TA biopsy, prednisone treatment, major GCA complication, major prednisone complication)

Decomposed method (Assess utility in parts)

Categorize the states as short term and long term

Short term (days to weeks)

- Temporary hospitalization
- Unpleasant diagnostic procedures

Long term(enduring impact)

- Chronic symptoms from a disease,
- Negative impact of being persistently on medication

Estimating utilities

Table 3 • Utility Estimates for Giant Cell Arteritis (GCA) Decomposed Health States

Health State	Baseline Utility*	Range*
No GCA	1.00	
GCA symptoms	0.85	0.70–0.95
Major GCA complication	0.60	0.20–0.85
Prednisone treatment	0.97	0.90–1.00
Major prednisone complication	0.75	0.60–0.90
Temporal artery biopsy	0.995	0.97–1.00

*Baseline utilities and ranges are based on consensus estimates of a group of expert physicians.

Long term states

Aggregate the separate utilities

- Adding
- Multiplying
- Adding some multiplying some

Underlying assumptions are different

Estimating utilities

Table 3 • Utility Estimates for Giant Cell Arteritis (GCA) Decomposed Health States

Health State	Baseline Utility*	Range*
No GCA	1.00	
GCA symptoms	0.85	0.70–0.95
Major GCA complication	0.60	0.20–0.85
Prednisone treatment	0.97	0.90–1.00
Major prednisone complication	0.75	0.60–0.90
Temporal artery biopsy	0.995	0.97–1.00

*Baseline utilities and ranges are based on consensus estimates of a group of expert physicians.

Long term states

One possible way

Convert short term state utility values to disutility values(1-utility)

Multiply the utility values of long term states and subtract the disutility

GCA, prednisone treatment, major prednisone complication, major GCA complication, TA biopsy

Assumption: Prednisone therapy eliminates all the symptoms of GCA, so the utility of having GCA symptoms on prednisone therapy is equal to the utility of no GCA (1)

=(1*0.97*0.75*0.60)-0.005=0.432

Estimating utilities

Table 4 • Rank Ordering of Giant Cell Arteritis (GCA) Outcome Values

Outcome State	Utility*	QALYs†
No GCA, no prednisone, no TA‡ biopsy	1.000	13.600
No GCA, no prednisone, TA biopsy	0.995	13.595
No GCA, prednisone	0.970	13.192
GCA, prednisone§	0.970	13.192
No GCA, prednisone, TA biopsy	0.965	13.187
GCA, prednisone, TA biopsy	0.965	13.187
GCA, no prednisone	0.850	11.560
GCA, no prednisone, TA biopsy	0.845	11.556
No GCA, prednisone, major prednisone complication	0.728	9.894
GCA, prednisone, major prednisone complication	0.728	9.894
No GCA, prednisone, major prednisone complication, TA biopsy	0.723	9.890
GCA, prednisone, major prednisone complication, TA biopsy	0.723	9.890
GCA, prednisone, major GCA complication	0.582	7.915
GCA, prednisone, major GCA complication, TA biopsy	0.577	7.912
GCA, no predinsone, major GCA complication	0.510	6.936
GCA, no prednisone, major GCA complication, TA biopsy	0.505	6.933
GCA, prednisone, major prednisone and GCA complications	0.437	5.936
GCA, prednisone, major prednisone and GCA complications, TA biopsy	0.432	5.934

*Utilities are calculated by multiplying the baseline utilities of the long-term states and then, when applicable, subtracting the disutility (i.e., 1 − utility) value for temporal artery biopsy.

†QALYs = quality-adjusted life years are calculated by, when applicable, subtracting the time period of negative impact of temporal artery biopsy from the life expectancy and then multiplying the difference by the product of the baseline utilities of the long-term states.

‡TA = temporal artery.

