

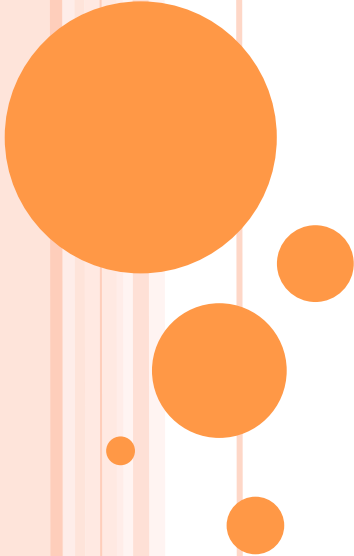


UNIVERSITÀ
DEGLI STUDI
DI MILANO



LA STATALE

PEDIATRIC CLINICAL
RESEARCH CENTER



METODOLOGIA DELLA RICERCA: UN'INTRODUZIONE (M1-METH-TEORIA)

Giorgio Bedogni

OBIETTIVO

- Fornire le basi metodologiche per il disegno e l'interpretazione degli studi clinici

PROGRAMMA

- Sessione teorica
 - Martedì 14 marzo 2017
- Sessione pratica
 - Martedì 21 marzo 2017

PROGRAMMA

- *Parte teorica*
 - *Martedì 14 marzo 2017*
- *Parte pratica*
 - *Martedì 21 marzo 2017*

PROGRAMMA

- Una visione d'insieme
- La domanda dello studio
- I soggetti dello studio
- Le variabili e i metodi dello studio
- L'ipotesi nulla e la stima della numerosità campionaria
- I disegni di studio
- Aspetti etici

PROGRAMMA

- *Una visione d'insieme*
- La domanda dello studio
- I soggetti dello studio
- Le variabili e i metodi dello studio
- L'ipotesi nulla e la stima della numerosità campionaria
- I disegni di studio
- Aspetti etici

STRUTTURA DELLO STUDIO

- Domanda dello studio
 - Qual è la domanda a cui lo studio vuole rispondere?
- Rilevanza
 - Perché questa domanda è importante?
- Disegno dello studio
 - Come è strutturato lo studio?

STRUTTURA DELLO STUDIO

○ Soggetti

- Chi sono i soggetti dello studio e come saranno selezionati?

○ Variabili

- Quali misurazioni saranno effettuate?
- Quale variabile è l'*outcome*?
- Quali variabili sono i *predittori*?

STRUTTURA DELLO STUDIO

- Analisi statistica
 - Qual è l'*ipotesi* dello studio?
 - Qual è l'approccio analitico migliore per rispondere alla domanda?
 - Qual è la numerosità campionaria necessaria per rispondere alla domanda?

LO STUDIO IN BREVE: ESEMPIO

- Domanda dello studio
 - La supplementazione di calcio a donne in post-menopausa riduce la frequenza di fratture?
- Rilevanza
 - I dati disponibili in letteratura sono contrastanti e ottenuti da disegni di studio inadeguati [*non* è vero, ma lo assumiamo per ragioni didattiche]
- Disegno di studio
 - Studio randomizzato controllato

LO STUDIO IN BREVE: ESEMPIO

○ Soggetti

- Donne in post-menopausa: un gruppo supplementato per 5 anni con calcio e uno supplementato con placebo

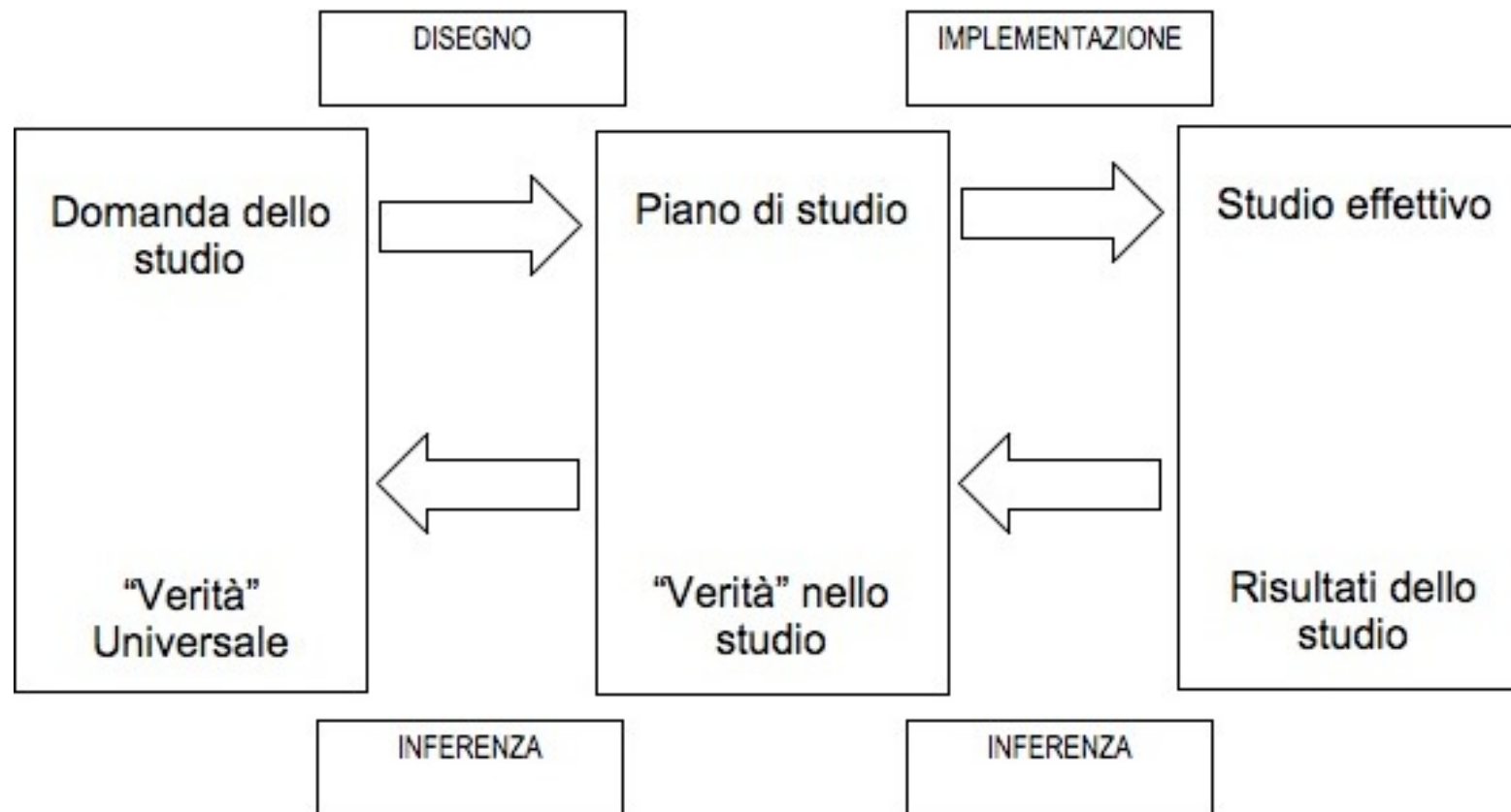
○ Variabili

- Outcome: numero di fratture nei 5 anni successivi alla somministrazione del trattamento
- Predittore: la supplementazione di calcio

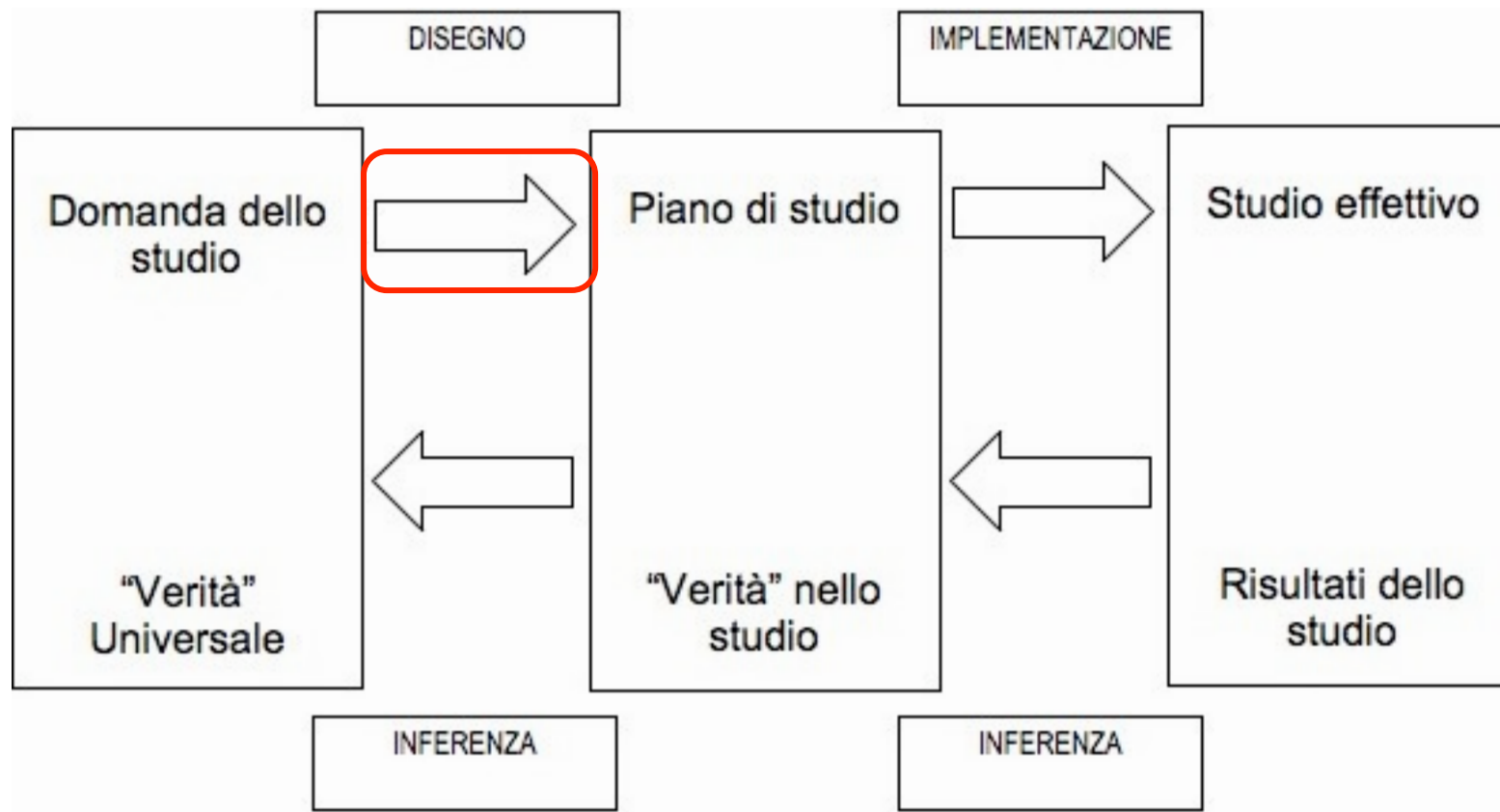
LO STUDIO IN BREVE: ESEMPIO

- Analisi statistica
 - Compileremo questa parte dopo aver imparato altri concetti ...

