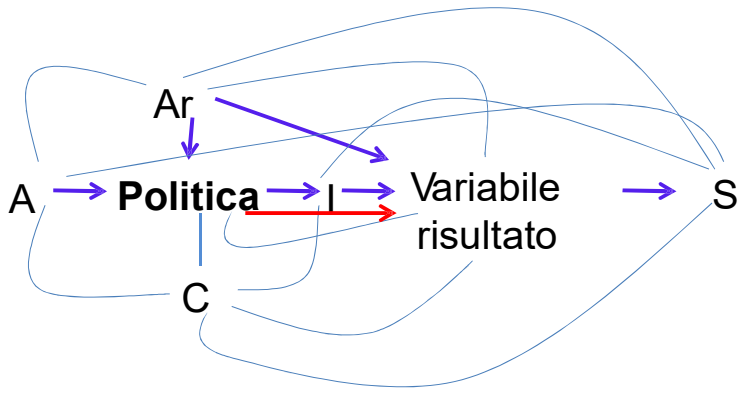
Progettazione e problemi dei metodi sperimentali

Manipolo la politica (X) per studiare il suo effetto sulla variabile risultato (Y)



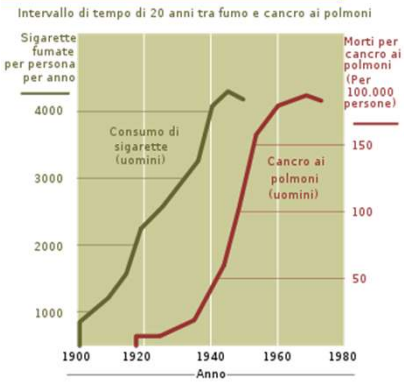
Three main approaches to causation:

- robust association → Regression analysis and similar tecniques
- generative process → Case Study

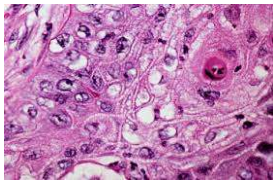
  (theory on actors, tested)
- consequential manipulation → RCT

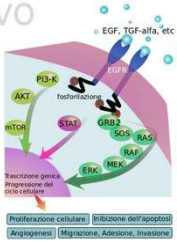
Tra le logiche di inferenza causale

Analisi di associazioni robuste



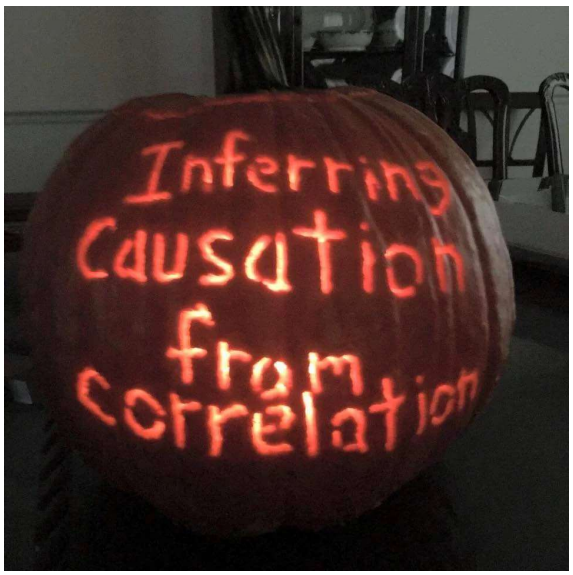
Identificazione del processo generativo





Manipolazione della variabile indipendente

In pratica, un esempio per avere il sentore dell'inferenza causale nei processi sociali…



What we observe:
Biased school guidance advice

Our true research question:
Does parental background affect teachers' advice?

The relevant research question:
How? Through which mechanisms?

Available Evidence

Ethnography

Secondary analysis on administrative data

Vignette study

Ethnography

Prof. E: lo sforzo più grosso che noi stiamo facendo è quello di riuscire a dare a ciascuno uno sbocco adeguato... quindi accanto al discorso delle attitudini, c'è poi comunque un discorso di ordine economico […] perché alcuni nostri studenti sono partiti già dalla prima con il bonus libri, per cui, alle elementari libri gratis, alle medie libri gratis, si troveranno alle superiori a pagare una barca di soldi per i libri! Poi tutte le richieste, con i laboratori eccetera. E bisogna stare attenti perché molti genitori non se ne rendono conto […] È brutto dirlo perché questo è discriminante, però va detto. Cioè io quando

Prof. B: capisci che un ragazzo che cresce in una certa famiglia, che sente parlare in un certo modo, dove i valori sono, come dire, di un certo livello. Cioè capisci che la sua struttura è quella per cui non puoi impegnarlo in studi filosofici, per esempio, non puoi pensare che vada a un liceo a studiare latino, greco, filosofia, perché è un mondo che... non gli appartiene. Non riuscirebbe, farebbe una fatica, non ha una situazione a casa che lo può supportare… anche se magari lo vorrebbe tanto perché, non so, i suoi amici vanno lì, lui ha in mente che vuole fare l'Università. Quindi quello che noi cerchiamo di fare con il consiglio orientativo è anche di fargli capire che magari lì sarebbe un pesce fuor d'acqua. [Estratto di intervista]

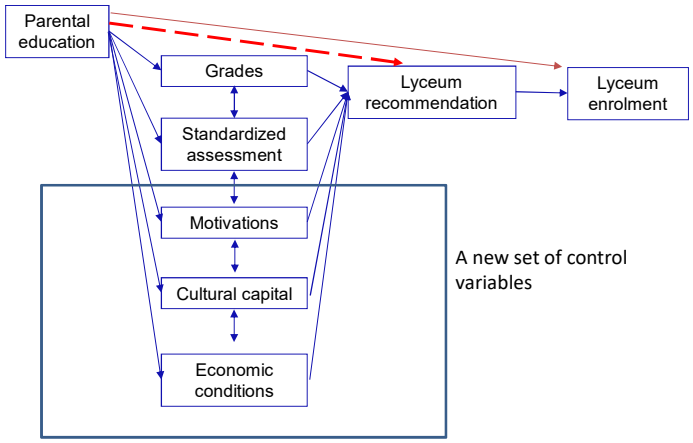
(Romito 2014)

Ethnography

Prof. B: capisci che un ragazzo che cresce in una certa famiglia, che sente parlare in un certo modo, dove i valori sono, come dire, di un certo livello. Cioè capisci che la sua struttura è quella per cui non puoi impegnarlo in studi filosofici, per esempio, non puoi pensare che vada a un liceo a studiare latino, greco, filosofia, perché è un mondo che... non gli appartiene. Non riuscirebbe, farebbe una fatica, non ha una situazione a casa che lo può supportare… anche se magari lo vorrebbe tanto perché, non so, i suoi amici vanno lì, lui ha in mente che vuole fare l'Università. Quindi quello che noi cerchiamo di fare con il consiglio orientativo è anche di fargli capire che magari lì sarebbe un pesce fuor d'acqua. [Estratto di intervista]

(Romito 2014)

Secondary analyses



Secondary analyses

		performance	motivations	cultural capital	economic conditions
Lyceum	Model 0	Model 1	Model 2	Model 3	Model 4
Upper secondarty	19.10***	8.537***	8.428***	7.964***	5.996***
	(0.755)	(0.673)	(0.673)	(0.677)	(0.718)
Tertiary	35.09***	14.23***	14.25***	13.10***	9.970***
	(0.868)	(0.802)	(0.802)	(0.832)	(0.921)
Classic/ Scientific Lyceum	Model 0	Model 1	Model 2	Model 3	Model 4
Upper secondarty	15.65***	3.740***	3.678***	3.071***	2.096***
	(0.756)	(0.642)	(0.642)	(0.646)	(0.685)
Tertiary	38.79***	15.17***	15.25***	13.66***	11.61***
	(0.868)	(0.765)	(0.765)	(0.794)	(0.879)

Vignette study

Simulazione_intro.
Una simulazione di orientamento alla scelta

Nella pagina seguente le verrà presentato un caso ipotetico di orientamento.

