Uncertainty and Sensitivity Analyses

Marcelo Coca Perraillon

University of Colorado
Anschutz Medical Campus

Cost-Effectiveness Analysis
HSMP 6609
2016

Outline

- Defining terms: variability, heterogeneity, uncertainty
- Sensitivity analysis: "deterministic" and "probabilistic"
- Base case, one-way, two-way, three-way, scenarios
- Influential variables: tornado diagrams
- More advanced methods: probabilistic sensitivity analysis (PSA)
- Probability distributions
- Example: HIV model
- ANCOVA and Expected Value of Perfect of Information (EVPI)

What is sensitivity analysis

- We have already performed some basic sensitivity analyses
- It is essentially the study of **how changes in model inputs affect model outputs**
- **Inputs** are quantities, prices, life years, probabilities, and so on. In essence, all we measured in the CEA study
- **Outputs** are the model outcomes we care about: average or expected costs, average or expected benefits, average or expected ICER or Net Monetary Benefit

Deterministic and probabilistic sensitivity analysis

- Two types of sensitivity analyses:
  1. **Deterministic**: We choose values for one or more parameters keeping the rest constant. For example, min or max or a case that has policy relevance
  2. **Probabilistic**: We assign parameters a probability distribution and use simulations to compute new ICERs

Defining terms

- **Variability**: Variation due to chance; random variability. Even with identical patients facing the same probability of an event, some experience the event and some do not. Sometimes called *stochastic uncertainty* or first-order uncertainty (less common)
- **Heterogeneity**: Differences between patients that can be attributed or explained by patient's characteristics (think sex, age, income, and so on)
- **Uncertainty**: What sensitivity analysis tries to measure. We **do not know the true value** of some input (parameter) or the true way a process is generated
- Two types of uncertainty: **parameter uncertainty** and **model uncertainty**

Parameter uncertainty

- We do not know the true value of a parameter. For example, the true average cost for a procedure in the population
- We could estimate the mean cost in a sample but the estimated parameter (the mean) itself has some variability
- In statistics, a measure of uncertainty in the parameter estimation is the **standard error** (do not confuse it with the standard deviation)
- We can reduce some of this uncertainty if we had a larger sample; the standard error also depends on the standard deviation

Model uncertainty

- Models are simplifications of reality; some models capture the key elements of reality and some don't
- Remember the classic "All models are wrong; some are useful"
- The doubts we have about the model assumptions is **model uncertainty**

Analogy with regression

- A standard linear regression model can be written as:

  $y = \beta_0 + \beta_1 X_1 + ... + \beta_k X_k + \epsilon$

  where ϵ is assumed to distribute normal with mean 0 and variance σ^2
- After collecting data on variables y and X_1 to X_k, we can estimate the model:
- $\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 X_1 + ... + \hat{\beta}_k X_k$
- Standard errors for the coefficient: $SE(\hat{\beta}_j)$ are also estimated
- With the coefficients and the SEs, it is possible to test hypotheses

Analogy with regression

- The random error ϵ corresponds to stochastic uncertainty or variability due to chance
- Parameter uncertainty is measured by the standard error $SE(\hat{\beta}_j)$ of the estimated parameter $\hat{\beta}_j$
- Heterogeneity of effects is represented by coefficients $\hat{\beta}_j$
- Model uncertainty is the uncertainty around the assumptions of the model. Is it correct to assume that the error, and consequently the outcome y, distribute normal? Are we including all the relevant covariates X? Should we model y or $log(y)$?

Cheat sheet

- Trying to come up with a common terminology. From Briggs et al (2012):

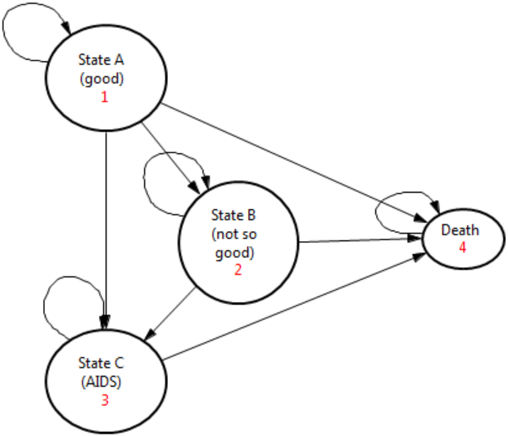
Table 1 – Uncertainty for decision modeling: Concepts and terminology.