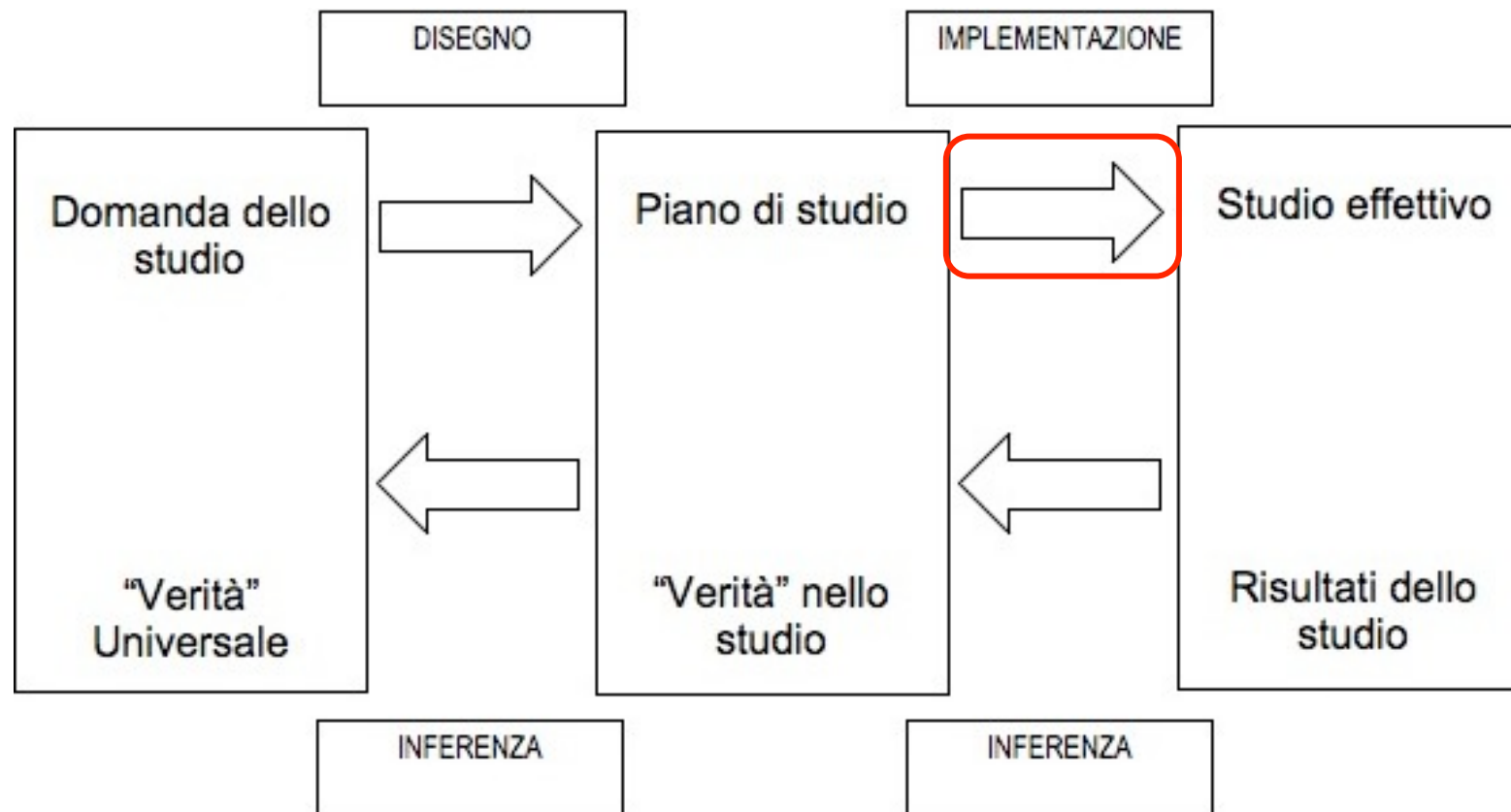
SVOLGIMENTO DI UNO STUDIO



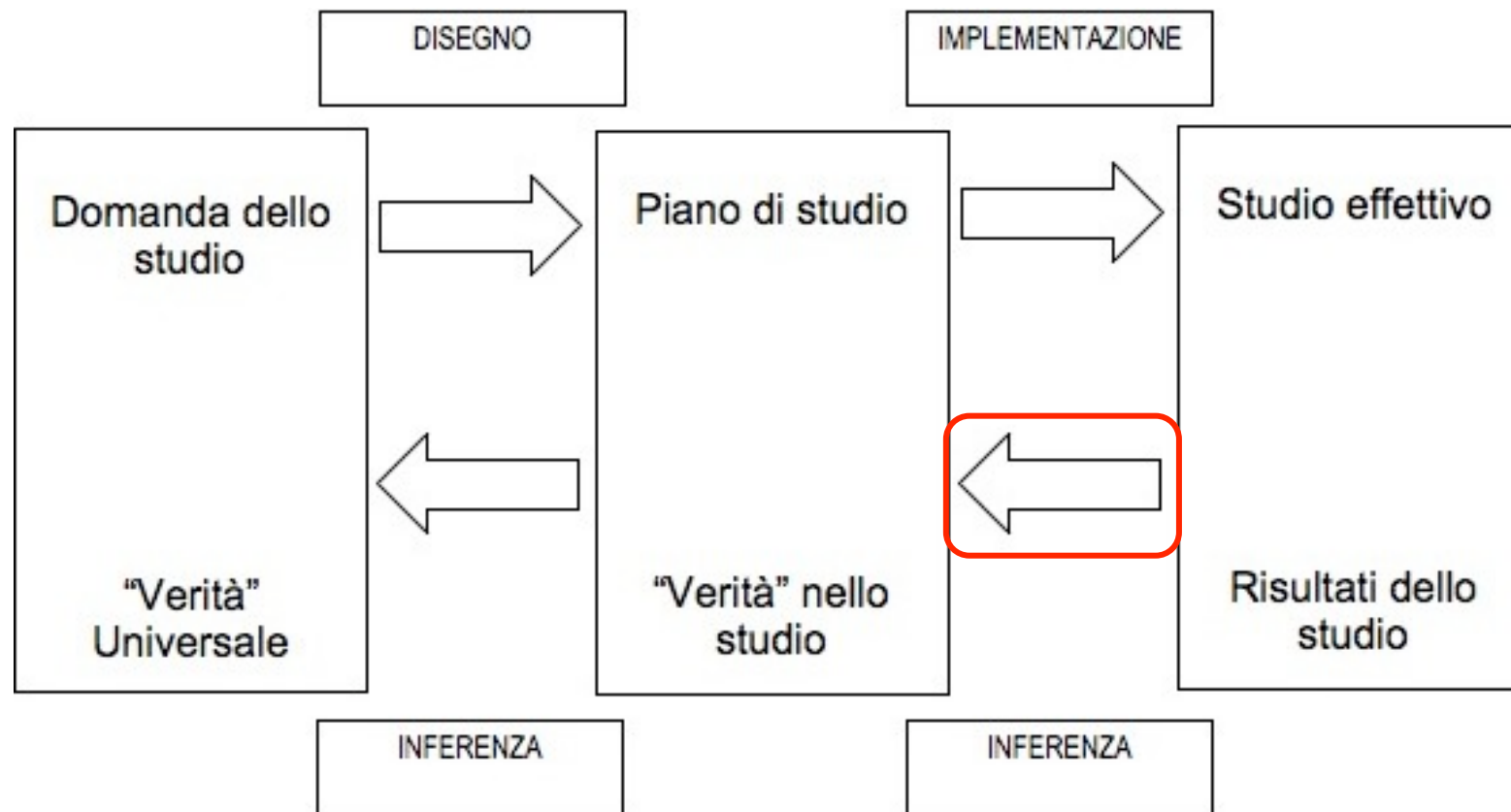
DISEGNO



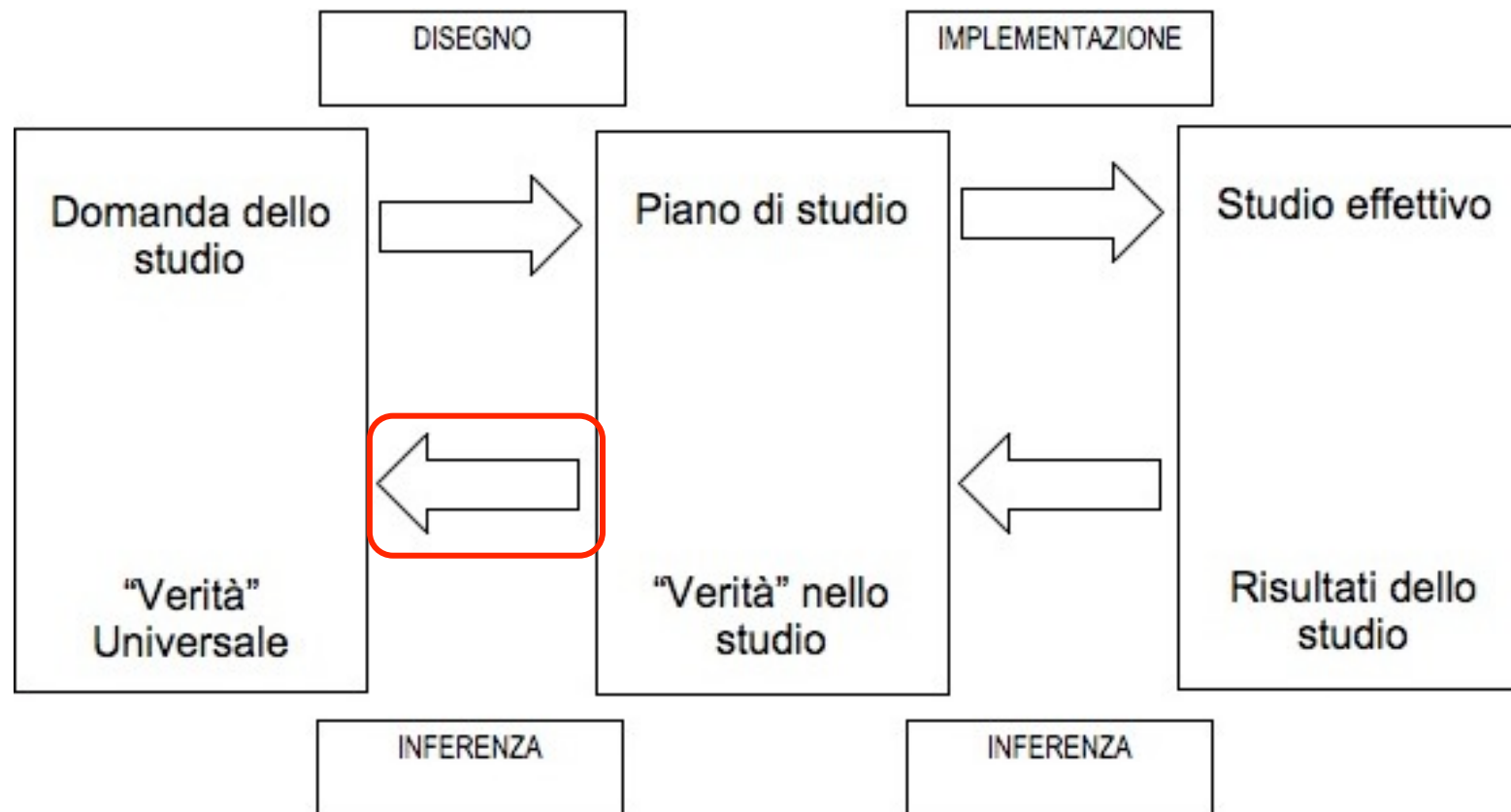
IMPLEMENTAZIONE



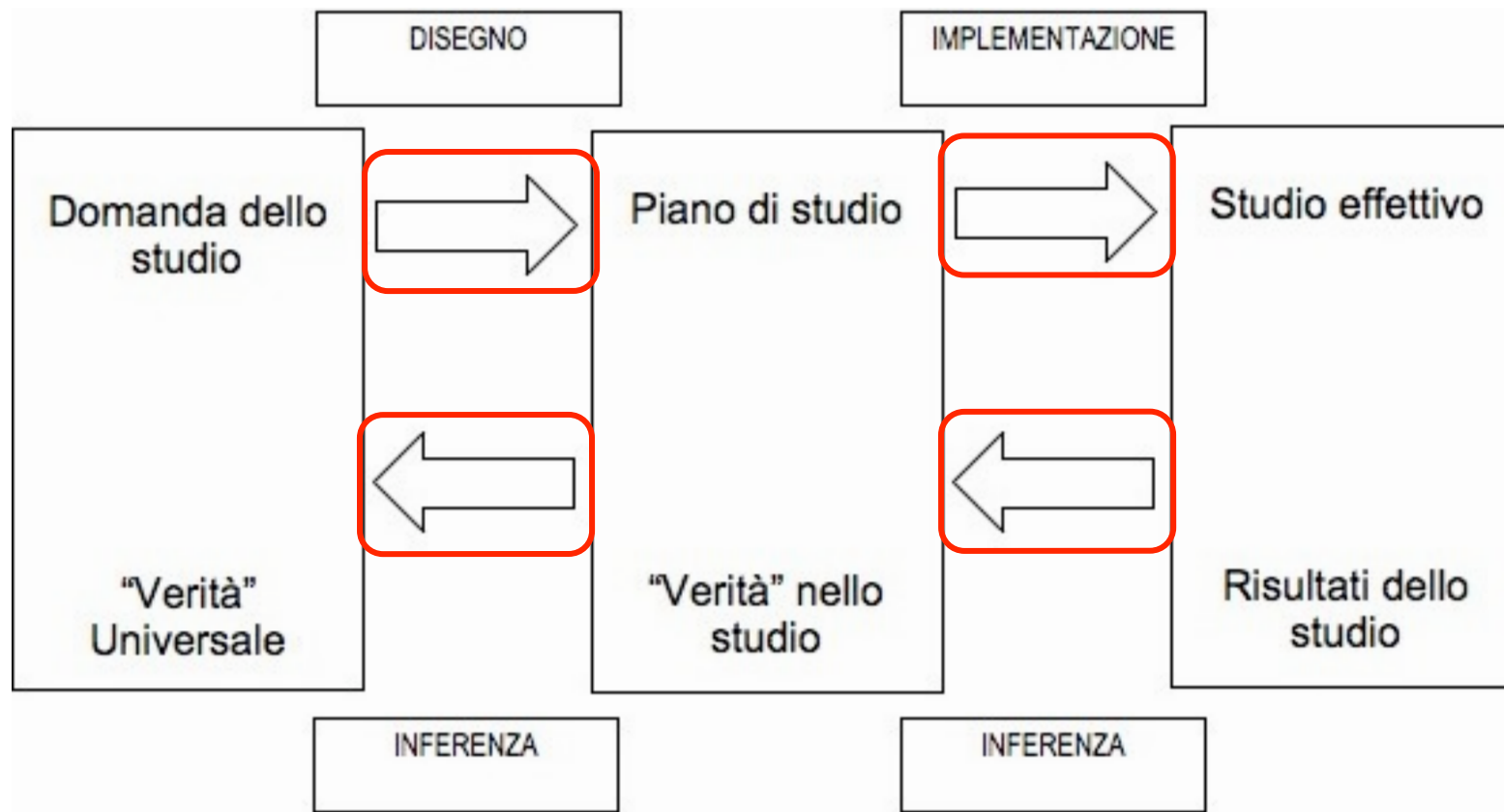
INFERENZA INTERNA



INFERENZA ESTERNA



POSSIBILITÀ DI ERRORE



QUALI ERRORI?

- Errore casuale (“random error”)
 - È dovuto al caso
 - Influenza la stima dell’outcome in entrambe le direzioni (“sopra e sotto”)
- Errore sistematico (“systematic error”)
 - È dovuto al *bias*
 - Influenza la stima dell’outcome in una sola direzione (“sopra o sotto”)

IL PROTOCOLLO DI STUDIO

- Il protocollo di studio è l'organizzazione coerente e logica del progetto di ricerca
- La versione iniziale del protocollo è *breve* e serve per fissare e valutare l'idea dello studio nel suo insieme
- Il protocollo evolve continuamente (in dimensione e contenuto) fino ad assumere la forma finale

PROGRAMMA

- Una visione d'insieme
- *La domanda dello studio*
- I soggetti dello studio
- Le variabili e i metodi dello studio
- L'ipotesi nulla e la stima della numerosità campionaria
- I disegni di studio
- Aspetti etici

LA DOMANDA DELLO STUDIO

- L'incertezza che il ricercatore desidera “risolvere” attraverso lo studio

DOVE ORIGINA LA DOMANDA DELLO STUDIO?

- Ricerca precedente (propria ed altrui)
- Frequentazione di congressi
- Osservazione dei pazienti
- Insegnamento

CONSULTAZIONE DELLA LETTERATURA

- Adottare un approccio sistematico
- (Effettuare una “racsegna sistematica” e, quando possibile, una “metanalisi”)

SCOPERTA DI WESTHEIMER

- Un paio di mesi in laboratorio spesso fanno risparmiare un paio d'ore in biblioteca