Sappiamo che per formulare un consiglio orientativo si valutano moltissimi elementi e informazioni. Le chiediamo però di esprimere comunque la sua opinione sulla base dei dati che le saranno forniti.

Legga con attenzione e memorizzi queste informazioni, anche prendendosi degli appunti se crede. Le servirà ricordare bene il caso proposto per le domande nelle pagine successive.

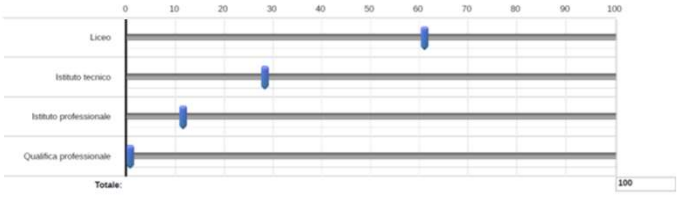
M_IT. **Lo studente P. L. è nato in Italia.**

Status_basso. **La madre lavora come cassiera in un supermercato, mentre il padre è operaio in piccola impresa di calzature.**

Int_equilibrati. **Va mediamente bene a scuola, con voti intorno a 7 e mezzo, e ha una discreta voglia di studiare. Non sa cosa scegliere dopo la scuola secondaria di primo grado e ha lo stesso interesse per le discipline tecnico/scientifiche e per quelle umanistiche, cavandosela altrettanto bene in entrambe.**

Vignette study

Consiglio. Alla luce delle informazioni disponibili, quanto è probabile che Lei raccomandi a P.L. e ai suoi genitori i seguenti percorsi? Esprima la sua scelta in percentuale: muova i cursori in modo da trovare la combinazione di probabilità che meglio rispecchia la sua opinione. Il totale dovrà fare 100.



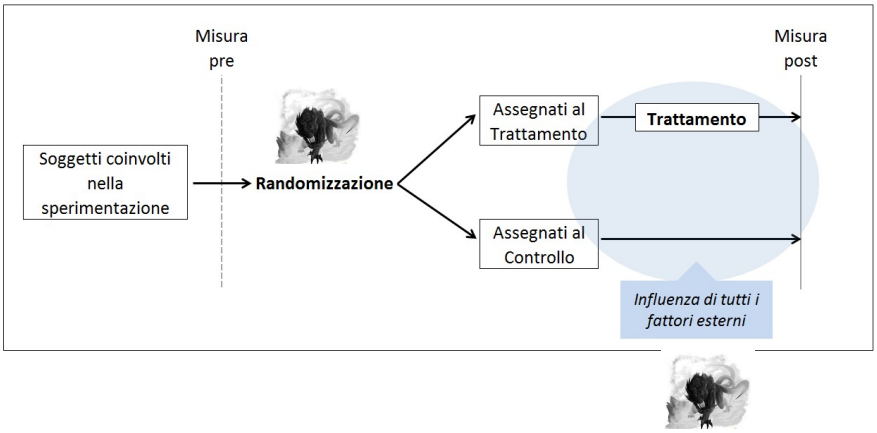
Vignette study

Consiglio_1	Coefficient	Std. err.	t	P>|t|	[95% conf.	interval]
origini						
Cmi	2.305906	1.279933	1.80	0.072	-.203865	4.815676
Bor	20.62375	1.280003	16.11	0.000	18.11384	23.13366
sesso						
F	2.327018	1.041875	2.23	0.026	.2840452	4.36999
nazionalita						
II gen	-1.420322	1.285939	-1.10	0.269	-3.941871	1.101226
I gen	-3.487592	1.275765	-2.73	0.006	-5.989191	-.9859933
interesse						
TecSc	-10.00445	1.278853	-7.82	0.000	-12.5121	-7.496796
Uguale	-7.553311	1.274927	-5.92	0.000	-10.05327	-5.053355
_cons	38.26929	1.481354	25.83	0.000	35.36456	41.17402

Perché gli RCT?

"IL" metodo, quello che preferiamo

"Gold standard"



ATTENZIONE alla polisemia del termine!

"SPERIMENTAZIONE"

"USUALE"

CONTROLLATA
Randomized Controlled Trial (RCT)

Ulteriore polisemia: Efficacy versus Effectiveness trial

- RCT in forma di interventi pilota
- RCT per stimare l'effetto di una politica a scala

Portare le politiche a scala riduce la fedeltà di implementazione e porta quindi a effetti diversi (inferiori) rispetto a quelli delle *super-realization* tipiche degli studi pilota!

Come progettare e realizzare un RCT

FASE 0 – verifica delle condizioni per fare una sperimentazione controllata (1 di 2)

- Siamo arrivati **prima dell'avvio** dell'intervento? In tempo per gestire il processo?
- La randomizzazione è eticamente, tecnicamente e politicamente **accettabile** (genuina incertezza, limitatezza risorse, dilazione intervento, esclusione praticabile)?
- Il **protocollo** dell'intervento è ben definito? In base alle risorse economiche e umane, può essere implementato in modo ragionevolmente fedele a quanto previsto?
- È possibile definire pochi **outcome** misurabili (primario e secondari)?

FASE 0 – verifica delle condizioni per fare una sperimentazione controllata (2 di 2)

- Posso scegliere il target per avere **maggiore validità esterna**?
- Il target è definito e si avrà un **numero adeguato** di soggetti trattabili, tale da consentire esclusione e garantire potenza statistica adeguata?
- Ci possiamo attendere **compliance** adeguata per trattati e controlli?
- Esiste il rischio di **substitution** per i controlli?
- **Nel tempo**, potremo tracciare trattati e controlli? Abbiamo risorse sufficienti per rilevare gli outcome?

FASE 1 – reperimento dei potenziali trattati

Necessità di sovra enrolment

Chiarezza nella successiva assegnazione casuale – punto spinoso per la validità esterna con importanti implicazioni etiche

Due scenari: opt in – opt out

Disegno e registrazione dell'esperimento

Come pre-registrare registrare un RCT

https://www.socialscienceregistry.org/

FASE 2 – randomizzazione

Come è possibile che la randomizzazione generi equivalenza?

Un parallelo tra randomizzazione e campionamento può aiutare a capire…

- ricordate gli intervalli di confidenza?
- ricordate come interpretare un p-value?

