



BANFF, CANADA September 7–11, 2014

# The International School on Research Impact Assessment

# Research Impact Assessment Designs

**Steven Wooding,**  
RAND Europe  
September 9<sup>th</sup>, 2014

Hosted by:



In partnership with:



# Learning Outcomes

- Appreciate the variety of evaluation designs
- Understand strengths and weaknesses of different selection frameworks
- Be able to apply a design that is appropriate for a RIA



# Types of inference 1

- Evaluation is about figuring out
  - whether x caused y
  - and if so
    - how x caused y
    - and how much y was caused by x
- There are different conceptual models for causal inference
- There doesn't seem to be a standardised vocabulary, edge cases are grey



# Types of inference 2

- Experimental/comparative
  - When x happens y always happens even when other things change
    - randomised trials
    - natural experiments
    - case control
    - pre/post
    - econometrics
- Theory based evaluation
  - Understanding the process by which x causes y
    - case studies
    - expert interviews
- Either
  - time series



# Comparison

	Theory based	Experimental
Internal validity (resistance to bias)	Low - subject to preconceptions of investigator/experts	High - can overcome bias by testing against data
Data requirements	Few cases understood in detail	Many cases matched for important characteristics and diverse in others
External validity (ability to generalise to other contexts)	Strong external validity - can compare other situations against evaluated context	Weak - unclear which are most important factors of success

# Example evaluations



BANFF, CANADA 2014

**The International School** on Research Impact Assessment

# The NIHR Leadership programme evaluation

## Theory of change

- Aimed to understand the role of management and leadership training in supporting the faculty of the NIHR
- Used literature review, interviews, and survey to understand how the programme was working
- By understanding how the programme seems to working will make recommendations for its improvement



# The Impact of Arthritis Research

## Natural experiment\comparative

- Used purposively selected case studies to compare success stories of different modes of funding
  - project
  - programme
  - fellowship
  - institute
- Showed contributions of different types of funding - eg valuable contribution of project funding
- But through narrative provided recommendations for improving funding process





# Showcase evaluation

## Case control

- The Wellcome Trust's Showcase scheme aimed fund high risk, high return research
- Project descriptions compared against control sample of normal project grants of similar size
- Project descriptions re-written to make them 'scheme agnostic' and reviewed applicants to ensure accuracy
- Showcase grants perceived to be more 'risky', 'novel', 'speculative', 'adventurous' and 'innovative' by expert panel members



# Unit of analysis

- The things you are looking at
- May be more than one type



# Comprehensive

- “Pick everything”
- Example
  - NIHR leadership evaluation survey
    - Aimed to survey every participant in the scheme
    - Tailored surveys for different seniority levels



# Purposive

- “Select interesting ones”
- Example
  - Evaluation of Health Technology Assessment programme of NIHR
    - Wanted to understand examples where new/existing treatments had been shown to be ineffective
    - Wanted to understand the journey from finding to influence on guideline



# Stratified random

- “Assign to groups and then select randomly”
  - **Mental Health Retrosight**
    - Collated all papers in mental health
    - Stratified by research type, country and number of citations

		Canada	UK	USA
Basic	With “schiz*”		1	2
	Without “schiz*”	2	1	
Clinical		2	2	2
Interventional	Biological	2		
	Psychosocial		1	1
	Health services/service delivery		1	1

# Random

- “Dice rolling”



# Tradeoffs in selection

## Structured Purposive

## Stratified

## Random

Can ensure representation of important but uncommon factors



Compensates for unknown factors

Ensures representation



Can be unbalanced

Complex



Easy

Requires information on candidates



Does not require information

Can be directed at hypothesis testing



Exploratory

Potentially biased



Unbiased



# Learning Activity



- Think of an evaluation you, or your organisation has done.
- Classify it into a design type: experimental, TBE, other
- Classify the sampling method(s) used
- Discuss in the group to see if patterns emerge
- 15 minutes





# Key Messages

- Different design approaches have different strengths and weaknesses
- There is lots of ambiguity about what 'research design' encompasses
- Different sampling methods are appropriate for different contexts



# Recommended Reading

- Stern, E., Stane, N., Mayne, J., Forss, K., Davies, R. & Befani, B. 2012. Broadening the range of designs and methods for impact evaluations. *Working Papers* – 38.



# Questions?

Steven Wooding

RAND Europe

@drstevenwooding



## Structured Purposive

## Stratified

## Random

Can ensure representation of important but uncommon factors



Compensates for unknown factors

Ensures representation



Can be unbalanced

Complex



Easy

Requires information on candidates



Does not require information

Can be directed at hypothesis testing



Exploratory

Potentially biased



Unbiased