Bloch A. *La legge di Murphy del 2000*.

Longanesi; 1999, p. 57.

UNA BUONA DOMANDA DI STUDIO

- È “FINER” !
 - “Feasible” = “fattibile”
 - “Interesting” = “interessante”
 - “Novel” = “nuova”
 - “Ethical” = “etica”
 - “Relevant” = “rilevante”

FATTIBILITÀ

- “**FINER**” = Fattibilità
 - Numero di soggetti adeguato?
 - Esperienza tecnica adeguata?
 - Sufficienti risorse (tempo e denaro)?
 - Possibilità di portare a termine il progetto?

INTERESSE

- “F**I**NER” = Interesse
 - Il *ricercatore* è interessato al progetto?

NOVITÀ

- “FINER” = Novità
 - Aumenta la conoscenza disponibile?
 - Conferma o confuta evidenze precedenti?
 - Offre nuove evidenze?

ETICA

- “FIN**E**R” = Etica
 - È “etica”?

RILEVANZA

- “FINE**R**” = Rilevanza
 - Per la conoscenza scientifica?
 - Per la medicina e la salute pubblica?
 - Per la ricerca futura?

PERCHÉ VIENE RIFIUTATO UN ARTICOLO SCIENTIFICO?

- L'ipotesi non è scientificamente rilevante
- Lo studio non è originale
- Lo studio non ha testato l'ipotesi degli Autori
- Il disegno dello studio è errato
- Il protocollo di studio è stato compromesso da problemi in corso di studio (es. perdita cospicua al follow-up)

Greenhalgh T. *How to read a paper.*

BMJ Books: London 2001.

PERCHÉ VIENE RIFIUTATO UN ARTICOLO SCIENTIFICO?

- Il numero di soggetti è inadeguato
- Lo studio non è “controllato”
- L’analisi statistica è inadeguata
- Le conclusioni non sono giustificate dai dati
- C’è un evidente conflitto di interesse degli Autori
- Il linguaggio è incomprensibile

Greenhalgh T. *How to read a paper*.
BMJ Books: London 2001.