Quali relazioni tra n, variabilità e incertezza?

Come fare una randomizzazione e come controllare il balancing

Effect size, MDES e calcolo di potenza

Metrica di stima degli effetti: valori assoluti e valori relativi

Effect size come misura relativa alla devizione standard dei controlli

Minimum Detectable Effect Size, ovvero la questione della **potenza statistica** del nostro esperimento

(vogliamo evitare gli errori di tipo II, ovvero dire che non c'è un effetto quando invece c'era)

Potenza statistica e margini di manovra

N e variabilità del"outcome sono dati

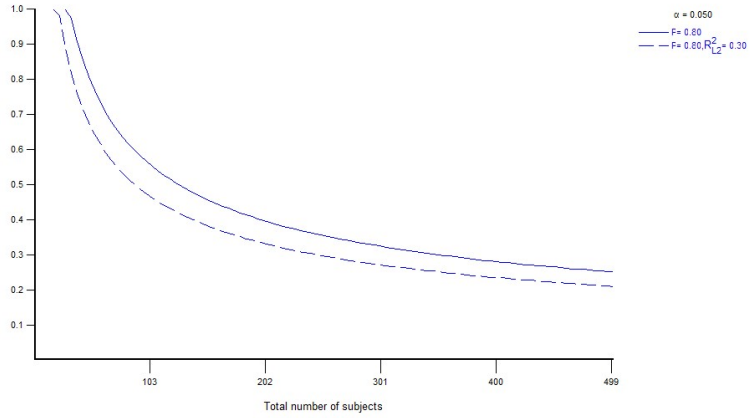
La clusterizzazione del trattamento è anche spesso data e fuori dal nostro controllo

Possiamo operare su

- misura pre-intervento
- blocchi di randomizzazione (stratificazione)

Esempio di tool per il calcolo di potenza

Optimal design



Esempio di tool per il calcolo di potenza

Power up!

B	C	D
Model 1.0: MDES Calculator for Individual Random Assignment (IRA) Designs—Completely Randomized Controlled Trials		
Assumptions		**Comments**
Alpha Level (α)	0,05	Probability of a Type I error
Two-tailed or One-tailed Test?	1	
Power (1-β)	0,80	Statistical power (1-probability of a Type II error)
P	0,50	Proportion of the sample randomized to treatment: n_T / (n_T + n_C)
R^2	0,30	Percent of variance in outcome explained by covariates
k*	5	Number of covariates used
n (Total Sample Size)	500	
M (Multiplier)	2,49	Computed from T_1 and T_2
T_1 (Precision)	1,65	Determined from alpha level, given two-tailed or one-tailed test
T_2 (Power)	0,84	Determined from given power level
MDES	0,186	Minimum Detectable Effect Size

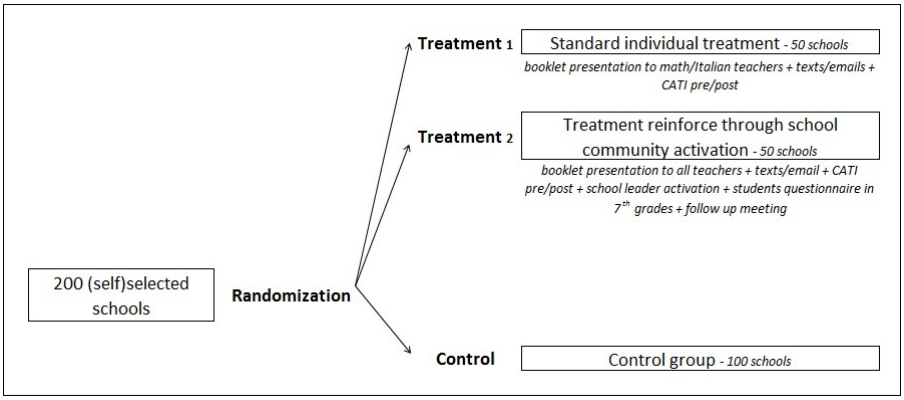
Note: The parameters in the yellow cells need to be specified. The MDES will be calculated automatically.

Come calcolare la Potenza

https://www.causalevaluation.org/power-analysis.html

Il disegno a due braccia è solo uno tra quelli possibili!

Utile per «spacchettare» interventi complessi o per studiarne diverse modalità di implementazione!



FASE 3 – intervento e suo monitoraggio

Fedeltà nella fase di implementazione

Compliance allo stato assegnato

Monitoraggio di nterventi concorrenti sui controlli

Contaminazione – spillover

assunto **SUTVA**

(Stable Unit Treatment Value Assumption)

La risposta di un individuo al trattamento non dipende dalle risposte degli altri individui allo stesso trattamento

FASE 4 – misurazione outcome

Outcome rilevanti osservabili

Outcome principale e outcome secondari

Tempo: manipulation check, outcome prossimi e outcome finali

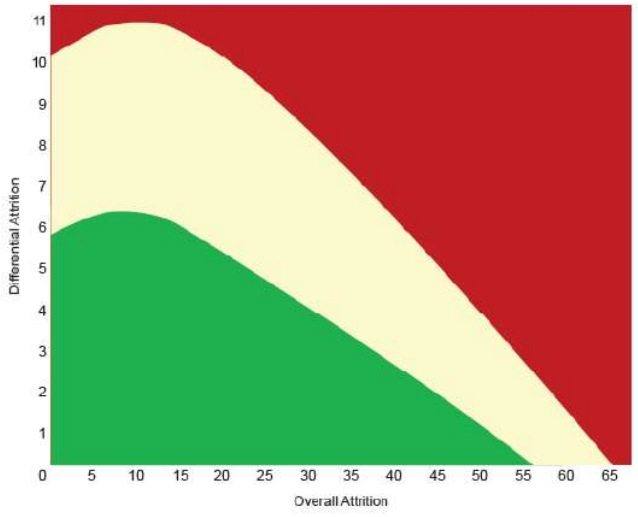
Blindness rispetto all'osservatore – rischio *yes-saying*

FASE 5 – analisi con stima dell'effetto

Check balancing
Check attrition
Confronto di medie
Confronto con controllo unbalance o blocchi
Stime ITT e TOT
Analisi di eterogeneità dell'effetto
(occhio al fishing for effects!)

