Evidence-based Decision Making: Session 11

Rui Mata, FS 2022

Efficacy isn't everything…

Relevance	To what extent are the programme objectives justified in relation to needs?
Efficiency	Have the objectives been achieved at the lowest cost?
Effectiveness	To what extent has the outcome been achieved?
Sustainability	Are the results and impacts, including institutional changes, durable over time?
Impact	Are the results still evident after the intervention is completed?

World Health Organization (2013). ***WHO evaluation practice handbook***.

Efficacy isn't everything…

Table 1 Characteristics of Efficacy vs. Effectiveness Trial Designs (after [8])

	Efficacy Trial	Effectiveness Trial
Validity Priority	Internal > External	External ≥ Internal
Population and Sample	• Highly selected for condition of interest, narrowly defined	• Selected for condition of interest, reflecting presentation in source population
	• Few comorbidities • Willing and motivated participants	• Comorbidities resemble those in population to which results will be applied; only those who cannot practically or ethically participate are excluded
Intervention	• Intervention staff are highly qualified • Training may be intensive • Fidelity monitoring may be similarly intensive	• Staff selection, training, and fidelity monitoring resemble those likely to be feasible in target sites outside of the protocol proper
Outcome Measures and Data Collection	• Outcome measurements can be extensive, casting a wide net for potential secondary effects, moderators and mediators, or adverse effects	• Outcome batteries minimize respondent burden (in terms of both frequency and length of assessments) since subjects are heterogeneous in their willingness and capability to participate
	• Since subjects are motivated, respondent burden less of a concern	• Accordingly, outcome measures chosen carefully to target fewer outcomes, and must be simple to complete
Data Analysis	• Standard statistical approaches suffice, and data-intensive analyses may be feasible	• Analyses to account for greater sample heterogeneity • Analyses account for more missing data and data not missing at random

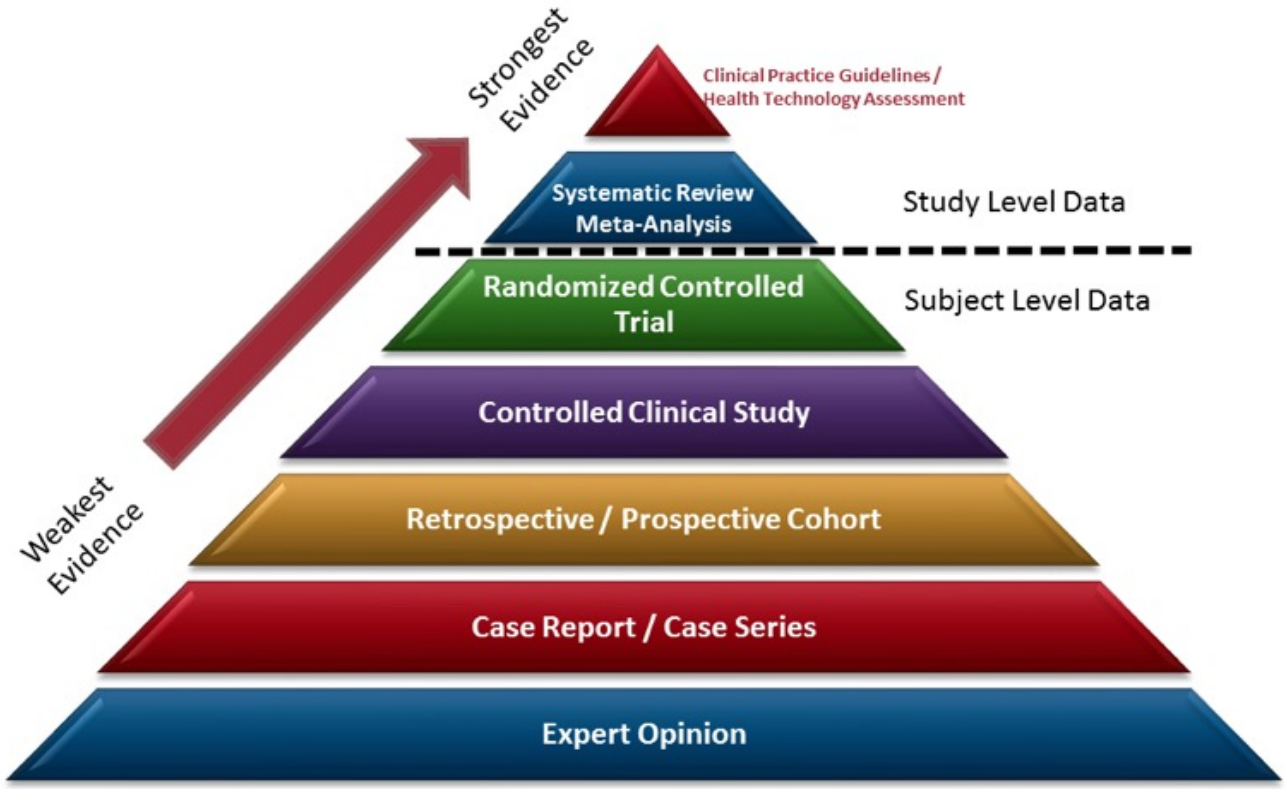
Bauer, M. S., Damschroder, L., Hagedorn, H., Smith, J., & Kilbourne, A. M. (2015). An introduction to implementation science for the non-specialist. BMC Psychology, 3(1), 65–12. http://doi.org/10.1186/S40359-015-0089-9