PROGRAMMA

- Una visione d'insieme
- La domanda dello studio
- *I soggetti dello studio*
- Le variabili e i metodi dello studio
- L'ipotesi nulla e la stima della numerosità campionaria
- I disegni di studio
- Aspetti etici

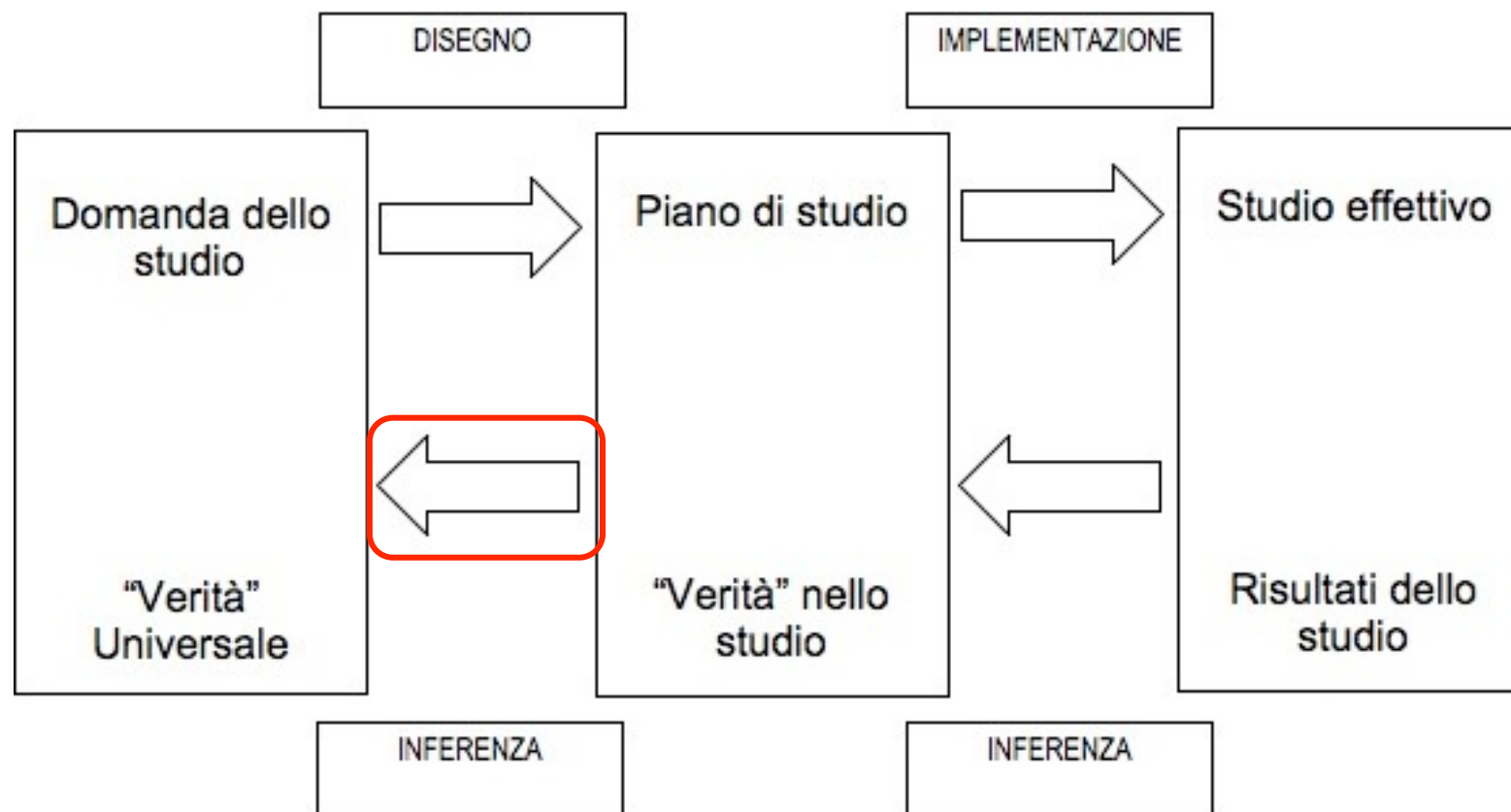
CAMPIONE E POPOLAZIONE

- La ricerca epidemiologica si basa sull'uso di un *campione* per rappresentare una *popolazione*
- Il *campionamento* consente di fare *inferenze* sulla popolazione studiando un numero inferiore di soggetti

CAMPIONE E POPOLAZIONE

- Il vantaggio del campionamento è (dunque) la sua *efficienza*
- Se il campione non è sufficientemente *rappresentativo*, la generalizzazione non è possibile

VERITÀ UNIVERSALE?



CHI HA BISOGNO DELLA RAPPRESENTATIVITÀ?

Published by Oxford University Press on behalf of the International Epidemiological Association
© The Author 2013; all rights reserved.

International Journal of Epidemiology 2013;**42**:1012–1014
doi:10.1093/ije/dys223

POINT COUNTERPOINT

Why representativeness should be avoided

Kenneth J Rothman,^{1,2} John EJ Gallacher³ and Elizabeth E Hatch¹

¹Department of Epidemiology, Boston University School of Public Health, Boston, MA, USA, ²RTI Health Solutions, RTI International, Research Triangle Park, NC, USA and ³Institute of Primary Care and Public Health, Cardiff University, Cardiff, UK

Accepted 21 November 2012

CHI HA BISOGNO DELLA RAPPRESENTATIVITÀ?

- “What exactly the principle of representativeness embodies?”
- “The mistake is to think that statistical inference is the same as scientific inference”
- “Surveys of opinions, of the prevalence of disease, of habits or of environmental exposures may be informative but are not science in the same the way that causal studies about how nature operates are science”

CHI HA BISOGNO DELLA RAPPRESENTATIVITÀ?



J. R. Statist. Soc. A (2016)
179, Part 2, pp. 319–376

Perils and potentials of self-selected entry to epidemiological studies and surveys

Niels Keiding

University of Copenhagen, Denmark

and Thomas A. Louis

Johns Hopkins Bloomberg School of Public Health, Baltimore USA

[Read before The Royal Statistical Society at the 2015 Joint Statistical Meetings of the American Statistical Association in Seattle on Wednesday, August 12th, 2015, the President, Professor P.J. Diggle, in the Chair]

DEFINIZIONE DELLA POPOLAZIONE

- Criteri di inclusione
 - Caratteristiche dei soggetti ritenuti idonei allo studio
- Criteri di esclusione
 - Caratteristiche dei soggetti ritenuti inidonei allo studio (o che non possono essere studiati per motivi etici)

CRITERI DI INCLUSIONE

- Caratteristiche demografiche
 - Es. donne “caucasiche” (“europidi”) in menopausa ...
- Caratteristiche cliniche
 - Es. senza malattie croniche o neoplasie all’ingresso nello studio ...
- Caratteristiche geografiche
 - Es. dalla popolazione generale della città di Milano
- ...

CRITERI DI ESCLUSIONE

- Alta probabilità di perdita al follow-up
- Incapacità di fornire dati adeguati
- Ad alto rischio di effetti collaterali
- Ragioni etiche

CRITERI DI ESCLUSIONE

- Essere parsimoniosi per evitare di compromettere la generalizzabilità dello studio (considerando l'obiezione di Rothman)

STRATEGIE DI CAMPIONAMENTO

- Campione di convenienza
- Campione di probabilità

CAMPIONE DI CONVENIENZA

- Il più utilizzato nella ricerca clinica (semplice, di basso costo e adatto a molti contesti)
- Necessita di arruolamento consecutivo per evitare bias di selezione
- La rappresentatività del campione così arruolato è *sempre* una questione di giudizio

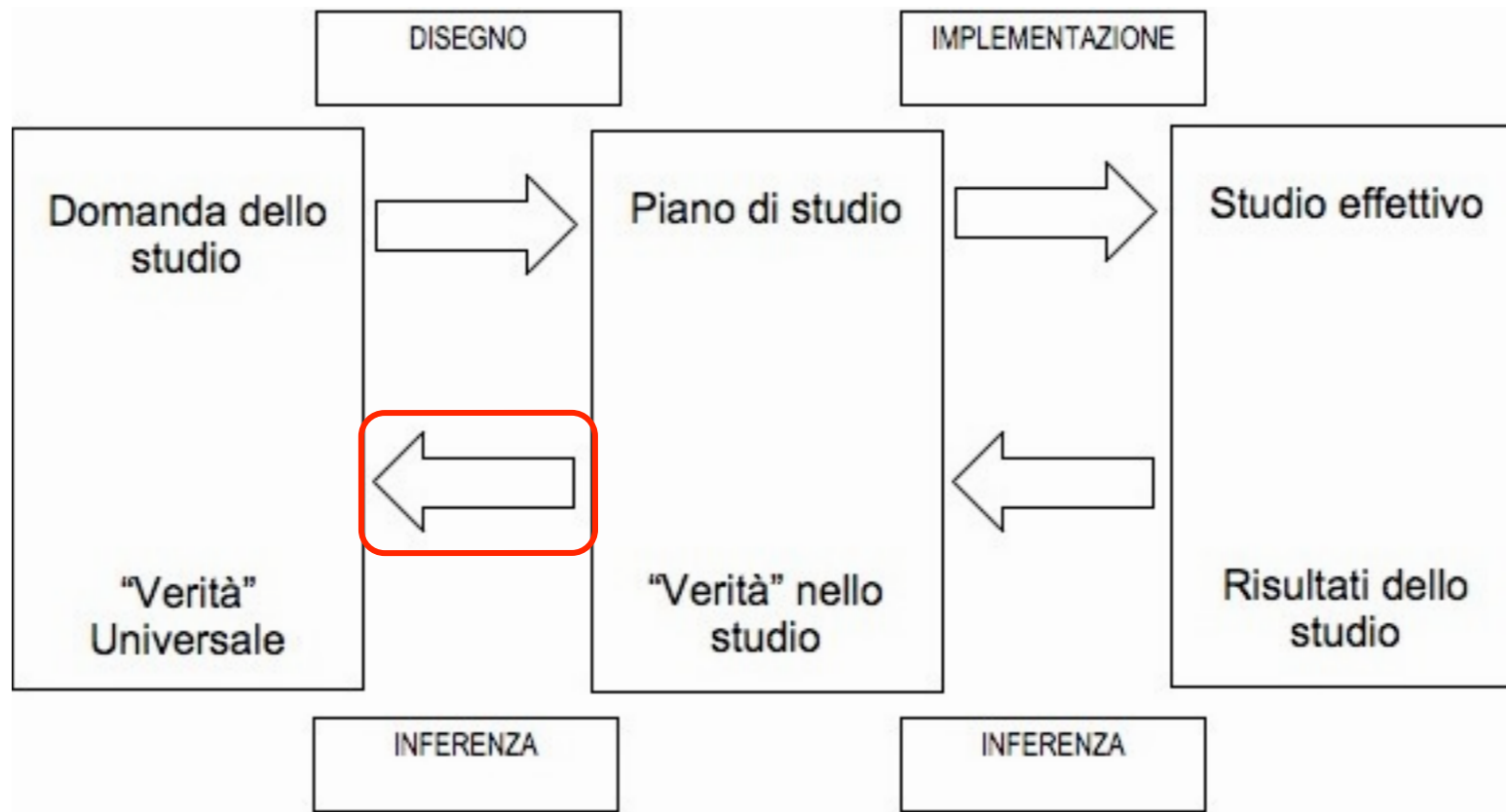
CAMPIONE DI PROBABILITÀ

- Uso di metodi statistici per garantire che ogni soggetto della popolazione abbia una *probabilità definita* di essere selezionato per lo studio