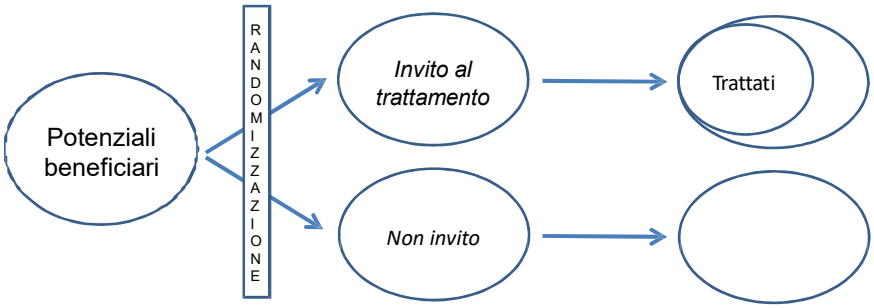
FASE 5 – attrition differenziale

FIGURE A.1. TRADEOFFS BETWEEN OVERALL AND DIFFERENTIAL ATTRITION



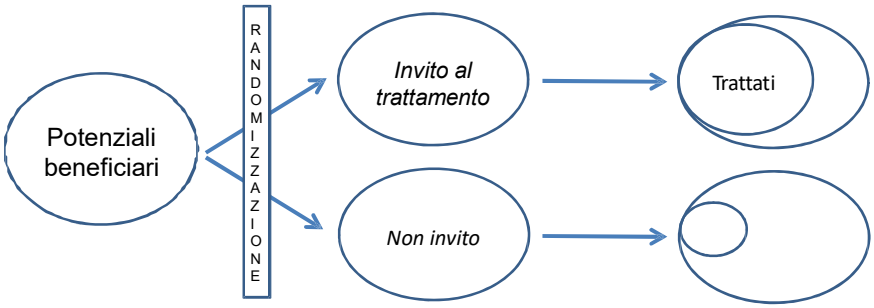
Analisi con compliance parziale dei trattati

Situazione tipica di moltissimi esperimenti…
stima ITT



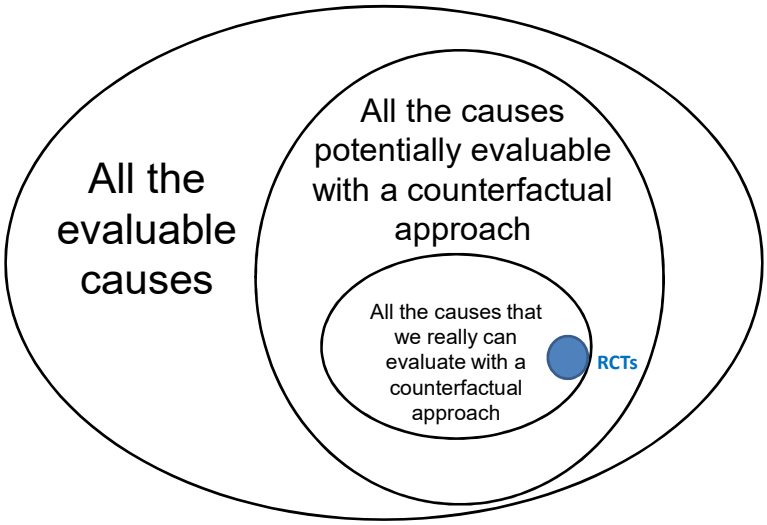
Se non possiamo randomizzare, possiamo provare a randomizzare l'invito

Confrontiamo due gruppi identici,
uno ha ricevuto più intensamente dell'altro il trattamento
perchè invitato efficacemente a farlo.



Ostacoli e problemi

Fattibilità



Additional limitations

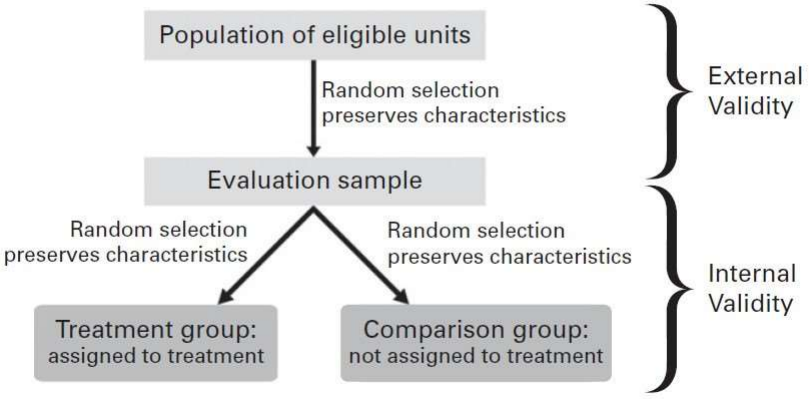
- Ethical objections
- Difficulties to reduce social complexity, due to
  - reverse causality
  - unexpected effects
  - relevance of local contexts and impacts' heterogeneity among subgroups
- Several causes cannot be (effectively) manipulated through interventions
- Costs and time
- Risk of failure

Light touch interventions and administrative data

41

La questione validità esterna!

Figure 4.2 Random Sampling and Randomized Assignment of Treatment



RCTs risk to produce weak evidence for several reasons…

- External validity (samples and contexts)

The need for a proper design

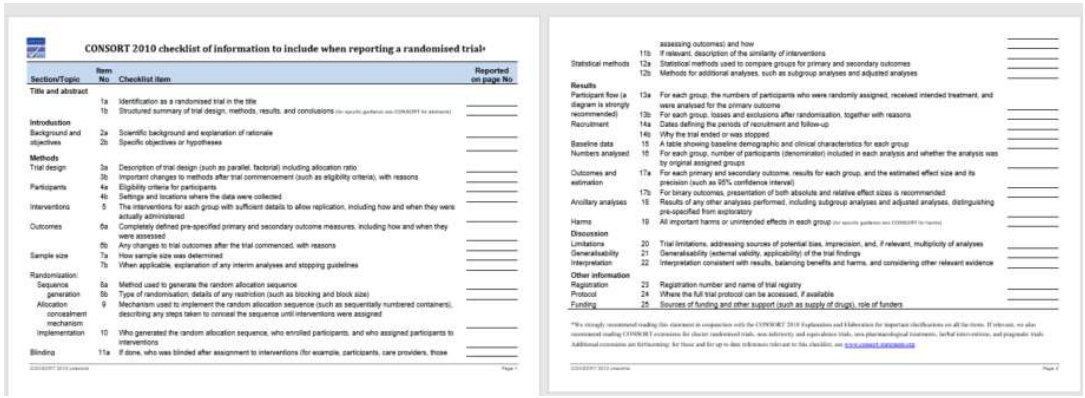
- Under-powered experiment
- Sloppy/Super implementation
- Pre-specification

And preserving randomization…

- Non compliance, contamination, substitution
- (differential) Attrition
- External observer

Quale qualità degli RCTs in Italia (nel campo dell'educazione)?

https://sciendo.com/it/article/10.2478/sjs-2022-0004



CONSORT 2010 checklist of information to include when reporting a randomised trial*

