



Spring Meeting Giovani Ricercatori



SID



SIIA



SIMI



SIPREC



SISA

Perché capire il Propensity Score?

Approccio statistico dagli studi randomizzati agli studi real world

Mario Luca Morieri

UOC Malattie del Metabolismo

Azienda Ospedaliera Universitaria di Padova (Padova)

Università degli Studi di Padova

Joslin Diabetes Center (Boston, USA)

WORKSHOP: "Come approcciarsi ai test statistici: road to Propensity Score"

Vantaggi e punti di forza dei trial clinici randomizzati

Piramide della medicina basata sulle evidenze



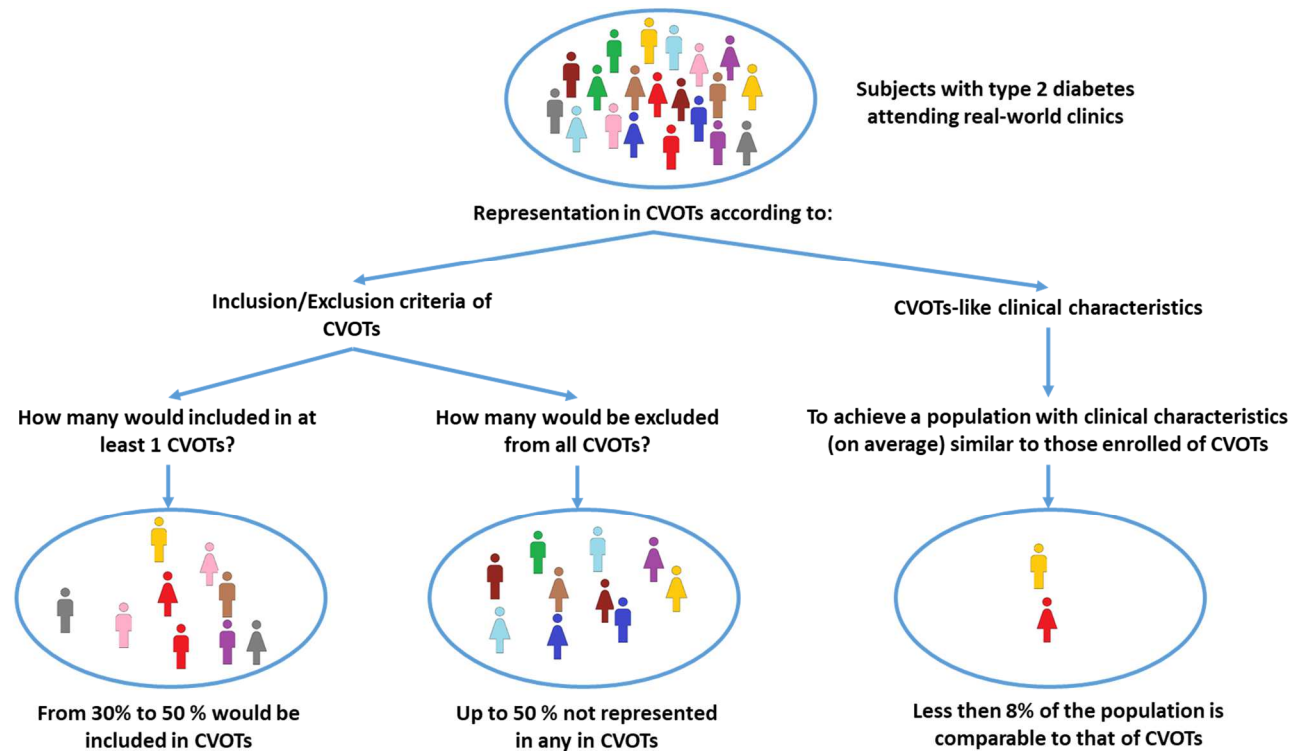
- Studi clinici randomizzati e controllati (RCT)
- Condotti e disegnati ad-hoc
- Metodologia rigorosa
- Contesti clinici controllati
 - Risultati Affidabili
 - Elevata validità interna
 - Grazie alla randomizzazione: liberi da possibili effetti confondenti (misurati, misurabili e non misurati o non misurabili).
- **«CAVILLO» → Rigoroso contesto sperimentale ≠ pratica clinica routinaria (Real-World)**

Opportunità dagli studi real-world

- Quanto sono rappresentativi gli RCTs vs popolazione vista in pratica clinica?