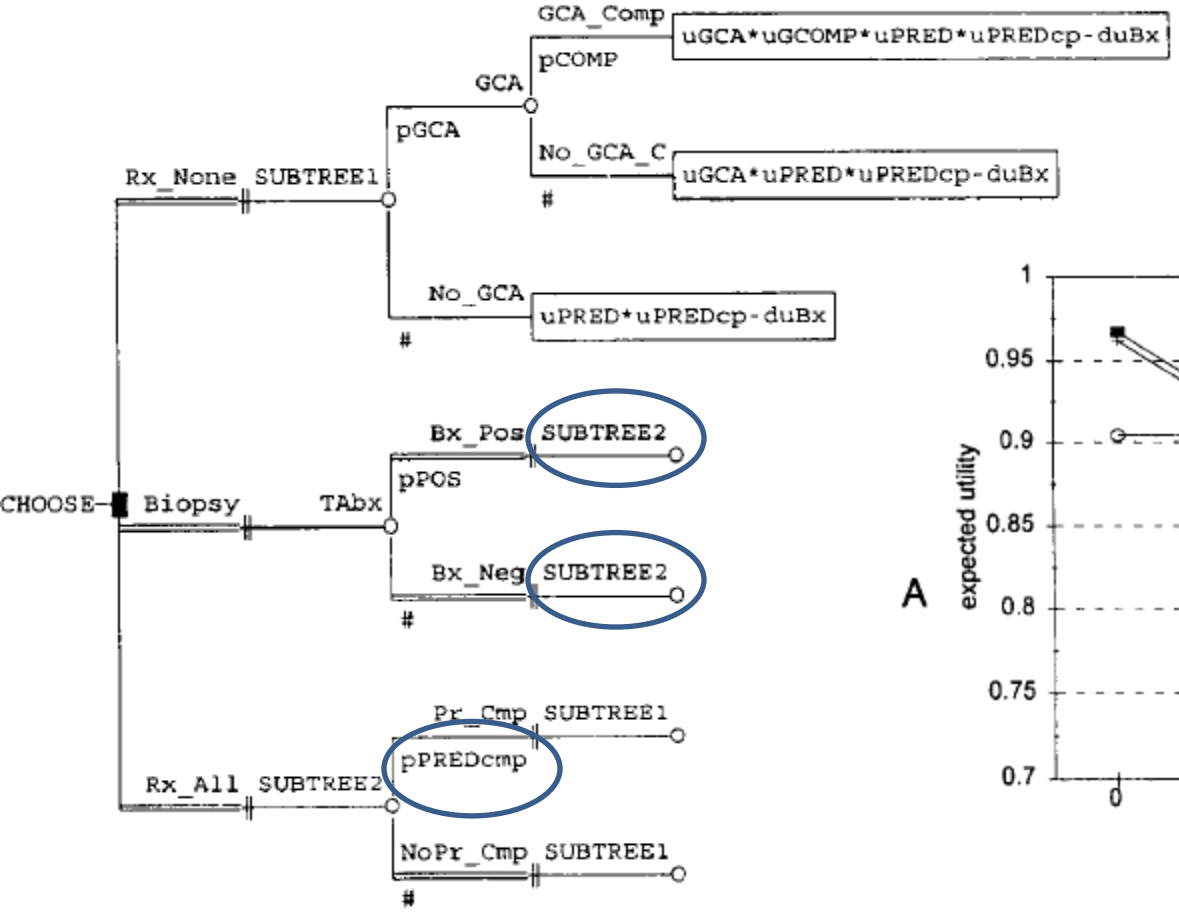
[utility of GCA symptoms on prednisone therapy × utility of taking prednisone therapy daily × utility of a major prednisone complication × utility of a major GCA complication] × [LE − time period of negative impact from temporal artery biopsy] = [1.0 × 0.97 × 0.75 × 0.60] × [13.6 years − 0.005 years] = 5.934 QALYs.