CAMPIONE DI PROBABILITÀ

- Semplice
 - Enumerazione dei soggetti e scelta casuale degli stessi
- Stratificato
 - Stratificazione dei soggetti e scelta casuale degli stessi
- A grappoli (cluster)
 - Scelta casuale all'interno di gruppi
- Sistemático (DA EVITARE)
 - Enumerazione dei soggetti e scelta con criterio preordinato

DAL CAMPIONE ALLA POPOLAZIONE



UN CAVEAT SULLA RAPPRESENTATIVITÀ

The New England Journal of Medicine

© Copyright, 1999, by the Massachusetts Medical Society

VOLUME 341

SEPTEMBER 2, 1999

NUMBER 10

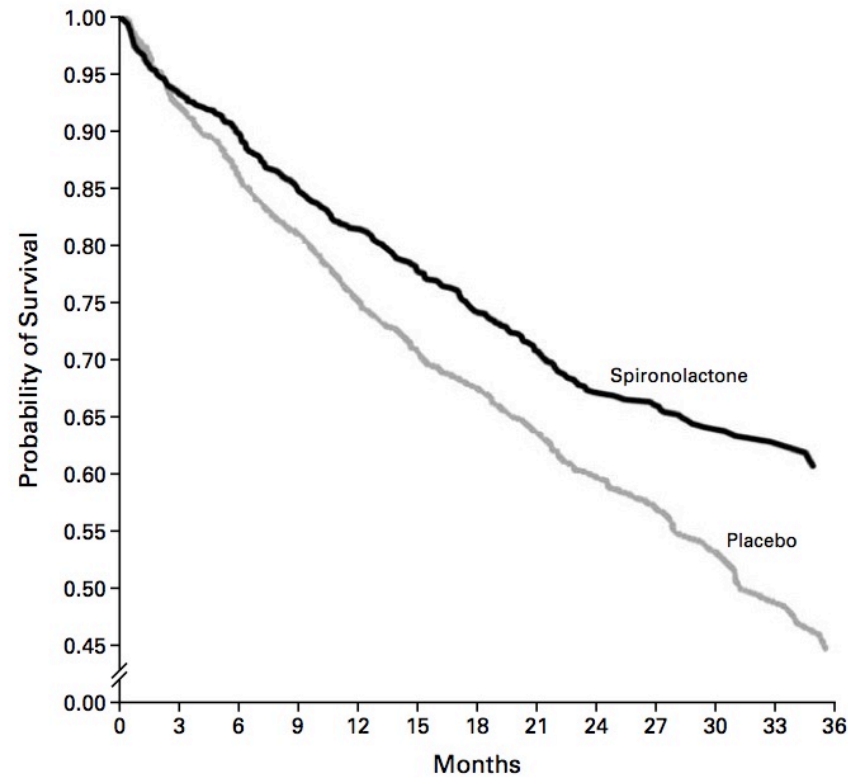


THE EFFECT OF SPIRONOLACTONE ON MORBIDITY AND MORTALITY IN PATIENTS WITH SEVERE HEART FAILURE

BERTRAM PITT, M.D., FAIEZ ZANNAD, M.D., WILLEM J. REMME, M.D., ROBERT CODY, M.D., ALAIN CASTAIGNE, M.D.,
ALFONSO PEREZ, M.D., JOLIE PALENSKY, M.S., AND JANET WITTES, PH.D.,
FOR THE RANDOMIZED ALDACTONE EVALUATION STUDY INVESTIGATORS*

www.giorgiobedogni.it

UN CAVEAT SULLA RAPPRESENTATIVITÀ



No. AT RISK													
Placebo	841	775	723	678	628	592	565	483	379	280	179	92	36
Spironolactone	822	766	739	698	669	639	608	526	419	316	193	122	43

Figure 1. Kaplan–Meier Analysis of the Probability of Survival among Patients in the Placebo Group and Patients in the Spironolactone Group.

The risk of death was 30 percent lower among patients in the spironolactone group than among patients in the placebo group ($P < 0.001$).

UN CAVEAT SULLA RAPPRESENTATIVITÀ

- “The median creatinine and potassium concentrations did not change in the placebo group during the first year of follow-up”
- “During the same period, however, the median creatinine concentration in the spironolactone group increased by approximately 0.05 to 0.10 mg/dl and the median potassium concentration increased by 0.30 mmol/l”
- “The differences between the two groups were significant ($p < 0.001$) *but were not clinically important*”

UN CAVEAT SULLA RAPPRESENTATIVITÀ

The NEW ENGLAND JOURNAL *of* MEDICINE

ORIGINAL ARTICLE

Rates of Hyperkalemia after Publication of the Randomized Aldactone Evaluation Study

David N. Juurlink, M.D., Ph.D., Muhammad M. Mamdani, Pharm.D., M.P.H.,
Douglas S. Lee, M.D., Alexander Kopp, B.A., Peter C. Austin, Ph.D.,
Andreas Laupacis, M.D., and Donald A. Redelmeier, M.D.

N Engl J Med 2004;351:543-51.

Copyright © 2004 Massachusetts Medical Society.

UN CAVEAT SULLA RAPPRESENTATIVITÀ

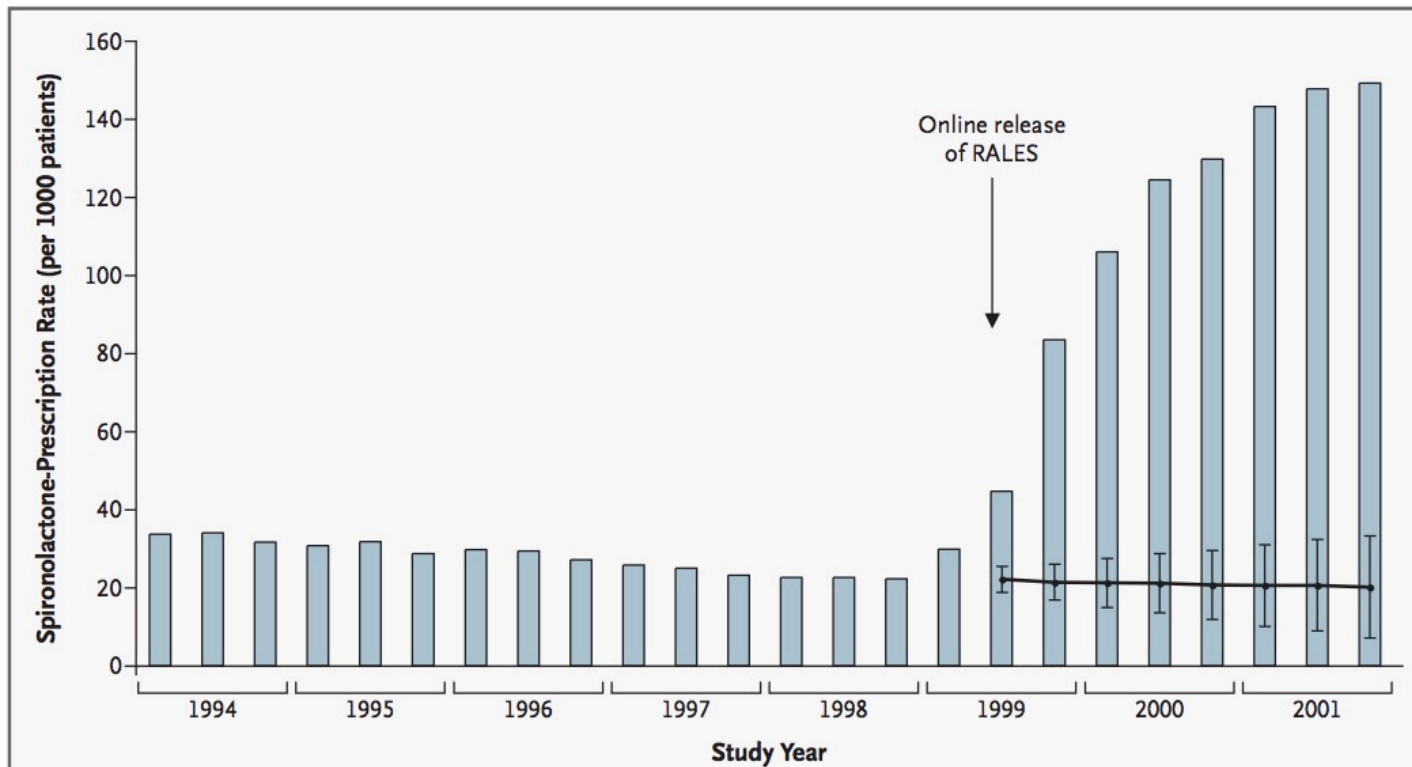


Figure 1. Rate of Prescriptions for Spironolactone among Patients Recently Hospitalized for Heart Failure Who Were Receiving ACE Inhibitors.

Each bar shows the observed spironolactone-prescription rate per 1000 patients during one four-month interval. The line beginning in the second interval of 1999 shows projected prescription rates derived from interventional autoregressive integrated moving-average (ARIMA) models, with I bars representing the 95 percent confidence intervals.