Opportunità dagli studi real-world

- Quanto sono rappresentativi gli RCTs vs popolazione vista in pratica clinica?



Opportunità dagli studi real-world

- Quanto sono rappresentativi gli RCTs vs popolazione vista in pratica clinica?
- Le caratteristiche dei soggetti trattati sono simili a quelle dei soggetti arruolati negli RCT?
- Se la popolazione è diversa, i risultati degli RCT sono ancora validi? (generalizzabilità)
- Se la popolazione trattata ha caratteristiche simili, ma il contesto è diverso (es frequenza visite di follow-up, contatti medico/infermiere, aderenza...) i risultati degli RCT sono trasferibili al contesto del mondo reale?
- Quali sono i fattori che potrebbero influenzare l'efficacia del trattamento nel RW (abitudini prescrittive, inerzia terapeutica)?

Opportunità e Potenzialità dagli studi real-world


The Economist

Menu Weekly edition The world in brief Search

Leaders | Regulating the internet giants

The world's most valuable resource is no longer oil, but data

The data economy demands a new approach to antitrust rules



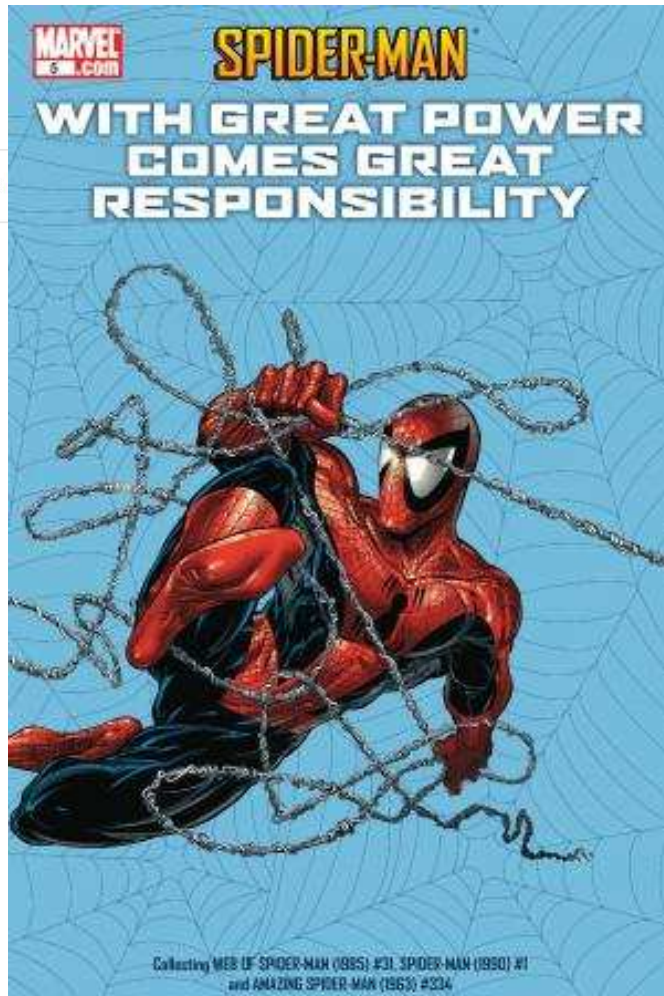
May 6th 2017

Share

- Raccolti quotidianamente e in maniera incrementale
- Popolazioni più ampie e "teoricamente" complete
- Costo molto inferiore rispetto RCTs
- Riutilizzabili e "illimitate"
- Crescita dei link con database amministrativi (info ambientali: socio-economiche, educazionali, abitative) e database di ricerca (es UK biobank, Coorti paesi scandinavi...).