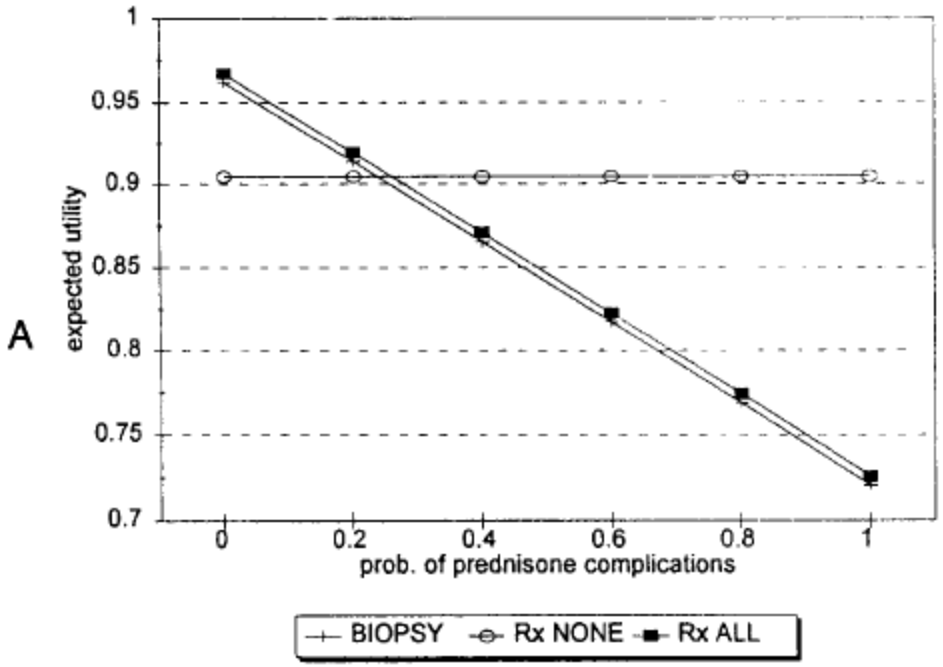
Analyzing the model and interpreting the results

- Folding back will give us the expected value of each strategy
- Before the results are interpreted sensitivity analysis is required
- Sensitivity Analysis: The process of repeatedly folding back the tree using different values for probability and utility variable
  - Debugging the tree
  - Evaluating Uncertainty

Debugging

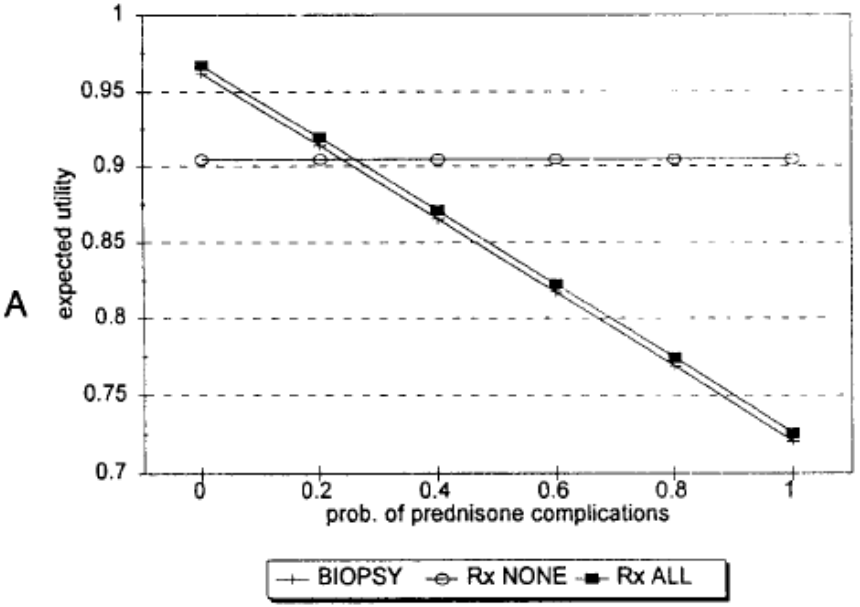
1. Change one variable at a time (one way sensitivity analysis) over its entire range
2. Evaluate the results graphically