Efficacy isn't everything…

Table 2 Types of Studies to Address Blockages in the Implementation Process

Implementation Process Gap	Types of Studies
Limited external validity of efficacy/effectiveness studies	• Design clinical interventions ready for implementation earlier in the research pipeline, emphasizing tools, products, and strategies that mitigate variations in uptake across consumer, provider, and or organizational contexts
Quality gaps across systems due to variations in organizational capacity (e.g., resources, leadership)	• Assess variations and customize implementation strategies based on organizational context
	• Data infrastructure development to routinely capture or assess implementation fidelity, patient-level processes/outcomes of care, and value/return-on-investment measures
	• Further refinement of implementation strategies involving organizational and/or provider behavior change
	• Development of provider/practice networks to conduct implementation studies or evaluation of national programs
Frontline provider competing demands (e.g., multiple clinical reminders)	• Refinement of implementation strategies using cross-disciplinary methods that address provider behavior/organizational change (e.g., business, economics, policy, operations research. etc.)
	• Positive deviation or adaptation studies especially to improve implementation at lower-resourced, later-adopter sites
Misalignment with national or regional priorities	• National policy/practice roll-outs
	• Randomized evaluations of national programs or policies

Bauer, M. S., Damschroder, L., Hagedorn, H., Smith, J., & Kilbourne, A. M. (2015). An introduction to implementation science for the non-specialist. BMC Psychology, 3(1), 65–12. http://doi.org/10.1186/S40359-015-0089-9


Strongest Evidence
Clinical Practice Guidelines / Health Technology Assessment
Systematic Review Meta-Analysis
Study Level Data
Randomized Controlled Trial
Subject Level Data
Controlled Clinical Study
Retrospective / Prospective Cohort
Case Report / Case Series
Expert Opinion
Weakest Evidence

Teaching evaluation

1. Via Transaktionsnummer: Q1YF1. Eingabe unter: https://k11331.evasys.de/evasys/online/.
2. Oder via Direktlink: https://k11331.evasys.de/evasys/online.php?pswd=Q1YF1.

#	Date	Topic	Slides	Questions
1	02.03.2021	Introduction	pdf	1
2	09.03.2021	The Scientific Method(s)	pdf	2
3	16.03.2021	Algorithms	pdf	5
4	23.03.2021	Algorithms	pdf	
5	30.03.2021	Consensus	pdf	3
6	06.04.2021	Consensus	pdf	
7	13.04.2021	Counterfactuals	pdf	3
8	20.04.2021	Counterfactuals	pdf	
9	27.04.2021	Synthesis	pdf	3
10	04.05.2021	Synthesis	pdf	
11	11.05.2021	Interventions	pdf	3
12	18.05.2021	Interventions	pdf	
13	25.05.2021	Exam		20

Which of the following statements is/are correct?

A) Naomi Oreskes' views on science are aligned with the Popperian view that sees falsification as the single form of advancing science.

B) Naomi Oreskes argues that science derives its trustworthiness largely from its organized structure aimed at critically scrutinizing evidence.

C) Naomi Oreskes argues that science is objective because scientists always strive for an automated measurement of evidence that is free from personal opinion.

D) Naomi Oreskes argues that because even false theories can make correct predictions, there is a fundamental uncertainty about the truth-value of scientific theories.

Which of the following statements is/are correct?

A) Initial large-scale efforts to estimate replication rates of empirical findings in psychology suggest that less than half of studies can be replicated (i.e., Open Science Collaboration, 2015), an estimate even lower than the replication rate found for studies in economics (e.g., Camerer et al., 2016).

B) One main reason underlying low replicability may be HARKing, which stands for Hypothesizing After the Results are Known; the main instrument proposed to deal with this problem is increasing the quality of peer-review by journal editors (cf. Munafo et al., 2017).

C) "Registered Reports" seek to eliminate various forms of bias in hypothesis-driven research by splitting the peer-review process into two stages. At the first stage, prior to data collection, the study plan is made available on an online platform and the scientific community collectively provides suggestions for improvement of the design and planned analyses (crowdsourcing). At the second stage, reviewers and editors assess the crowdsourced protocol and analysis plan.

D) Some journals in psychology have started providing badges as an incentive to open science (open materials and data) but the percentage of articles reporting available data has stayed constant in such journals, suggesting that voluntary (non-coercive) strategies are not able to increase open science practices in psychology.

Which of the following statements is/are correct?

A) Regression is a statistical method often used in the behavioral sciences because it allows drawing strong inferences about causality from observational data.

B) The method of instrumental variables is sometimes used to estimate causal relationships when controlled experiments are not feasible. A valid instrumental variable induces changes in the explanatory variable but has no independent effect on the dependent variable, allowing a researcher to uncover the causal effect of the explanatory variable on the dependent variable using a regression model.

C) Regression discontinuity involves designs in which participants have been assigned to the intervention and the control conditions based on a cut-off score on a pre-intervention measure, for example, assessing merit or past performance.

D) The difference-in-differences approach makes the assumption that the change in outcomes from pre- to post-intervention in the control group is a good proxy for the counterfactual change in untreated potential outcomes in the treated group, which may not be true if the two groups differ in some important way.