Opportunità e Potenzialità dagli studi real-world

The
Economist



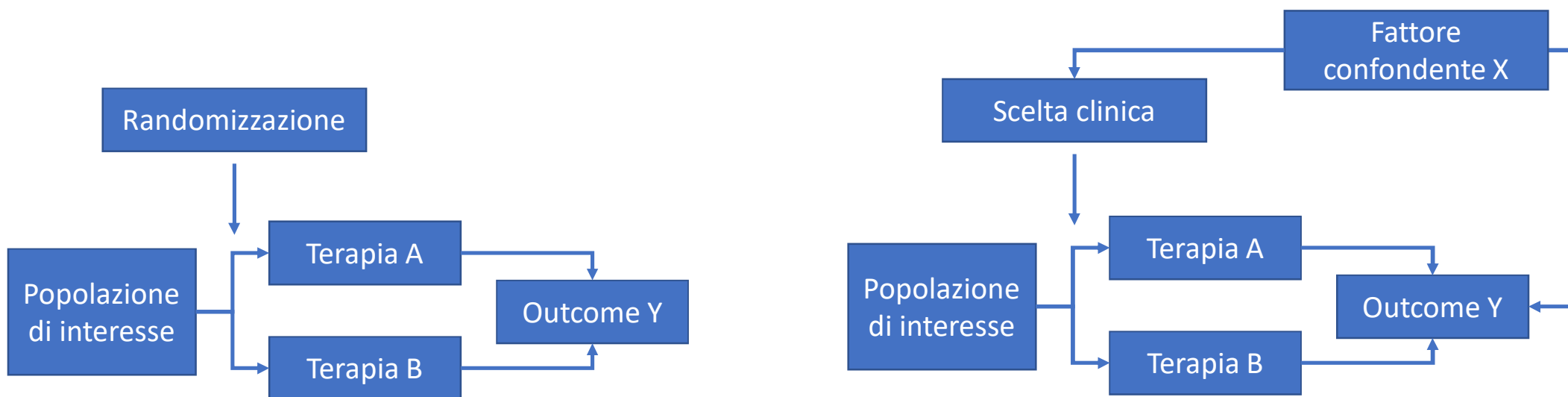
Real World Data
≠
Real World Study
≠
Real World Evidence

database di ricerca (es UK biobank, Coorti paesi scandinavi...).

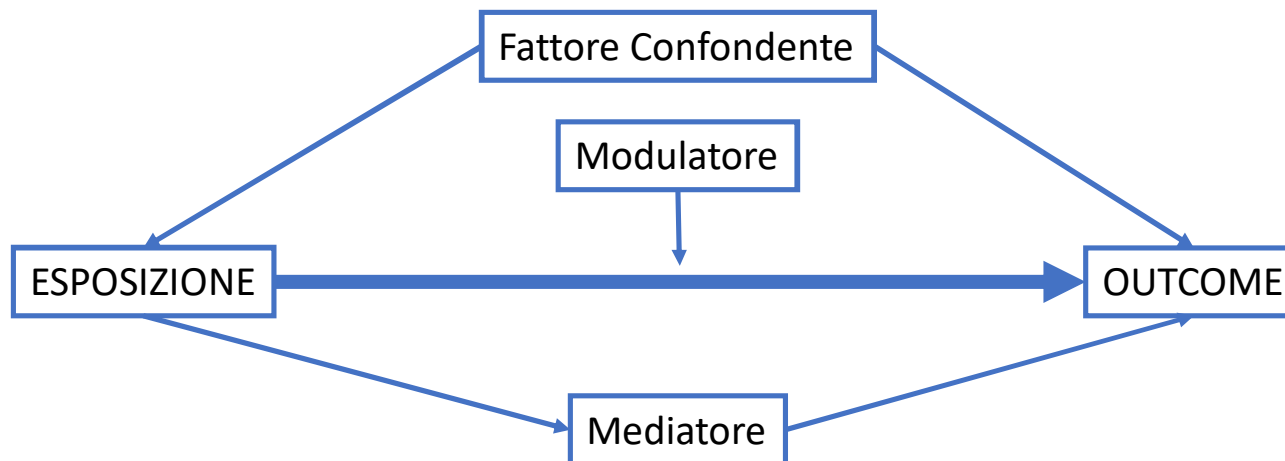
RCT e RWE a confronto

Quesito: quale trattamento è migliore tra A e B per ridurre l'incidenza dell'outcome Y ?