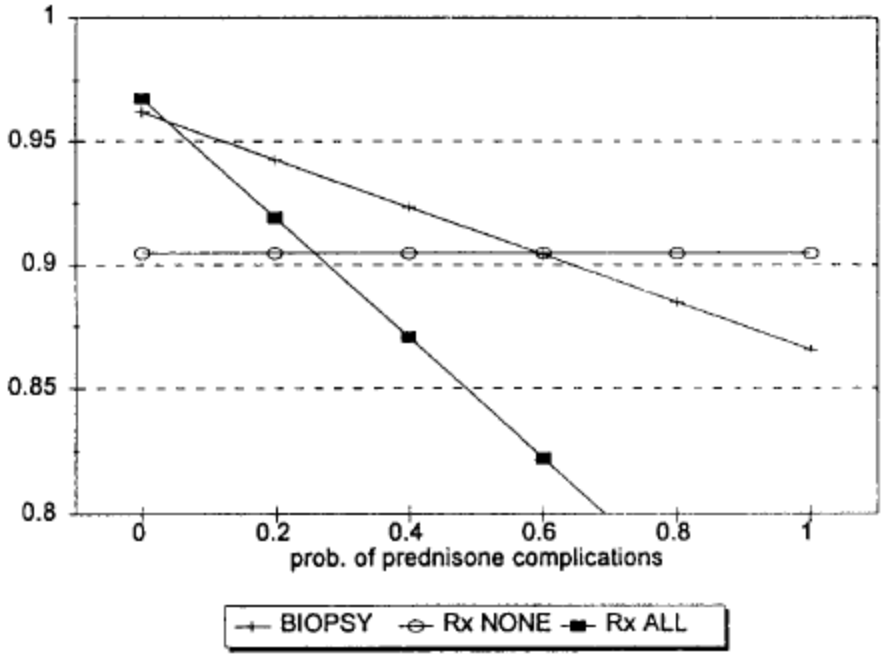





```
Bindings from TAbx to Bx_Neg:
   pGCA     := ((1-SENS)*pGCA)/((1-SENS)*pGCA+SPEC*(1-pGCA))
   pPREDcmp := 0
   uPRED    := 1
```