Preferred term	Concept	Other terms sometimes employed	Analogous concept in regression
Stochastic uncertainty	Random variability in outcomes between identical patients	Variability Randomness Monte Carlo error First-order uncertainty	Error term
Parameter uncertainty	The uncertainty in estimation of the parameter of interest	Second-order uncertainty	Standard error of the estimate
Heterogeneity	The variability between patients that can be attributed to characteristics of those patients	Variability Observed or explained heterogeneity	Beta coefficients (or the extent to which the dependent variable varies by patient characteristics)
Structural uncertainty	The assumptions inherent in the decision model	Model uncertainty	The form of the regression model (e.g., linear, log-linear)

Deterministic sensitivity analysis

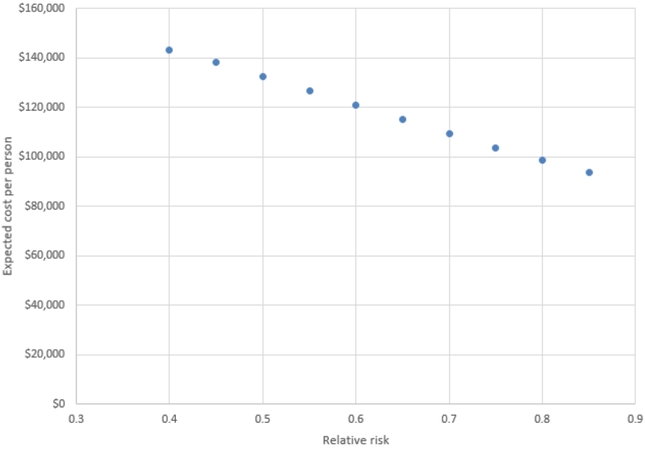
- **One-way**: Change one parameter at a time keeping all others constant
- Standard way of presenting one-way sensitivity analyses results is to plot the parameter you are changing in the x-axis and an output of interest on the y-axis
- In the HIV example, we could change the relative risk and analyze what happens to average costs in the combination therapy group and the ICER

HIV transition diagram



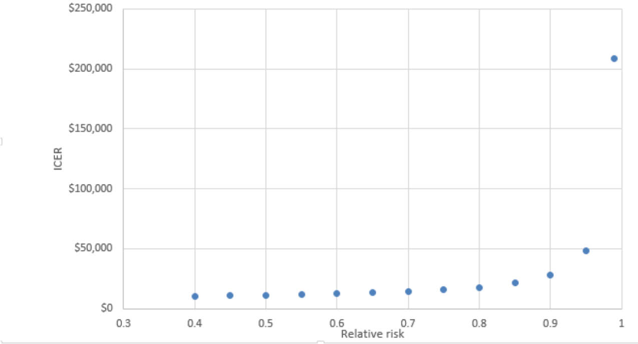
- See Excel file for more details

One-way sensitivity analysis of relative risk (cost)



- Lower RR implies combo therapy more effective, so more people stay in states that are costly

One-way sensitivity analysis of relative risk (ICER)



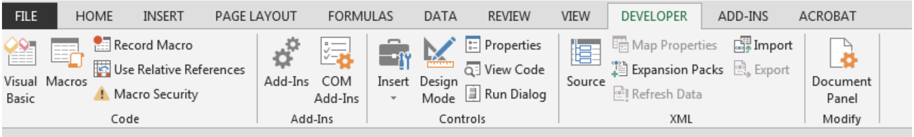
- ICER shows more nuance because of more moving parts. More effective drug increases costs but it also increases life expectancy. Less effective combo therapy becomes more similar to mono therapy but costs keep going up

Using a macro for sensitivity analysis in Excel

- It is easy to change values in Excel but it is not the most practical way to perform many sensitivity analyses
- Excel has a **macro language** that is very flexible and relatively easy to use, although it can take a while to master it
- It's based on a version of a programming language, Visual Basic (Visual Basic for Applications, VBA)
- Fortunately, you don't need to master it in order to use it for simple analyses
- We will use a simple macro for quick sensitivity analyses

Macros in Excel

- First, make sure that you can see the Developer ribbon in Excel:
  1. Click the File tab.
  2. Click Options.
  3. Click Customize Ribbon.
  4. Under Customize the Ribbon and under Main Tabs, select the Developer check box



Record a macro

- The easiest way to use macros is to **record** a macro of a repetitive task like changing the RR and then copying and pasting the resulting ICER into a new column
- Excel will create VBA code to do the repetitive task
- It is then fairly easy to modify the code to complete a task
- See accompanying Excel file for an example