UN CAVEAT SULLA RAPPRESENTATIVITÀ

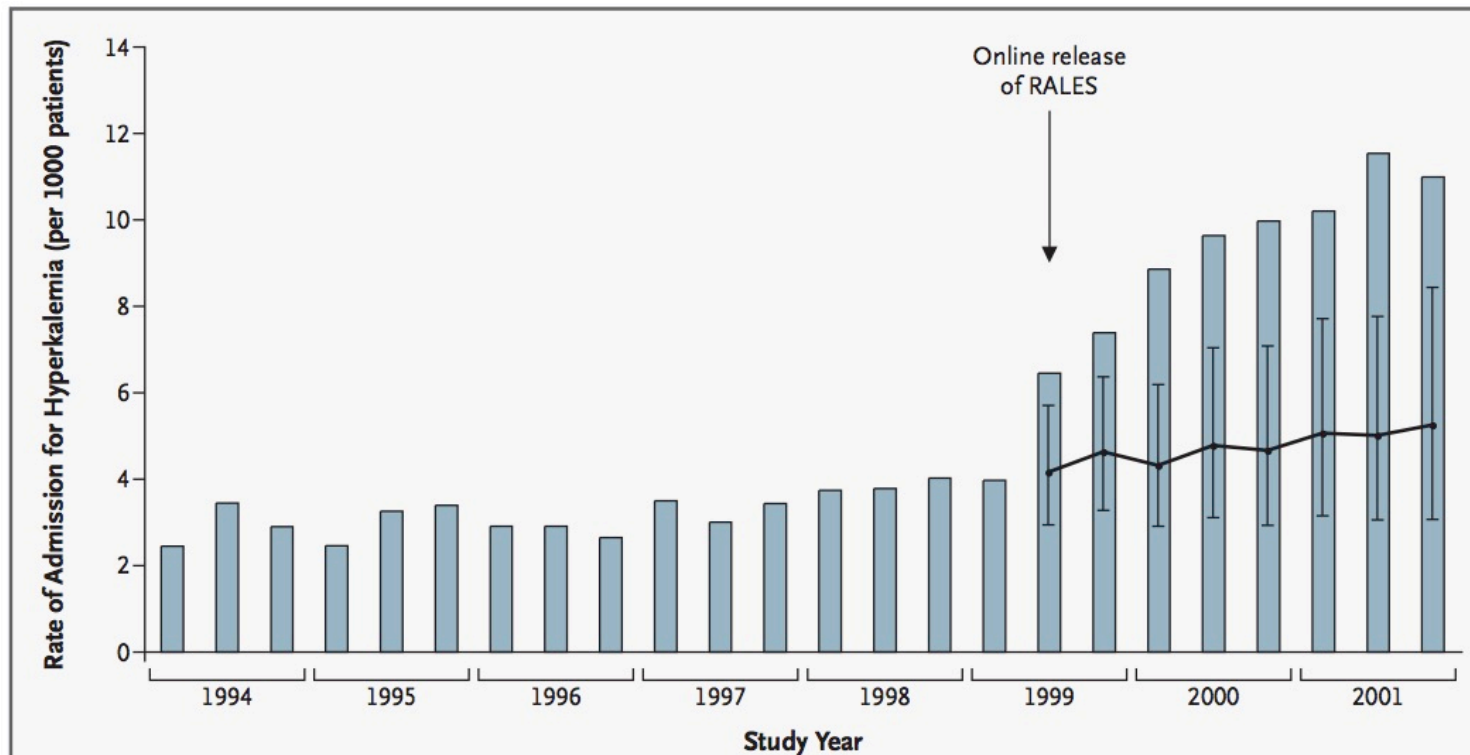


Figure 2. Rate of Hospital Admission for Hyperkalemia among Patients Recently Hospitalized for Heart Failure Who Were Receiving ACE Inhibitors.

Each bar shows the rate of hospital admission for hyperkalemia per 1000 patients during one four-month interval. The line beginning in the second interval of 1999 shows projected admission rates for hyperkalemia derived from interventional ARIMA models, with I bars representing the 95 percent confidence intervals.

UN CAVEAT SULLA RAPPRESENTATIVITÀ

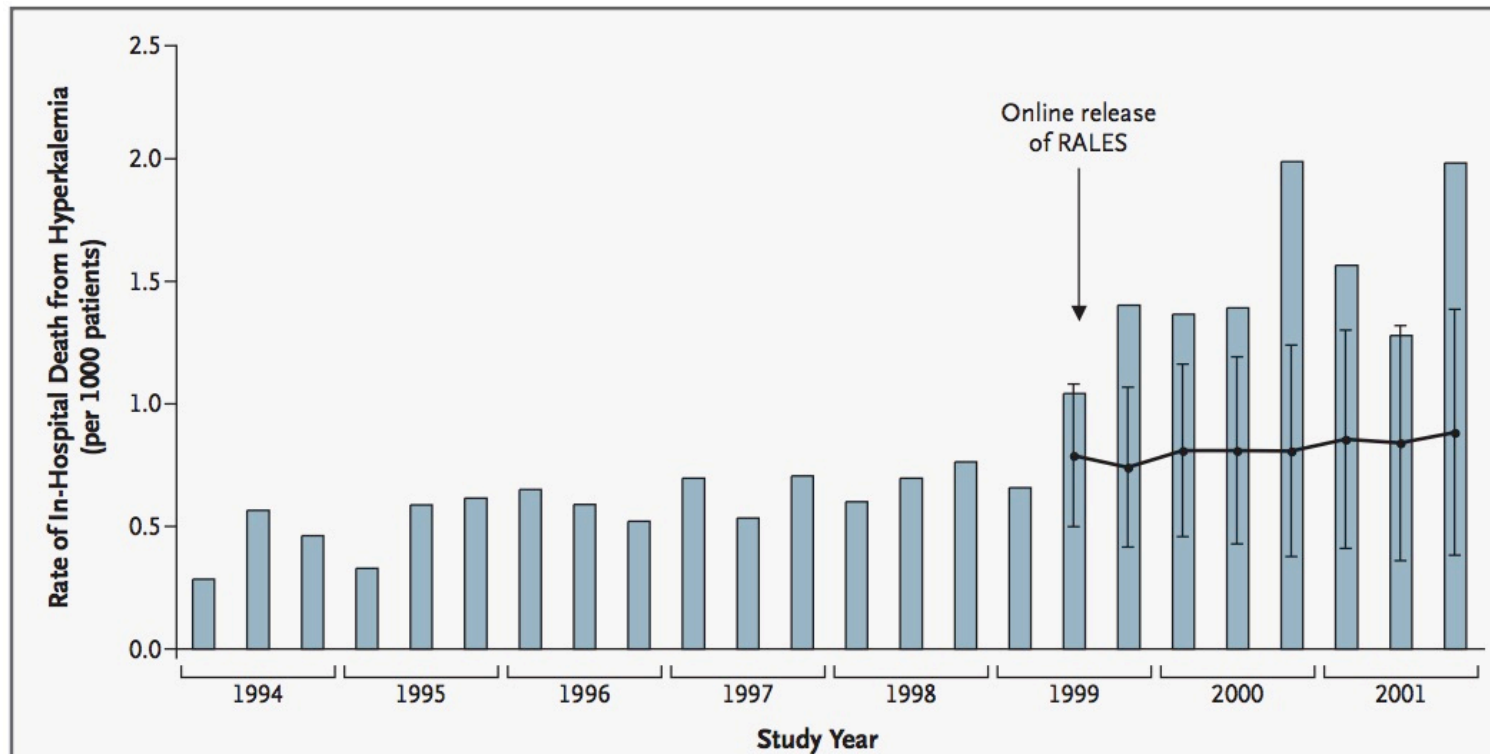


Figure 3. Rate of In-Hospital Death Associated with Hyperkalemia among Patients Recently Hospitalized for Heart Failure Who Were Receiving ACE Inhibitors.

Each bar shows the rate of in-hospital death associated with hyperkalemia per 1000 patients during one four-month interval. The line beginning in the second interval of 1999 shows projected death rates derived from interventional ARIMA models, with I bars representing the 95 percent confidence intervals.

UN CAVEAT SULLA RAPPRESENTATIVITÀ



A landmark for popperian epidemiology: refutation of the randomised Aldactone evaluation study

Elard Koch, Alvaro Otarola and Aida Kirschbaum

J. Epidemiol. Community Health 2005;59;1000-1006
doi:10.1136/jech.2004.031633

Updated information and services can be found at:
<http://jech.bmjjournals.com/cgi/content/full/59/11/1000>

PROGRAMMA

- Una visione d'insieme
- La domanda dello studio
- I soggetti dello studio
- *Le variabili e i metodi dello studio*
- L'ipotesi nulla e la stima della numerosità campionaria
- I disegni di studio
- Aspetti etici

VARIABILI

- Le variabili sono le misure* dei fenomeni di interesse

* Nota: la misura è prodotta da una *misurazione*.

TIPI DI VARIABILI

- Continue
 - Quantificabili con scala numerica *teoricamente* infinita
- Categorie
 - Nominali: senza ordine intrinseco
 - Ordinali: con ordine intrinseco
- Discrete
 - Quantificabili su scala numerica *finita*

DESCRIZIONE DELLE VARIABILI

- Continue
 - Media e deviazione standard (se distribuzione Gaussiana), mediane e percentili ...
- Categorie
 - Numero e percentuale
- Nominali
 - Numero e percentuale
- Ordinali
 - Numero e percentuale, (percentili)
- Discrete
 - Numero e percentuale, (percentili, media e deviazione standard)

TEORIA DELLA MISURAZIONE (SÌ, TEORIA!)

SCIENCE

Vol. 103, No. 2684

Friday, June 7, 1946

On the Theory of Scales of Measurement

S. S. Stevens

Director, Psycho-Acoustic Laboratory, Harvard University

TEORIA DELLA MISURAZIONE (SÌ, TEORIA!)

TABLE 1

Scale	Basic Empirical Operations	Mathematical Group Structure	Permissible Statistics (invariantive)
NOMINAL	Determination of equality	<i>Permutation group</i> $x' = f(x)$ $f(x)$ means any one-to-one substitution	Number of cases Mode Contingency correlation
ORDINAL	Determination of greater or less	<i>Isotonic group</i> $x' = f(x)$ $f(x)$ means any monotonic increasing function	Median Percentiles
INTERVAL	Determination of equality of intervals or differences	<i>General linear group</i> $x' = ax + b$	Mean Standard deviation Rank-order correlation Product-moment correlation
RATIO	Determination of equality of ratios	<i>Similarity group</i> $x' = ax$	Coefficient of variation

TEORIA DELLA MISURAZIONE (SÌ, TEORIA!)