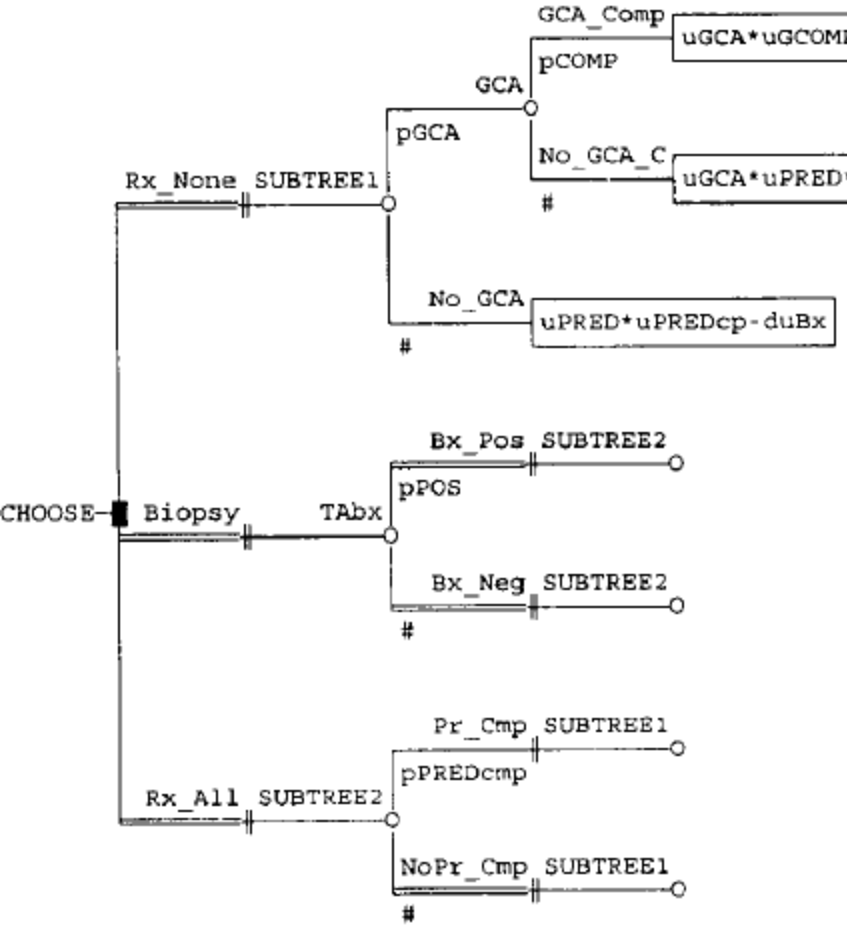
Debugging

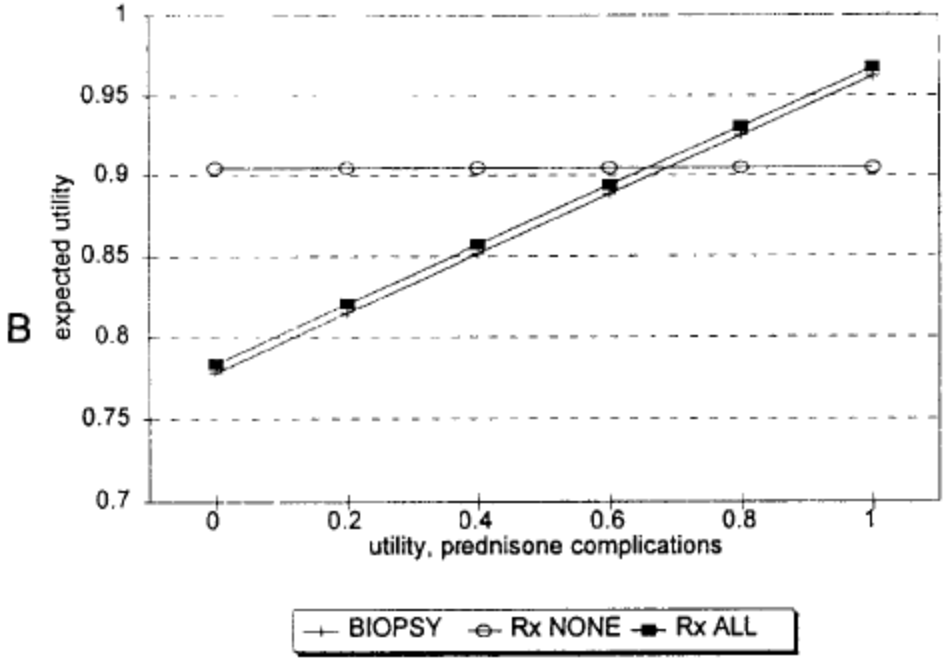




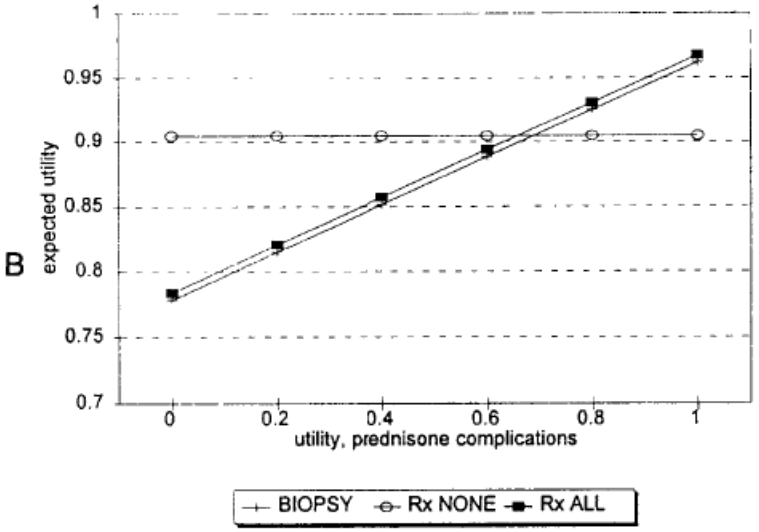
Evaluate the slopes
Evaluate the rank order