Final code

```
Sub oneway()
'
' oneway Macro
' Performs a basic sensitivity analysis
'
' Keyboard Shortcut: Ctrl+Shift+O

Dim i As Integer

For i = 5 To 17
    Range("D" & i).Select
    Selection.Copy
    Sheets("Markov").Select
    Range("D26").Select
    ActiveSheet.Paste
    Range("U32").Select
    Application.CutCopyMode = False
    Selection.Copy
    Sheets("Macro").Select
    Range("E" & i).Select
    Selection.PasteSpecial Paste:=xlPasteValues, Operation:=xlNone, SkipBlanks _
        :=False, Transpose:=False
Next i

End Sub
```

Logic

- For each copy-paste step, Excel creates a chunk of code
- We used this chunk of code to create a loop so Excel does the same in a range of cells, from rows 5 to 17
- We added
  1. Dim i As Integer
  2. For i = 5 To 17
  3. Next i
- See `http://www.excel-easy.com/vba.html` for more Examples
- Remember, this code only repeats interactions with Excel that you could do "by hand"

Tornado diagrams

- Tornado diagrams are a nice method to depict results of one-way sensitivity analyses
- Provide a way to quickly figure out which **parameters has the most influence** on outcomes
- Not the only way to present results but a common way to do so in CEA
- We will do a tornado diagram of the the HIV model

Tornado diagram HIV model

- We need to come up with a **range of values** for the parameters
- The most important variables are the cost of the medications and the relative risk

Variable	Base case	Lower-bound	Upper-bound
Drug price mono	2278	1595	2961
Drug price combined	4364	3055	5673
Relative risk	0.509	0.4	0.6

Where do we get these numbers?

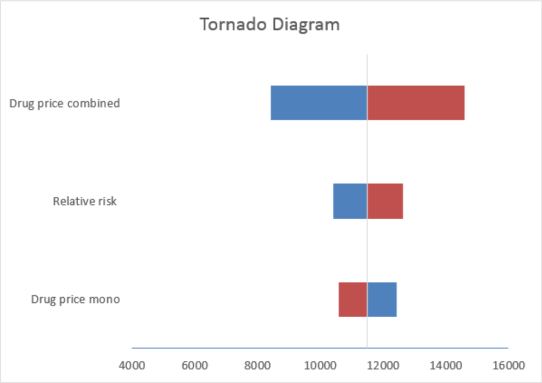
- The **base-case** scenario is our best estimate (sometimes called **baseline** values)
- The range could be obtained from multiple sources, like drug prices that are common for hospitals or prices after discounts
- The relative risk was obtained from a meta-analysis, so a confidence interval could be available
- We can then calculate the ICER corresponding to each of the values in the table

Table for tornado diagram

Variable	ICER-Lb	ICER-Ub	Range
Drug price mono	12426	10572	1854
Relative risk	10422	12621	2199
Drug price combined	8413	14585	6172

- Lb = lower bound; Ub = upper bound
- The range is the absolute value of the difference between the upper and the lower bound
- Table is sorted from lowest to largest range
- **A tornado diagram is a graphical representation of this table**

Tornado diagram

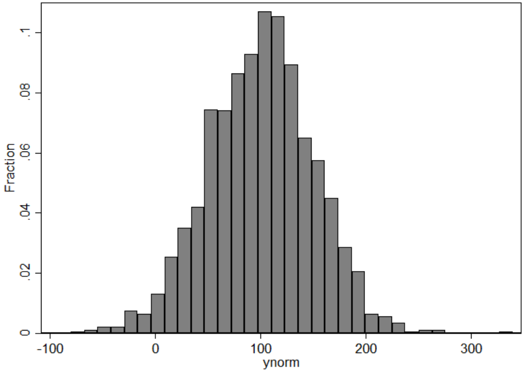


- Vertical line is the ICER for base case ($11,499); the most influential variable is the drug price for combination therapy
- See Excel file for details; see this video to create a similar plot: `https://vimeo.com/34389151`

Probabilistic Sensitivity Analysis

- Instead of changing one or more variables at a time we will change all (or many) variables simultaneously using a simulation (often called **second-order Markov simulation**)
- Instead of coming up with a range of possible values for each variable we will assume that a parameter has a **probability distribution**
- Remember that a probability distribution assigns a probability to each possible outcome of a random variable
- We can **describe** a distribution using its parameters
- For example, a normal distribution can be described by its **mean** and **variance** (two-parameters)

Normal distribution



- Simulated values for a normal variable with mean 100 and standard deviation of 50; $N(100, 50^2)$