Nominal, Ordinal, Interval, and Ratio Typologies Are Misleading

PAUL F. VELLEMAN and LELAND WILKINSON*

The American Statistician, February 1993, Vol. 47, No. 1

TEORIA DELLA MISURAZIONE (SÌ, TEORIA!)

- “Statistics programs based on Stevens’s typology suggest that doing statistics is simply a matter of declaring the scale type of data and picking a model”
- “*Worse*, they assert that the scale type is evident from the data independent of the questions asked of the data”
- “They thus restrict the questions that may be asked of the data”
- “Such restrictions lead to bad data analysis and bad science”

TEORIA DELLA MISURAZIONE (SÌ, TEORIA!)

J. R. Statist. Soc. A (1996)
159, Part 3, pp. 445–492

Statistics and the Theory of Measurement

By D. J. HAND†

The Open University, Milton Keynes, UK

[*Read before The Royal Statistical Society on Wednesday, March 20th, 1996, the President,
Professor A. F. M. Smith, in the Chair*]

SUMMARY

Just as there are different interpretations of probability, leading to different kinds of inferential statements and different conclusions about statistical models and questions, so there are different theories of measurement, which in turn may lead to different kinds of statistical model and possibly different conclusions. This has led to much confusion and a long running debate about when different classes of statistical methods may legitimately be applied. This paper outlines the major theories of measurement and their relationships and describes the different kinds of models and hypotheses which may be formulated within each theory. One general conclusion is that the domains of applicability of the two major theories are typically different, and it is this which helps apparent contradictions to be avoided in most practical applications.

Keywords: CLASSICAL MEASUREMENT; MEASUREMENT THEORY; OPERATIONAL MEASUREMENT; REPRESENTATIONAL MEASUREMENT; STATISTICAL MODELS; TRANSFORMATIONS

TIPI DI VARIABILI

- Le variabili con il maggior contenuto di informazione vanno preferite *e analizzate come tali*:

continue >> ordinali >> nominali

NON PERDERE INFORMAZIONE

Psychological Methods
2002, Vol. 7, No. 1, 19–40

Copyright 2002 by the American Psychological Association, Inc.
1082-989X/02/\$5.00 DOI: 10.1037//1082-989X.7.1.19

On the Practice of Dichotomization of Quantitative Variables

Robert C. MacCallum, Shaobo Zhang, Kristopher J. Preacher, and Derek D. Rucker
Ohio State University

The authors examine the practice of dichotomization of quantitative measures, wherein relationships among variables are examined after 1 or more variables have been converted to dichotomous variables by splitting the sample at some point on the scale(s) of measurement. A common form of dichotomization is the median split, where the independent variable is split at the median to form high and low groups, which are then compared with respect to their means on the dependent variable. The consequences of dichotomization for measurement and statistical analyses are illustrated and discussed. The use of dichotomization in practice is described, and justifications that are offered for such usage are examined. The authors present the case that dichotomization is rarely defensible and often will yield misleading results.

NON PERDERE INFORMAZIONE

STATISTICS IN MEDICINE

Statist. Med. 2006; **25**:127–141

Published online 11 October 2005 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/sim.2331

Dichotomizing continuous predictors in multiple regression: a bad idea

Patrick Royston^{1,*†}, Douglas G. Altman² and Willi Sauerbrei³

¹*MRC Clinical Trials Unit, 222 Euston Road, London NW1 2DA, U.K.*

²*Centre for Statistics in Medicine, University of Oxford, Wolfson College Annexe, Linton Road, Oxford OX2 6UD, U.K.*

³*Institute of Medical Biometry and Medical Informatics, University Hospital of Freiburg, Stefan-Meier-Str. 25, 79104 Freiburg, Germany*

SUMMARY

In medical research, continuous variables are often converted into categorical variables by grouping values into two or more categories. We consider in detail issues pertaining to creating just two groups, a common approach in clinical research. We argue that the simplicity achieved is gained at a cost; dichotomization may create rather than avoid problems, notably a considerable loss of power and residual confounding. In addition, the use of a data-derived ‘optimal’ cutpoint leads to serious bias. We illustrate the impact of dichotomization of continuous predictor variables using as a detailed case study a randomized trial in primary biliary cirrhosis. Dichotomization of continuous data is unnecessary for statistical analysis and in particular should not be applied to explanatory variables in regression models. Copyright © 2005 John Wiley & Sons, Ltd.

NON PERDERE INFORMAZIONE

Bennette and Vickers *BMC Medical Research Methodology* 2012, **12**:21
<http://www.biomedcentral.com/1471-2288/12/21>



DEBATE

Open Access

Against quantiles: categorization of continuous variables in epidemiologic research, and its discontents

Caroline Bennette¹ and Andrew Vickers^{2*}

Abstract

Background: Quantiles are a staple of epidemiologic research: in contemporary epidemiologic practice, continuous variables are typically categorized into tertiles, quartiles and quintiles as a means to illustrate the relationship between a continuous exposure and a binary outcome.

Discussion: In this paper we argue that this approach is highly problematic and present several potential alternatives. We also discuss the perceived drawbacks of these newer statistical methods and the possible reasons for their slow adoption by epidemiologists.

Summary: The use of quantiles is often inadequate for epidemiologic research with continuous variables.

NON PERDERE INFORMAZIONE

The screenshot shows the top of a web page for the Department of Biostatistics at Vanderbilt University. The header includes the department's name and logo, a search bar, and links to 'Register' and 'Log In'. Below the header, a breadcrumb trail indicates the current page location: 'You are here: Vanderbilt Biostatistics Wiki > Main Web > Education > HandoutsBioRes > BioMod > CatContinuous (07 Jul 2015, FrankHarrell)'. The main content area displays the title 'Problems Caused by Categorizing Continuous Variables | References | Key Reference | Key Reference for Response Variables' in a mix of red and blue text. To the left, a sidebar contains links to 'Main', 'Department Home Page', and 'Wiki Home'. To the right of the title, there are 'Edit' and 'Attach' buttons.

DEPARTMENT OF BIostatISTICS
VANDERBILT UNIVERSITY

Jump Search Register | Log In

Main You are here: Vanderbilt Biostatistics Wiki > Main Web > Education > HandoutsBioRes > BioMod > CatContinuous (07 Jul 2015, FrankHarrell) Edit Attach

Department Home Page

Wiki Home
The Department
People

Problems Caused by Categorizing Continuous Variables | References | Key Reference | Key Reference for Response Variables

www.giorgiobedogni.it

<http://biostat.mc.vanderbilt.edu/wiki/Main/CatContinuous>

LA “TRIAD” DEI METODI

- Precisione
- Accuratezza
- Validità

LA “TRIADE” DEI METODI

- *Precisione*
- Accuratezza
- Validità

PRECISIONE

- Valuta il grado in cui misurazioni ripetute della stessa variabile producono il medesimo valore
- La precisione è influenzata dall'*errore casuale*: più grande è l'errore casuale, più bassa è la precisione

PRECISIONE E OPERATORE

- Precisione intra-operatore
 - Lo stesso operatore effettua la stessa misura sullo stesso soggetto
- Precisione inter-operatore
 - Operatori differenti effettuano la stessa misura sullo stesso soggetto

PRECISIONE E STRUMENTO

- Precisione intra-strumentale
 - Lo stesso strumento viene utilizzato per effettuare la stessa misura sullo stesso soggetto
- Precisione inter-strumentale
 - Strumenti diversi vengono utilizzati per effettuare la stessa misura sullo stesso soggetto

COME SI VALUTA LA PRECISIONE?

- La misura più utilizzata per le variabili continue è il coefficiente di variazione (CV), che si ottiene dividendo la deviazione standard (SD) per la media (X):

$$CV = SD / X$$

COME SI VALUTA LA PRECISIONE?

- Le misure più utilizzate per le variabili categoriche sono l'agreement percentuale e la statistica k (semplice o pesata)

COME CONTROLLARE LA PRECISIONE?

- Standardizzazione dei metodi di misurazione
- Training e certificazione degli operatori
- Controllo e taratura degli strumenti
- Impiego di strumenti automatici
- Ripetizione delle misurazioni

LA “TRIADE” DEI METODI

- Precisione
- *Accuratezza*
- Validità

ACCURATEZZA

- Valuta il grado in cui la misura ottenuta si avvicina al valore reale
- L'accuratezza è influenzata dall'errore sistematico (*bias*): più grande è l'errore, più bassa è l'accuratezza

DA DOVE ORIGINA IL BIAS

- Operatore
 - Distorsione conscia o inconscia nella percezione o trascrizione della misurazione
- Soggetto
 - Distorsione della misurazione dovuta al soggetto
- Metodo
 - Malfunzionamento dello strumento

COME SI VALUTA L'ACCURATEZZA?