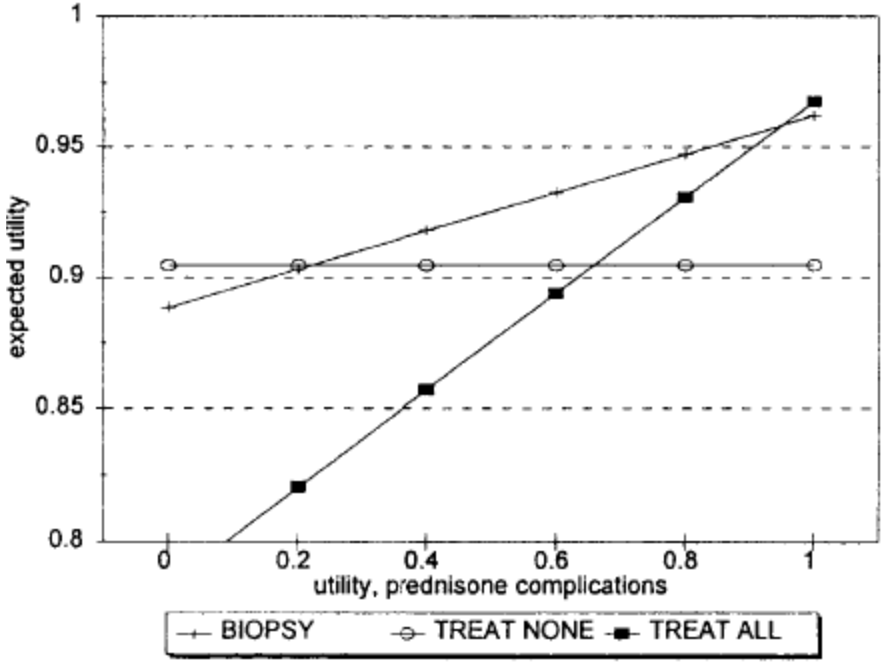
Debugging


GCA_Comp
uGCA*uGCOMP*uPRED*uPREDcp-duBx
pCOMP
GCA
pGCA
No_GCA_C
uGCA*uPRED*uPREDcp-duBx
#
Rx_None SUBTREE1
No_GCA
uPRED*uPREDcp-duBx
#
Bx_Pos SUBTREE2
pPOS
CHOOSE
Biopsy
TAbx
Bx_Neg SUBTREE2
#
Pr_Cmp SUBTREE1
pPREDcmp
Rx_All SUBTREE2
NoPr_Cmp SUBTREE1
#

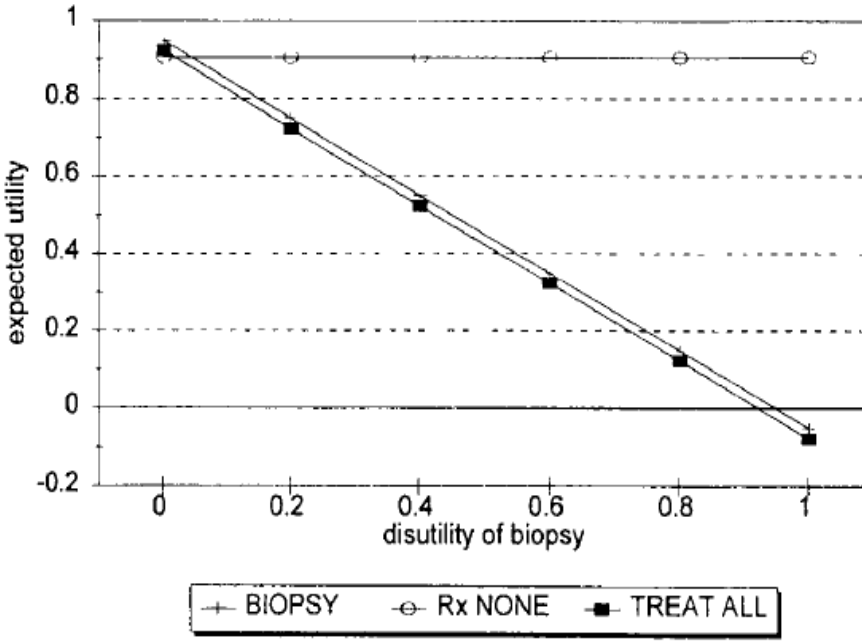

B
expected utility
1
0.95
0.9
0.85
0.8
0.75
0.7
0
0.2
0.4
0.6
0.8
1
utility, prednisone complications
BIOPSY
Rx NONE
Rx ALL