Section/Topic	Item No	Checklist item	Reported on page No
Title and abstract			
	1a	Identification as a randomised trial in the title	
	1b	Structured summary of trial design, methods, results, and conclusions (for specific guidance see CONSORT for abstracts)	
Introduction			
Background and objectives	2a	Scientific background and explanation of rationale	
	2b	Specific objectives or hypotheses	
Methods			
Trial design	3a	Description of trial design (such as parallel, factorial) including allocation ratio	
	3b	Important changes to methods after trial commencement (such as eligibility criteria), with reasons	
Participants	4a	Eligibility criteria for participants	
	4b	Settings and locations where the data were collected	
Interventions	5	The interventions for each group with sufficient details to allow replication, including how and when they were actually administered	
Outcomes	6a	Completely defined pre-specified primary and secondary outcome measures, including how and when they were assessed	
	6b	Any changes to trial outcomes after the trial commenced, with reasons	
Sample size	7a	How sample size was determined	
	7b	When applicable, explanation of any interim analyses and stopping guidelines	
Randomisation:			
Sequence generation	8a	Method used to generate the random allocation sequence	
	8b	Type of randomisation; details of any restriction (such as blocking and block size)	
Allocation concealment mechanism	9	Mechanism used to implement the random allocation sequence (such as sequentially numbered containers), describing any steps taken to conceal the sequence until interventions were assigned	
Implementation	10	Who generated the random allocation sequence, who enrolled participants, and who assigned participants to interventions	
Blinding	11a	If done, who was blinded after assignment to interventions (for example, participants, care providers, those assessing outcomes) and how	
	11b	If relevant, description of the similarity of interventions	
Statistical methods	12a	Statistical methods used to compare groups for primary and secondary outcomes	
	12b	Methods for additional analyses, such as subgroup analyses and adjusted analyses	
Results			
Participant flow (a diagram is strongly recommended)	13a	For each group, the numbers of participants who were randomly assigned, received intended treatment, and were analysed for the primary outcome	
	13b	For each group, losses and exclusions after randomisation, together with reasons	
Recruitment	14a	Dates defining the periods of recruitment and follow-up	
	14b	Why the trial ended or was stopped	
Baseline data	15	A table showing baseline demographic and clinical characteristics for each group	
Numbers analysed	16	For each group, number of participants (denominator) included in each analysis and whether the analysis was by original assigned groups	
Outcomes and estimation	17a	For each primary and secondary outcome, results for each group, and the estimated effect size and its precision (such as 95% confidence interval)	
	17b	For binary outcomes, presentation of both absolute and relative effect sizes is recommended	
Ancillary analyses	18	Results of any other analyses performed, including subgroup analyses and adjusted analyses, distinguishing pre-specified from exploratory	
Harms	19	All important harms or unintended effects in each group (for specific guidance see CONSORT for harms)	
Discussion			
Limitations	20	Trial limitations, addressing sources of potential bias, imprecision, and, if relevant, multiplicity of analyses	
Generalisability	21	Generalisability (external validity, applicability) of the trial findings	
Interpretation	22	Interpretation consistent with results, balancing benefits and harms, and considering other relevant evidence	
Other information			
Registration	23	Registration number and name of trial registry	
Protocol	24	Where the full trial protocol can be accessed, if available	
Funding	25	Sources of funding and other support (such as supply of drugs), role of funders	

*We strongly recommend reading this statement in conjunction with the CONSORT 2010 Explanation and Elaboration for important clarifications on all the items. If relevant, we also recommend reading CONSORT extensions for cluster randomised trials, non-inferiority and equivalence trials, non-pharmacological treatments, herbal interventions, and pragmatic trials. Additional extensions are forthcoming: for those and for up to date references relevant to this checklist, see www.consort-statement.org.

The review – inclusion criteria

- Experiments conducted (at least in part) in Italy
- Field: education (K-13)
- Type: field experiments
- Results publicly circulating during the period 2009-2020 (papers, book chapters, reports, conference presentations etc.)

The review – methodology

1. First draft of the list among the «community»
2. Scopus search by keyword
3. Blog “Studi randomizzati” and informal survey with colleagues https://studirandomizzati.wordpress.com

Final list: 26 studies

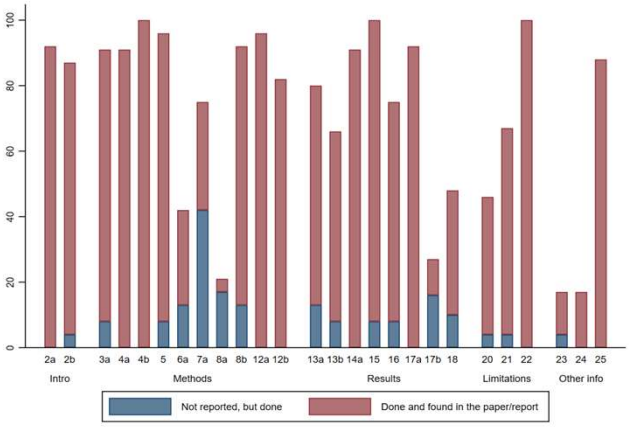
Coding

1. We asked the authors of the studies included in the review for any material about their experiments (also non-published ones)
2. We started to code a subset of studies indipendently, distinguishing elements that were found in the publicly available material to those available only in internal notes and reports
3. We met several times to compare our strategies in case of doubts
4. We forwarded each experiments’ grid to the authors, to validate our results and eventually to discuss changes

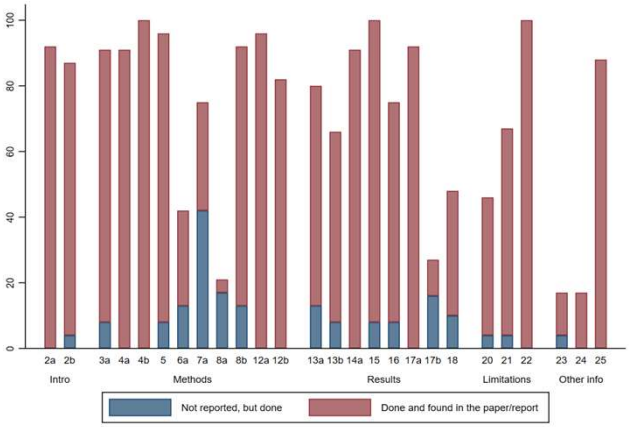
Description of the trials

- ❑All levels and all actors involved (as beneficiaries or as target)
- ❑Good level of geographical diffusion
- ❑Interventions acting on different theoretical levers; plurality of outcomes
- ❑Uniformity of designs (mostly school-centered)
- ❑Average good sample sizes (>1,000 students)
- ❑Driver for the evaluation: mostly coming from researchers (7/26: res. proposed PM to evaluate an intervention; 10/26: res. designed it theirselves)
- ❑Genuine demand for evaluation from PM rare (6/26); 3 times induced by funding regulations (e.g. EU grants)