Ipotesi nulla: pazienti trattati con A e B hanno Y uguale



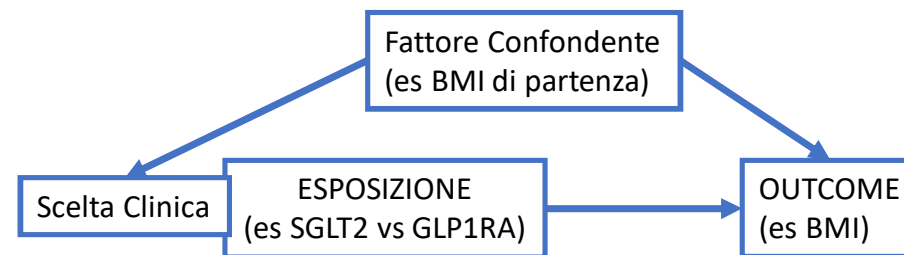
RCT e RWE a confronto



RCT



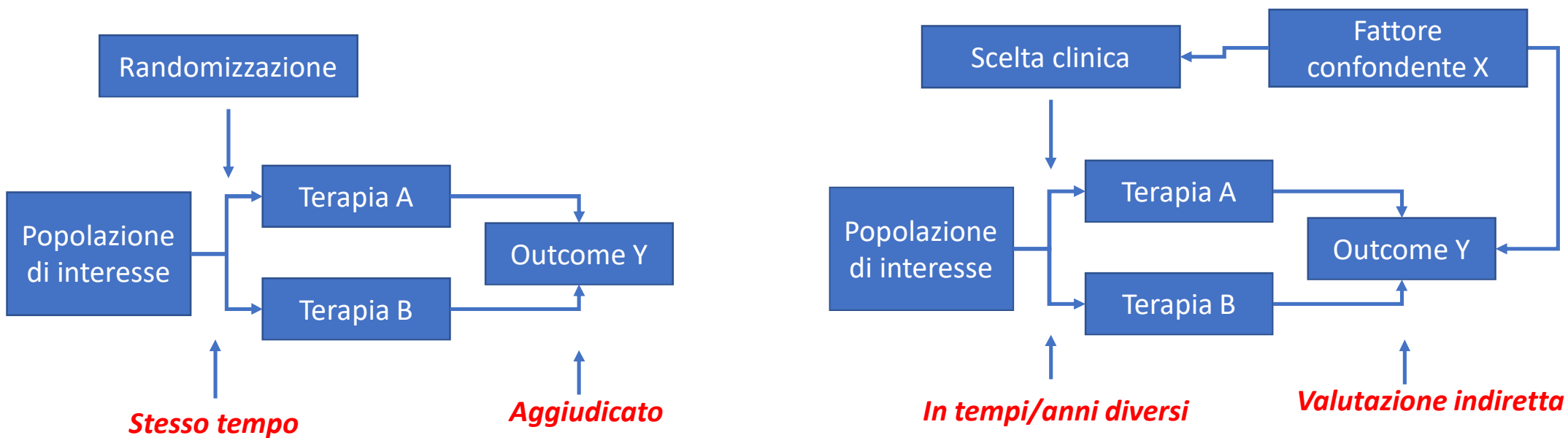
Real World Data



RCT e RWE a confronto

Quesito: quale trattamento è migliore tra A e B per ridurre l'incidenza dell'outcome Y ?