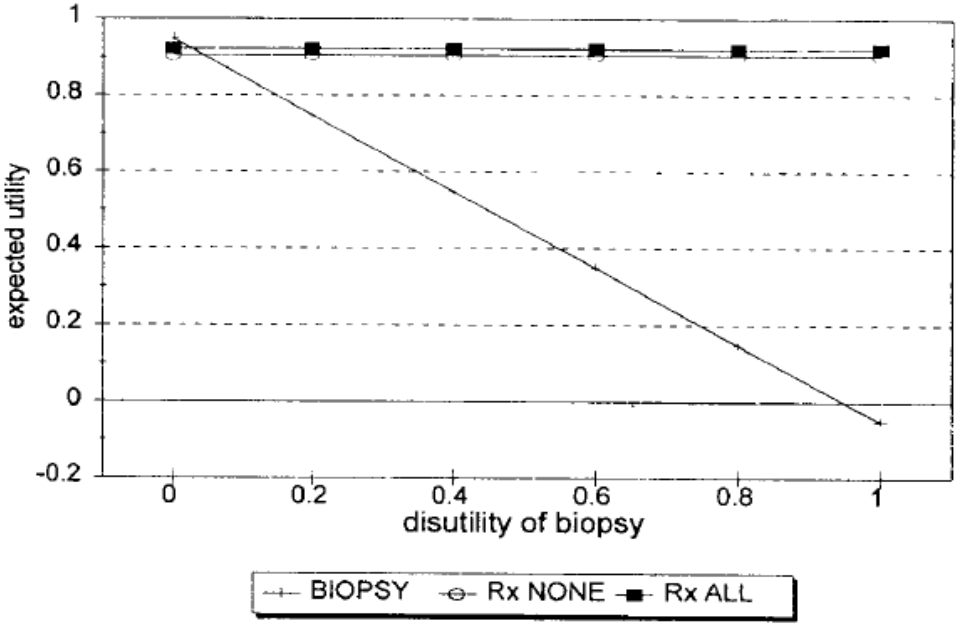
Debugging


B
1
0.95
0.9
0.85
0.8
0.75
0.7
expected utility
0
0.2
0.4
0.6
0.8
1
utility, prednisone complications
BIOPSY
Rx NONE
Rx ALL


1
0.95
0.9
0.85
0.8
expected utility
0
0.2
0.4
0.6
0.8
1
utility, prednisone complications
BIOPSY
TREAT NONE
TREAT ALL

Debugging


1
0.8
0.6
0.4
0.2
0
-0.2
expected utility
0
0.2
0.4
0.6
0.8
1
disutility of biopsy
BIOPSY
Rx NONE
TREAT ALL


1
0.8
0.6
0.4
0.2
0
-0.2
expected utility
0
0.2
0.4
0.6
0.8
1
disutility of biopsy
BIOPSY
Rx NONE
Rx ALL

More on debugging

- Structural problems
  - Symmetry problems
  - Linkage Problems: All common events in separate branches (treatment efficacy, predictive values of tests) linked by
    - Subtrees
    - Common variable names
    - Common expressions (efficacy equations, predictive value expressions for test results conditioned on disease prevalence)