Results

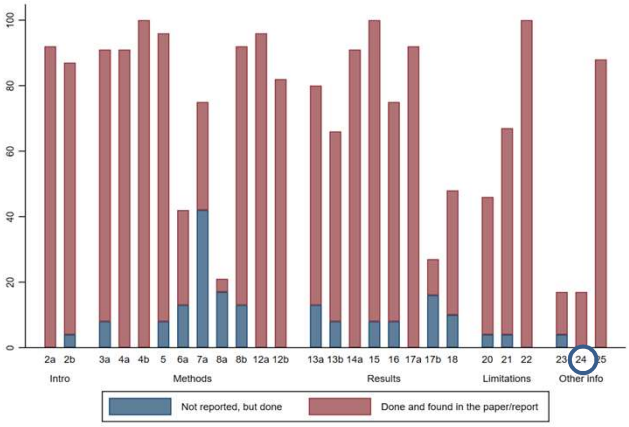


Results



1. Average high quality
2. Moderate variability
3. Rather high completeness

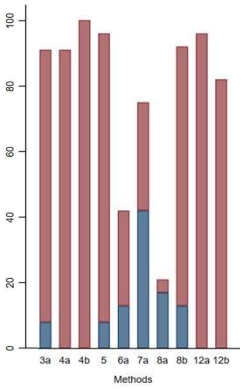
Results



1. Average high quality
2. Moderate variability
3. Rather high completeness

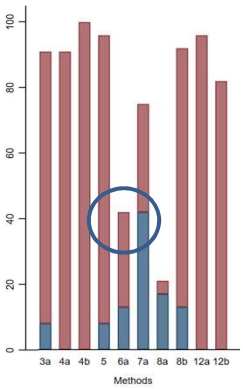
The foe: lack of pre-registration

Results - methods



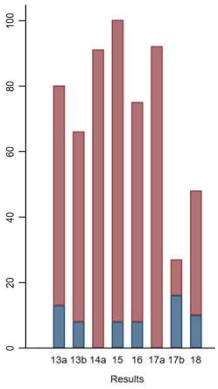
Trial design	3a	Description of trial design (such as parallel, factorial) including allocation ratio
Participants	4a	Eligibility criteria for participants
	4b	Settings and locations where the data were collected
Interventions	5	The interventions for each group with sufficient details to allow replication, including how and when they were actually administered
Outcomes	6a	Completely defined pre-specified primary and secondary outcome measures, including how and when they were assessed
Sample size	7a	How sample size was determined
Sequence generation	8a	Method used to generate the random allocation sequence
	8b	Type of randomisation; details of any restriction (such as blocking and block size)
Statistical methods	12a	Statistical methods used to compare groups for primary and secondary outcomes
	12b	Methods for additional analyses, such as subgroup analyses and adjusted analyses

Results - methods



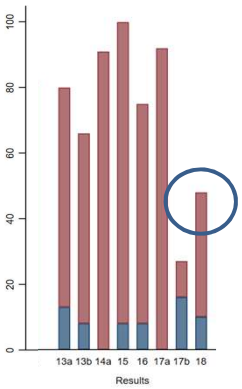
Trial design	3a	Description of trial design (such as parallel, factorial) including allocation ratio
Participants	4a	Eligibility criteria for participants
	4b	Settings and locations where the data were collected
Interventions	5	The interventions for each group with sufficient details to allow replication, including how and when they were actually administered
Outcomes	6a	Completely defined pre-specified primary and secondary outcome measures, including how and when they were assessed
Sample size	7a	How sample size was determined
Sequence generation	8a	Method used to generate the random allocation sequence
	8b	Type of randomisation; details of any restriction (such as blocking and block size)
Statistical methods	12a	Statistical methods used to compare groups for primary and secondary outcomes
	12b	Methods for additional analyses, such as subgroup analyses and adjusted analyses

Results - results



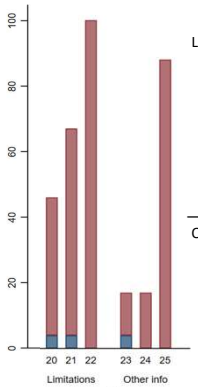
Participant flow	13a	For each group, the numbers of participants who were randomly assigned, received intended treatment, and were analysed for the primary outcome
	13b	For each group, losses and exclusions after randomisation, together with reasons
Recruitment	14a	Dates defining the periods of recruitment and follow-up
Baseline data	15	A table showing baseline demographic and clinical characteristics for each group
Numbers analysed	16	For each group, number of participants (denominator) included in each analysis and whether the analysis was by original assigned groups
Outcomes and estimation	17a	For each primary and secondary outcome, results for each group, and the estimated effect size and its precision (such as 95% confidence interval)
	17b	For binary outcomes, presentation of both absolute and relative effect sizes is recommended
Ancillary analyses	18	Results of any other analyses performed, including subgroup analyses and adjusted analyses, distinguishing pre-specified from exploratory

Results - results



Participant flow	13a	For each group, the numbers of participants who were randomly assigned, received intended treatment, and were analysed for the primary outcome
	13b	For each group, losses and exclusions after randomisation, together with reasons
Recruitment	14a	Dates defining the periods of recruitment and follow-up
Baseline data	15	A table showing baseline demographic and clinical characteristics for each group
Numbers analysed	16	For each group, number of participants (denominator) included in each analysis and whether the analysis was by original assigned groups
Outcomes and estimation	17a	For each primary and secondary outcome, results for each group, and the estimated effect size and its precision (such as 95% confidence interval)
	17b	For binary outcomes, presentation of both absolute and relative effect sizes is recommended
Ancillary analyses	18	Results of any other analyses performed, including subgroup analyses and adjusted analyses, distinguishing pre-specified from exploratory

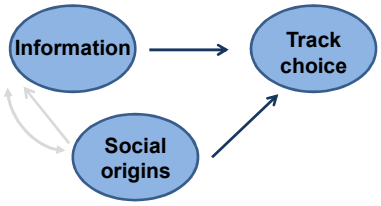
Results – limitations and other infos



Limitations	Limitations	20	Trial limitations, addressing sources of potential bias, imprecision, and, if relevant, multiplicity of analyses
	Generalisability	21	Generalisability (external validity, applicability) of the trial findings
	Interpretation	22	Interpretation consistent with results, balancing benefits and harms, and considering other relevant evidence
Other info	Registration	23	Registration number and name of trial registry
	Protocol	24	Where the full trial protocol can be accessed, if available
	Funding	25	Sources of funding and other support, role of funders

Example
ARTI Puglia

https://journals-sagepub-com.proxy.unimib.it/doi/full/10.1177/0001699317729872



The design

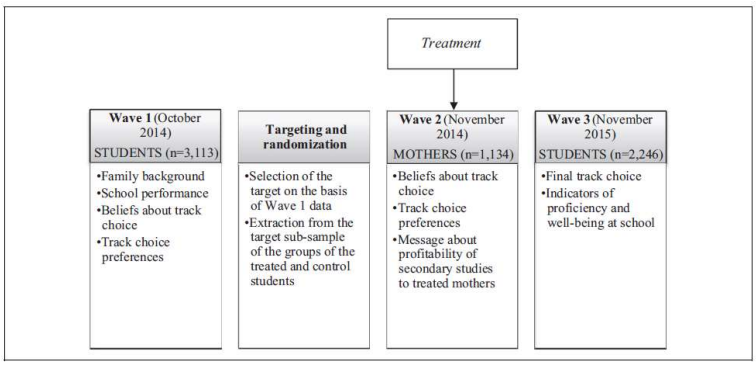


Figure 1. The structure of the longitudinal and experimental study.

Fonte: Barone et al 2018

60

The intervention

«Suo figlia/a ha un rendimento scolastico superiore a quello di chi s'iscrive al liceo e vi riesce senza difficoltà. Se dopo il diploma Suo/a figlio/a scegliesse di non proseguire all'università, una maturità liceale in Puglia offrirebbe comunque opportunità occupazionali simili a quelle degli altri indirizzi di studio».

61

ARTI Puglia: results

Table 3. Effects of the treatment on enrolment in the academic track (logit regression models, average marginal effects).