Ipotesi nulla: pazienti trattati con A e B hanno Y uguale



Biases → errori sistematici nella stima dell'associazione/effetto

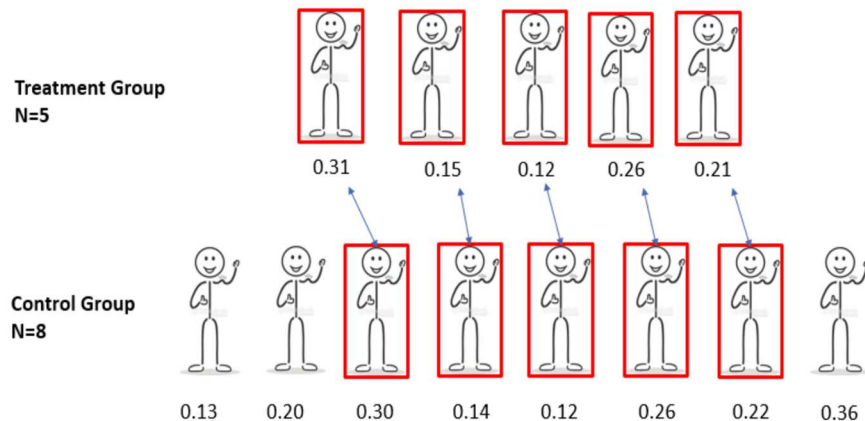
Domains	Description	Details/example
Confounding Biases	Factors affecting treatment choice and outcomes	Residual confounding (known and measured confounders) or unmeasured confounding.
Selection Biases	Exclusion of subjects (or events/outcomes or follow-up time)	Bias related to both intervention and outcome (immortal-time-bias, exclusion of subjects with missing data)
Information Biases (measurement bias)	Misclassification of intervention status or outcomes.	Misclassification of intervention in retrospective cohort studies when information on interventions are influenced by outcomes (a.k.a. <i>recall bias</i>). Misclassification of outcomes (aka <i>detection bias</i>) → intensity of observations/meas. of outcomes ≠ between intervention groups.
Reporting Biases	Selection of the reported results arising from a desire for findings to merit publication.	Selection of results based on P-values, magnitude or direction of the estimated effect of intervention. Selection of outcomes, analyses, or selective reporting of a subgroup of participants.

Confounding Bias – Propensity score

A **propensity score** is the probability (or propensity) for having the factor of interest (e.g., receiving a particular treatment) given the factors present at baseline.

→ **Multivariable logistic regression model to obtain predicted probabilities (i.e., propensity score values - PS) for both the probability of being treated, p , and the probability of not being treated $1 - p$.**

PS matching



Inverse probability weighting (IPW) (*probilità inversa ponderata*)

Each observation is weighted by the reciprocal (i.e., the inverse) of the predicted probability of being in the group that was observed for each patient

The weight for each treated unit is $1 / PS$
The weight for each control unit is $1 / (1 - PS)$

Propensity score, selezione delle variabili e bilanciamento

L'obiettivo di PS è bilanciare i gruppi trattati e non trattati sui fattori confondenti che influenzano sia l'assegnazione del trattamento che l'outcome

→ **Scelta delle variabili** (diversi punti di vista)

- Includere tutti i fattori disponibili, anche se il modello è «overfit», è accettabile quando si costruisce un PS, poiché in questo caso la performance finale (predizione) prevale sulla «parsimonia».
- Attenzione: overfitting del PS può portare a una varianza più ampia nei punteggi stimati, in particolare una sovrabbondanza di punteggi vicino a 0 o 1 → quindi maggiore varianza del ETE (estimated treatment effect) → includere fattori che influiscono sul trattamento ma non sull'outcome può portare a punteggi più estremi.

Propensity score, selezione delle variabili e bilanciamento

L'obiettivo di PS è bilanciare i gruppi trattati e non trattati sui fattori confondenti che influenzano sia l'assegnazione del trattamento che l'outcome

→ **Dimostrare che il PS raggiunge un buon bilanciamento è essenziale**

- più importante della discriminazione del modello del PS (es c-statistic)
- Valutazione delle differenze medie standardizzate (SMD, differenza nelle medie trattate e non trattate per ciascun fattore divisa per la deviazione standard aggregata). **Se SMD <0,1 OK!** (non guardare solo a p-value)
- Se squilibri importanti, il PS può essere modificato (es fattori aggiuntivi, interazioni, effetti non lineari ...)

→ **Lo sviluppo di un punteggio di propensione è un processo iterativo.**

→ **Va ottimizzato prima di eseguire le analisi sull'outcome e sempre documentato (meglio se deciso a priori)**