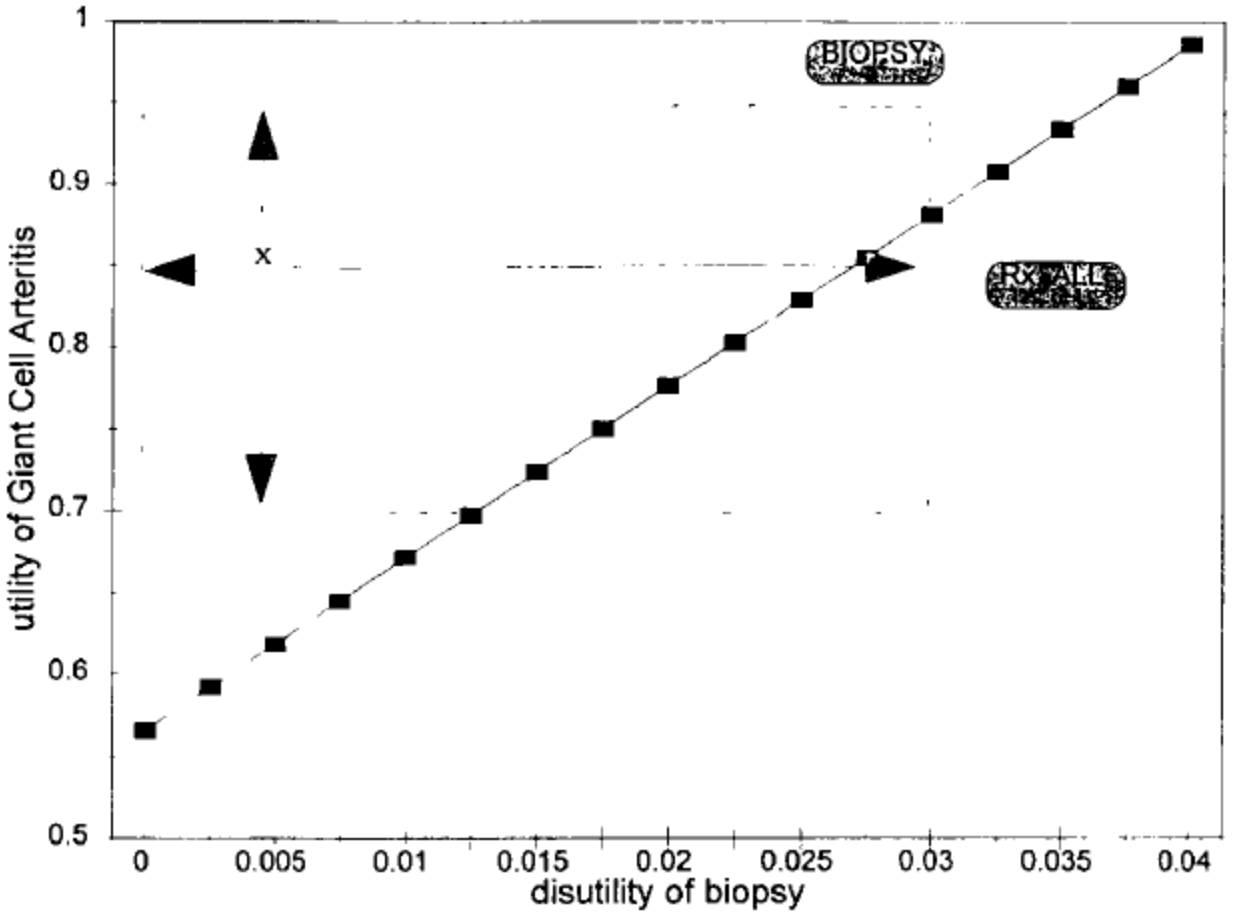
- Technical Errors
  - Lower case-upper case
  - Spelling errors
  - Inconsistent abbreviations
  - Wrong variable names
  - Errors in temporary bindings: very frequent source of error!

Evaluating uncertainty

Table 2 • Sensitivity Analyses

Variable	Baseline Value	Plausible Range	Threshold Value*	Sensitive?†
Prevalence of giant cell arteritis (GCA)	0.50	0.00–1.00	NA‡	NA‡
Probability of adverse outcome from GCA	0.12	0.05–0.40	0.86	N
Sensitivity of temporal artery biopsy	0.80	0.58–0.97	0.40	N
Specificity of temporal artery biopsy	1.00	0.90–1.00	0.42	N
Effectiveness of prednisone	0.89	0.60–1.00	NT§	N
Probability of iatrogenic side effects from prednisone	0.19	0.05–0.40	0.04	N
Disutility of biopsy	0.005	0.00–0.03	0.003	Y
Utility of GCA	0.85	0.70–0.95	0.95	Y
Utility of complications of GCA	0.60	0.20–0.85	NT§	N
Utility of taking prednisone	0.95	0.90–1.00	0.87	N
Utility of prednisone complications	0.75	0.60–0.90	0.95	N
Bias to Rx ALL	—	—	—	Y¶
Bias to Rx NONE	—	—	—	Y¶

Evaluating uncertainty



Evaluating uncertainty