	Model 1[a]	Model 2[a]	Model 3[a]
Treated	0.065	0.071	0.101*
	(0.066)	(0.064)	(0.059)
Control variables			
Sampling strata	Yes	Yes	Yes
Background and demographics[b]	Yes	Yes	Yes
School Proficiency and attitudes[c]	No	Yes	Yes
Intentions	No	No	Yes
Observations	333	333	333
R-squared	0.101	0.122	0.203

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Notes: [a]Standard errors in parentheses are clustered at the school-level.

[b]Background and demographic controls include: gender, parental education (high school diploma vs less than high school diploma), age and nationality.

[c]School proficiency variables include performance in mathematics and Italian.

Fonte: Barone et al 2018

62

Ulteriori problem e limiti…

Back to ARTI's intervention

«Suo figlia/a ha un rendimento scolastico superiore a quello di chi s'iscrive al liceo e vi riesce senza difficoltà. Se dopo il diploma Suo/a figlio/a scegliesse di non proseguire all'università, una maturità liceale in Puglia offrirebbe comunque opportunità occupazionali simili a quelle degli altri indirizzi di studio».

Basic intervention ?

Several bits of information
performance, academic returns, labor market returns

Empowerment

(academic) Authority

64

Tirando le somme…

- Causal inference through correlational analysis is a common practice, but less and less convincing and less and less rewarding
- RCTs are a rigorous tool to face and estimate causal effects
- RCTs can be not so rigorous
- RCTs have limitations and require effort
- RCTs are risky scientific enterprises
- Light touch interventions and administrative data are an easy way to implement RCTs
- Even a well implemented experiment may be difficult to be interepreted (especially in the case of null results)

Il Vostro report

Struttura del report

Abstract strutturato

Background

Rassegna della letteratura e rilevanza esperimento: in che modo contribuisce a quanto già sappiamo?

Obiettivi

Interrogativo secondo schema PICO(S)

Il disegno della sperimentazione, con particolare attenzione alla descrizione della randomizzazione e di cosa è seguito

Risultati

Flow chart, bilanciamento, compliance, effetti e loro eterogeneità

Conclusioni

Sintesi dei risultati, lezioni apprese, limiti del lavoro + implicazioni per la ricerca futura e per il disegno delle politiche che seguiranno

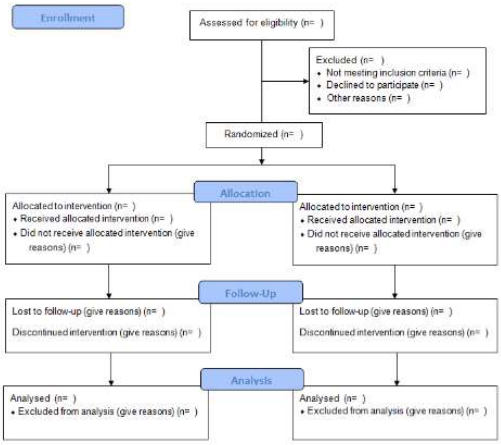
Come fare un buon esperimento

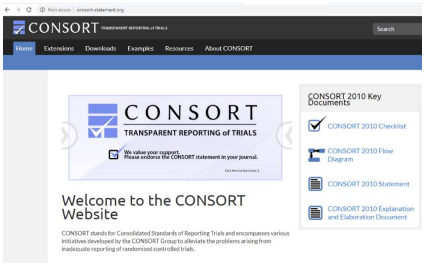
CONSORT 2010 checklist of information to include when reporting a randomised trial*

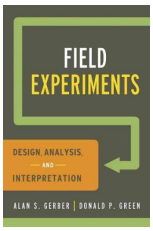
Section/Topic	Item No	Checklist item	Reported on page No
Title and abstract			
	1a	Identification as a randomised trial in the title	
	1b	Structured summary of trial design, methods, results, and conclusions (for specific guidance see CONSORT for abstracts)	
Introduction			
Background and objectives	2a	Scientific background and explanation of rationale	
	2b	Specific objectives or hypotheses	
Methods			
Trial design	3a	Description of trial design (such as parallel, factorial) including allocation ratio	
	3b	Important changes to methods after trial commencement (such as eligibility criteria), with reasons	
Participants	4a	Eligibility criteria for participants	
	4b	Settings and locations where the data were collected	
Interventions	5	The interventions for each group with sufficient details to allow replication, including how and when they were actually administered	
Outcomes	6a	Completely defined pre-specified primary and secondary outcome measures, including how and when they were assessed	
	6b	Any changes to trial outcomes after the trial commenced, with reasons	
Sample size	7a	How sample size was determined	
	7b	When applicable, explanation of any interim analyses and stopping guidelines	
Randomisation:			
Sequence generation	8a	Method used to generate the random allocation sequence	
	8b	Type of randomisation; details of any restriction (such as blocking and block size)	
Allocation concealment mechanism	9	Mechanism used to implement the random allocation sequence (such as sequentially numbered containers), describing any steps taken to conceal the sequence until interventions were assigned	
Implementation	10	Who generated the random allocation sequence, who enrolled participants, and who assigned participants to interventions	
Blinding	11a	If done, who was blinded after assignment to interventions (for example, participants, care providers, those assessing outcomes) and how	
	11b	If relevant, description of the similarity of interventions	
Statistical methods	12a	Statistical methods used to compare groups for primary and secondary outcomes	
	12b	Methods for additional analyses, such as subgroup analyses and adjusted analyses	
Results			
Participant flow (a diagram is strongly recommended)	13a	For each group, the numbers of participants who were randomly assigned, received intended treatment, and were analysed for the primary outcome	
	13b	For each group, losses and exclusions after randomisation, together with reasons	
Recruitment	14a	Dates defining the periods of recruitment and follow-up	
	14b	Why the trial ended or was stopped	
Baseline data	15	A table showing baseline demographic and clinical characteristics for each group	
Numbers analysed	16	For each group, number of participants (denominator) included in each analysis and whether the analysis was by original assigned groups	
Outcomes and estimation	17a	For each primary and secondary outcome, results for each group, and the estimated effect size and its precision (such as 95% confidence interval)	
	17b	For binary outcomes, presentation of both absolute and relative effect sizes is recommended	
Ancillary analyses	18	Results of any other analyses performed, including subgroup analyses and adjusted analyses, distinguishing pre-specified from exploratory	
Harms	19	All important harms or unintended effects in each group (for specific guidance see CONSORT for harms)	
Discussion			
Limitations	20	Trial limitations, addressing sources of potential bias, imprecision, and, if relevant, multiplicity of analyses	
Generalisability	21	Generalisability (external validity, applicability) of the trial findings	
Interpretation	22	Interpretation consistent with results, balancing benefits and harms, and considering other relevant evidence	
Other information			
Registration	23	Registration number and name of trial registry	
Protocol	24	Where the full trial protocol can be accessed, if available	
Funding	25	Sources of funding and other support (such as supply of drugs), role of funders	