Simulation

- If we assume that a variable has a distribution and have the parameters of the distribution, we can use software to generate a sample of values
- Think of this as asking the computer to pick random numbers for you. The only condition is that these numbers must come from a probability distribution with certain characteristics
- For example, I used Stata to simulate 10 draws of a normal distribution with mean of 100 and standard deviation of 50
- The first ten values were: 52.32808, 125.6597, 187.7224, 113.5894, 113.8938, 138.1172, 74.36327, 77.44591, -21.80569, 37.02974
- Digression: technically, computers generate **pseudo-random** numbers; need a "seed"

PSA

- The procedure to conduct PSA is simple:
  1. Assign a probability distribution to each variable
  2. Draw a number from these distributions
  3. Calculate ICER (or any other outcome of interest)
  4. Save ICER
  5. Repeat procedure many times (1000 or 10,000 times)
  6. Plot all ICERs in a cost-effectiveness plane
  7. Calculate statistics (like the percent of time the ICER is below threshold)
- The resulting "cloud" of ICERs will give you a sense of the overall uncertainty

PAS for HIV example

- Campell et al (2014) performed a PSA of the HIV example
- Assumed distributions

Table 1 Input Parameters and Ranges (Preliminary Model)

	Mean	Standard Error	Distribution	95% Interval for Inputs Low	High	Source
Costs (1994–1995 British pounds)						
Direct medical costs State A	£1701	£1701	$\gamma(1, 1701)$	£42	£6360	a
Direct medical costs State B	£1774	£1774	$\gamma(1, 1774)$	£47	£6541	a
Direct medical costs State C	£6948	£6948	$\gamma(1, 6948)$	£175	£24,776	a
Community care costs State A	£1055	£1055	$\gamma(1, 1055)$	£23	£3894	a
Community care costs State B	£1278	£1278	$\gamma(1, 1278)$	£31	£4728	a
Community care costs State C	£2059	£2059	$\gamma(1, 2059)$	£49	£7577	a
Utilities						
Utility score State A	0.880	0.132	$\beta(5.33, 0.73)$	0.553	0.999	b
Utility score State B	0.780	0.117	$\beta(9.78, 2.76)$	0.527	0.954	b
Utility score State C	0.700	0.105	$\beta(13.33, 5.71)$	0.486	0.877	b
Combination therapy treatment effect						
Relative risk[c]	0.509	0.088	Log-normal(–0.68, 0.17)	0.365	0.711	a

a. Chancellor and colleagues[8] and Briggs and colleagues.[5]
b. Kauf and colleagues[10] and Tsevat and colleagues.[11]
c. Assumed to only apply to the first 2 years of the model (consistent with Chancellor and colleagues[8]).

PAS for HIV example

- Performed 10,000 simulation draws

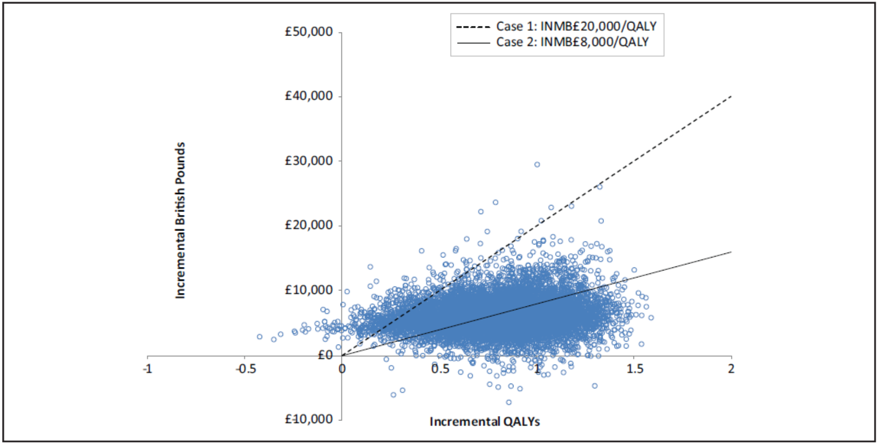
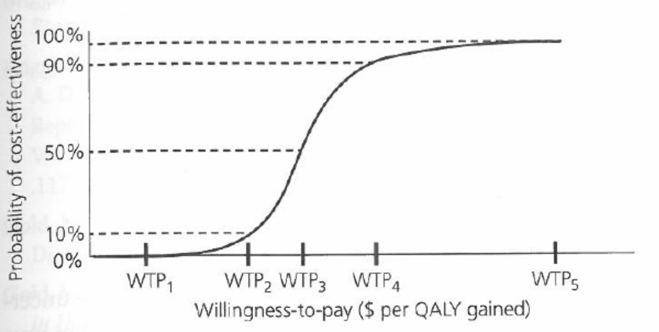