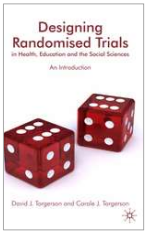
28

Consort flow chart

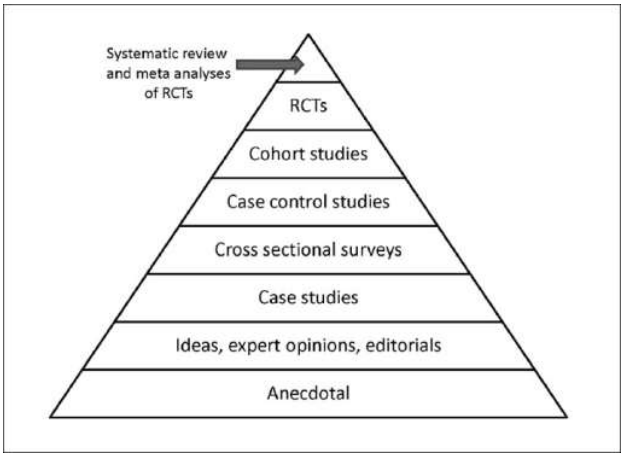








Beyond RCTs



Fonte: Mhaskar et al 2009

CENNI SULLA SYSTEMATIC REVIEW

Forme di evidenza (gerarchizzabili)

Hierarchy of evidence
I-I Systematic review and meta-analysis of two or more double blind randomized controlled trials.
I-2 One or more large double-blind randomized controlled trials.
II-1 One or more well-conducted cohort studies.
II-2 One or more well-conducted case-control studies.
II-3 A dramatic uncontrolled experiment.
III Expert committee sitting in review; peer leader opinion.
IV Personal experience.

How can we get solid results from literature: The use of systematic review design and methods

Professor Carole Torgerson
Durham University, UK
carole.torgerson@durham.ac.uk

Professor David Torgerson
University of York, UK
david.torgerson@york.ac.uk

Milan March 15th 2016

"How to get solid results from literature: the systematic reviews method"

http://centridiricerca.unicatt.it/crc-impact-laboratorio-sulla-valutazione-delle-seminiario-can-we-learn-something-from

Tipi di rassegna

Scoping review
- Broad mapping of the literature in a field

Systematic review
- In depth review relating to specific review question (may or may not include meta-analysis; may be used to address substantive or methodological question, for synthesis or hypothesis generation)

Tertiary review
- Overview of systematic reviews in a field

Cosa è una systematic review?

A systematic review
- is a synthesis of all the studies relevant to address a specific research question
- has explicit, transparent, replicable methods
- states in advance the criteria for including studies
- searches exhaustively for all the relevant studies within a pre-defined area
- negative and positive studies included to give an overall impartial view of the field
- narrative synthesis or meta-analysis

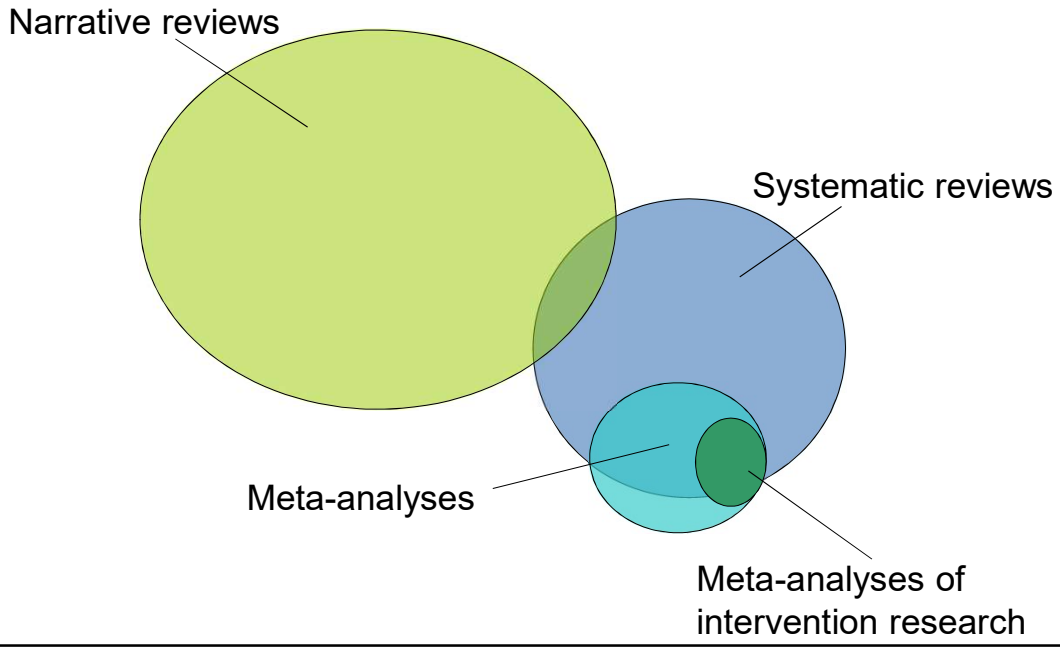
Key features of a systematic review

Explicit, transparent, replicable search

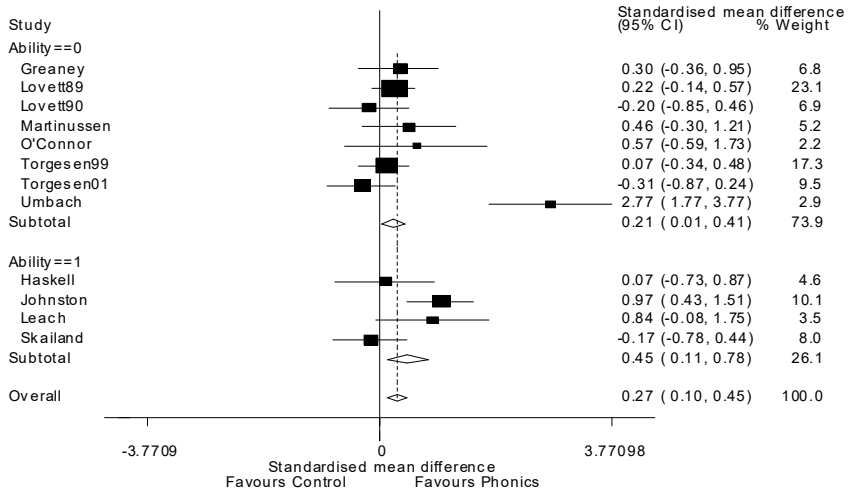
Critical evaluation (quality appraisal) of all included studies

Synthesis: narrative or quantified (meta-analysis)

Literature reviews - conceptual relations



Cosa è una meta-analisi



I 9 passi di una *systematic review*

1. Rassegna esplorativa, definizione dell'interrogativo e stesura del protocollo
2. Ricerca della letteratura
3. Selezione di titoli e abastract
4. Recupero dei paper
5. Selezione dei paper davvero pertinenti
6. Valutazione della qualità degli studi
7. Estrazione dei risultati
8. Sintesi e analisi (eventuale meta-analisi)
9. Scrittura della rassegna

Una domanda di ricerca viene tradotta in

(PICOS)

TERMINI collegati da operatori booleani

MOTORI di ricerca su cui impiegarli

CRITERI di selezione