Figure 1 *Cost-effectiveness plane with 10,000 Monte Carlo draws and willingness-to-pay lines for preliminary model Case 1 (λ = £20,000/QALY) and Case 2 (λ = £8,000/QALY).*

Cost-effectiveness acceptability curve (from M&B, 2016)

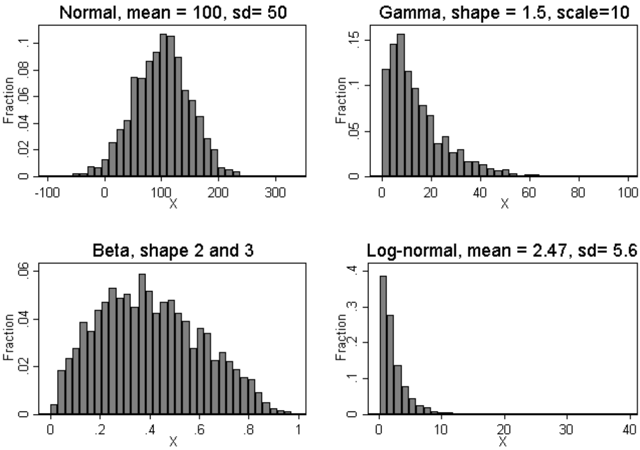


- For a given threshold (WTP), it shows the proportion of time the intervention would be considered cost-effective

Choosing probability distributions

- The normal distribution is the first one you learn about but in CEA few parameters distribute normal
- Costs, for example, cannot distribute normal because they have to be > 0 and costs are usually skewed
- Utilities and probabilities must be bounded between 0 and 1
- Common distributions in PSA are Gamma, Beta, and Log-normal

Example distributions



- See Stata code for simulations

Characteristics

- **Gamma**: $\gamma(k, \theta)$ has two parameters (shape and scale), support from > 0 to infinity. Used to simulate costs because costs are positive and skewed
- **Beta**: $\beta(a, b)$ has two shape parameters, support from 0 to 1. Good for simulating utilities or probabilities because they are bounded between 0 and 1
- **Log-normal**: the exponent of a normal distribution $exp(N(\mu, \sigma^2))$. Support from 0 to infinity
- Note: support is also called the **domain** of a distribution; the range of values for which the distribution is defined

How to simulate numbers

- Most statistical packages (SAS, Stata, R, Excel) have random numbers generators for common distributions
- **Inverse Transform Method**: It is possible to simulate numbers from any (continuous) distribution by using a uniform random generator
- We only need 1) a random generator for the uniform distribution and 2) the formula of the inverse cumulative distribution (may not exist in closed form)
- See Excel file for examples

Other ways of presenting and analyzing PSA

- Plots of the cost-effectiveness plane and acceptability curves are common ways to present results of PSA
- But we may want to **quantify** the **relative effect of each parameter** on the uncertainty of the ICER (or incremental cost or incremental benefit)
- One approach: run a regression with ICER as the dependent variable and different values of the inputs as predictors to understand the relative contribution of each input
- This is called **ANCOVA** (Analysis of Covariance); present proportion of sum of squares
- Another one: **Expected Value of Perfect Information (EVPI)**. An attempt to quantify the value of perfect information, or the opportunity cost of uncertainty

ANCOVA

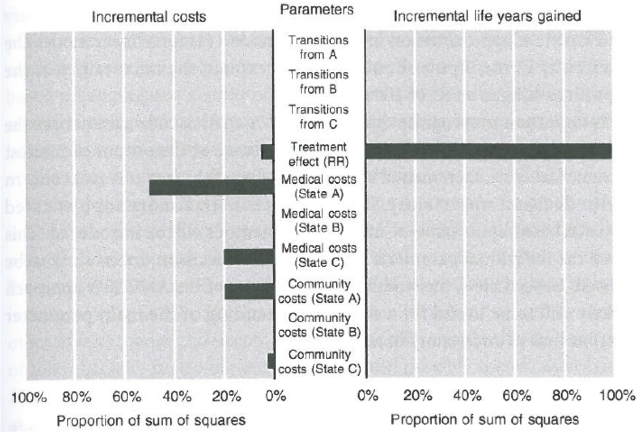


Fig. 5.5 ANCOVA analysis of proportion of sum of squares for incremental cost (left-hand side) and incremental life years gained (right-hand side) explained by uncertainty in the model input parameters.

- From Briggs et al (2006)

Summary

- We have used Markov models as examples, but everything we have learned apply to decision trees
- Sensitivity analyses are vital parts of CEA
- PSA may be complicated to implement but it is very easy to understand