- In senso stretto, è necessario il confronto con una tecnica di riferimento (“gold standard”)
- Quando la tecnica di riferimento non è disponibile, si considera la *validità*

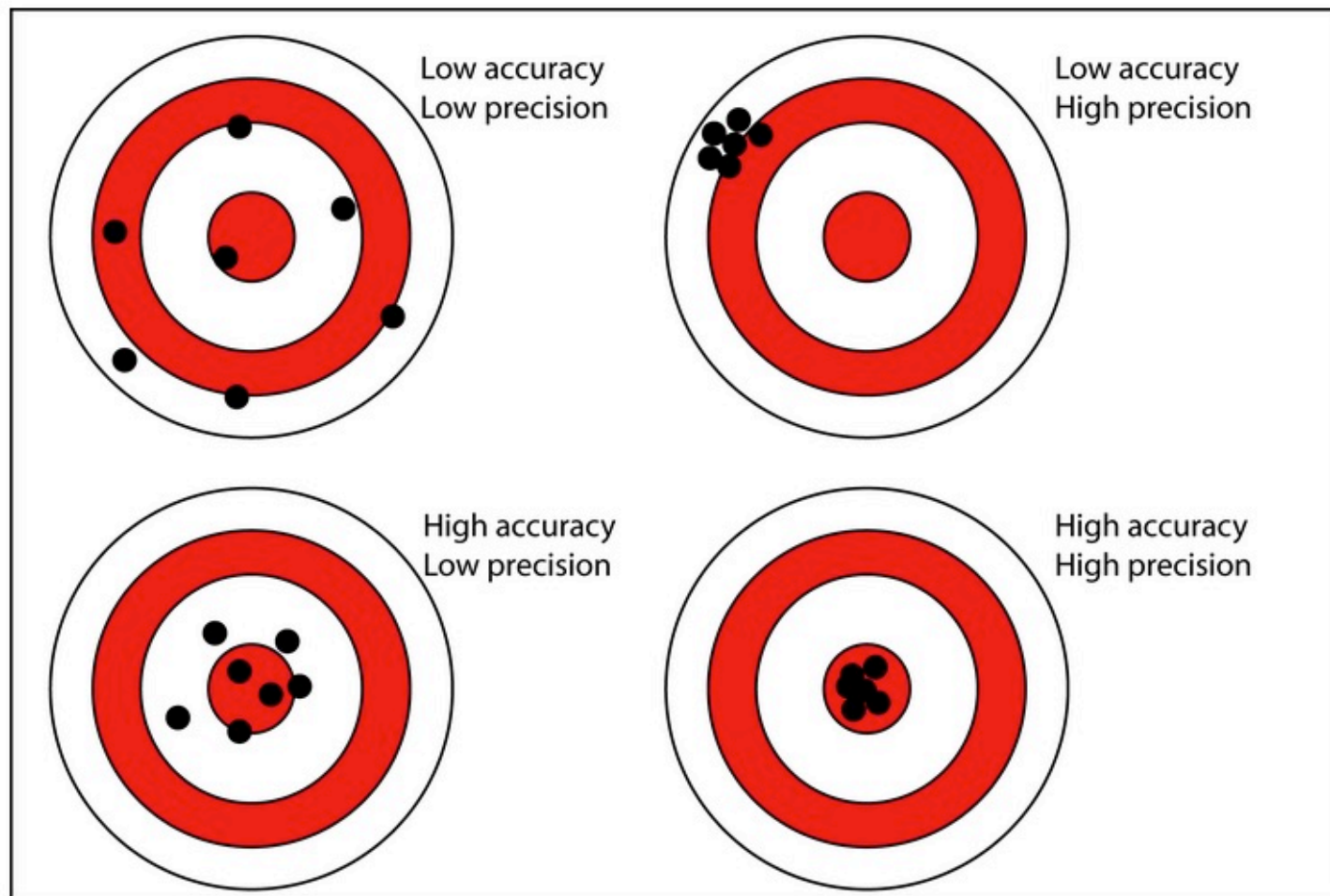
COME CONTROLLARE ACCURATEZZA E PRECISIONE?

- Standardizzazione dei metodi di misurazione
- Training e certificazione degli operatori
- Controllo e taratura degli strumenti
- Impiego di strumenti automatici

COME CONTROLLARE L'ACCURATEZZA?

- Ridurre il grado di interferenza con la vita del soggetto
- “Cecità” (*non* è una prerogativa dei trial clinici)
- Calibrazione dello strumento contro “gold standard”

PRECISIONE E ACCURATEZZA: UNA METAFORA



LA “TRIAD” DEI METODI

- Precisione
- Accuratezza
- *Validità*

VALIDITÀ

- Da considerare quando non esiste una tecnica di riferimento (“gold standard”) per la misurazione di una variabile di interesse

VALIDITÀ

- “It is fair to say that the process is often *less scientific and conclusive* than the world ‘validation’ connotes”

Hulley SB *et al.* *Designing Clinical Research*. Lippincot: Philadelphia 2007, p. 43.

VALIDITÀ*

- Content validity
- Construct validity
- Criterion-related validity

*A riprova dell'affermazione di Hulley: la nomenclatura è tutt'altro che standardizzata.

VALIDITÀ

- *Content validity*
- Construct validity
- Criterion-related validity

VALIDITÀ

- Content validity
 - Face validity
 - Giudizio soggettivo basato sul “buon senso”
 - Sampling validity
 - Giudizio sul fatto che lo strumento consideri o meno le principali variabili relative al fenomeno

VALIDITÀ

- Content validity
- *Construct validity*
- Criterion-related validity

VALIDITÀ

- Construct validity
 - Attinenza ai concetti teorici (“costrutti”) relativi all’outcome di interesse

VALIDITÀ

- Content validity
- Construct validity
- *Criterion-related validity*

VALIDITÀ

- Criterion-related validity
 - Grado di associazione con un criterio esterno

UNA “TETRADE” AGGIUNTIVA

- Sensibilità
- Specificità
- Appropriatezza
- Distribuzione delle risposte

UNA “TETRADE” AGGIUNTIVA

- *Sensibilità*
- Specificità
- Appropriatezza
- Distribuzione delle risposte

SENSIBILITÀ*

- Valuta il grado entro il quale il metodo impiegato consente di rilevare *differenze importanti* per il ricercatore

*Da non confondere con la “sensibilità” dei test di laboratorio né quella dei test diagnostici (“true positive fraction”)

UNA “TETRADE” AGGIUNTIVA

- Sensibilità
- *Specificità*
- Appropriatezza
- Distribuzione delle risposte

SPECIFICITÀ*

- Valuta il grado entro il quale il metodo impiegato rappresenta *solo* la misura di interesse
- (Domanda: sto misurando davvero soltanto o *soprattutto* il fenomeno che mi interessa?)

*Da non confondere con la “specificità” dei test diagnostici (“true negative fraction”).

UNA “TETRADE” AGGIUNTIVA

- Sensibilità
- Specificità
- *Appropriatezza*
- Distribuzione delle risposte

APPROPRIATEZZA*

- Valuta il grado in cui la misura è appropriata per l'*obiettivo* dello studio

* Da non confondere con l'“appropriatezza clinico-organizzativa”

DISTRIBUZIONE DELLE RISPOSTE

- Una misura *discriminante* produce valori alti in alcuni soggetti e bassi in altri
- È necessario uno studio pilota *sulla popolazione di interesse* prima di procedere con la validazione su larga scala del metodo di misurazione

PROGRAMMA

- Una visione d'insieme
- La domanda dello studio
- I soggetti dello studio
- Le variabili e i metodi dello studio
- *L'ipotesi nulla e la stima della numerosità campionaria*
- I disegni di studio
- Aspetti etici

STIMA DELLA NUMEROSITÀ CAMPIONARIA: PREMESSA

- Quanto diremo in questo modulo (M1) sull'ipotesi nulla è finalizzato *unicamente* al calcolo della numerosità campionaria

STIMA DELLA NUMEROSITÀ CAMPIONARIA: PREMESSA

- L'ipotesi nulla poggia su un'interpretazione frequentista della probabilità
- *Tutte* le interpretazioni della probabilità hanno dei limiti. L'importante è conoscerli e *non* schierarsi in maniera dogmatica.
http://www.phil.vt.edu/dmayo/personal_website/
<http://andrewgelman.com/>
<https://www.ucl.ac.uk/sts/staff/gillies>
- Diremo qualcosa di più sulle differenti interpretazioni della probabilità nel modulo M3

STIMA DELLA NUMEROSITÀ CAMPIONARIA: PREMESSA

Editorial Commentary

Surrogate Science: The Idol of a Universal Method for Scientific Inference

Gerd Gigerenzer

Max Planck Institute for Human Development

Julian N. Marewski

University of Lausanne

Journal of Management

Vol. 41 No. 2, February 2015 421–440

DOI: 10.1177/0149206314547522

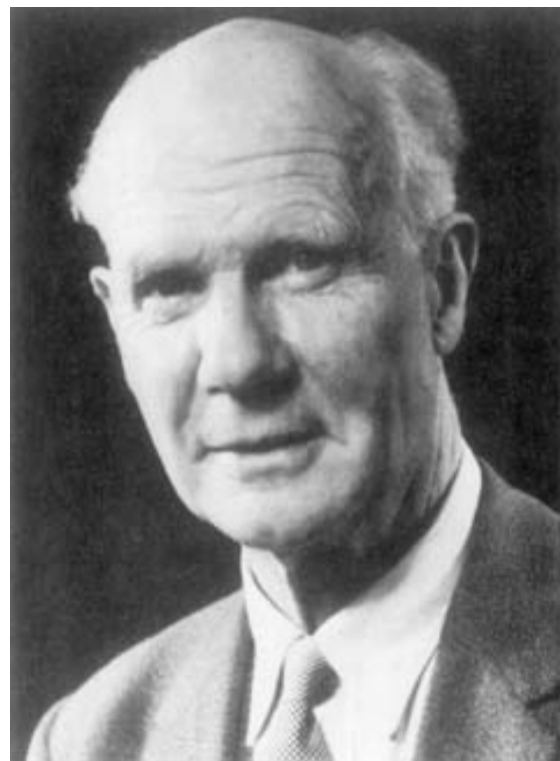
© The Author(s) 2014

The application of statistics to science is not a neutral act. Statistical tools have shaped and were also shaped by its objects. In the social sciences, statistical methods fundamentally changed research practice, making statistical inference its centerpiece. At the same time, textbook writers in the social sciences have transformed rivaling statistical systems into an apparently monolithic method that could be used mechanically. The idol of a universal method for scientific inference has been worshipped since the “inference revolution” of the 1950s. Because no such method has ever been found, surrogates have been created, most notably the quest for significant p values. This form of surrogate science fosters delusions and borderline cheating and has done much harm, creating, for one, a flood of irreproducible results. Proponents of the “Bayesian revolution” should be wary of chasing yet another chimera: an apparently universal inference procedure. A better path would be to promote both an understanding of the various devices in the “statistical toolbox” and informed judgment to select among these.

STIMA DELLA NUMEROSITÀ CAMPIONARIA: PREMESSA



Jerzy Neyman
(1894-1981)



Egon Sharpe Pearson
(1895-1980)

STIMA DELLA NUMEROSITÀ CAMPIONARIA: PREMESSA



Ronald Aylmer Fisher
(1890-1962)

STIMA DELLA NUMEROSITÀ CAMPIONARIA

- Indispensabile per

- Evitare di avere un numero troppo basso di soggetti per rispondere alla domanda dello studio (non etico!)
- (Evitare dispersione di risorse arruolando un numero troppo grande di soggetti)

STIMA DELLA NUMEROSITÀ CAMPIONARIA

- Strettamente correlata al test statistico
- Implica la trasformazione dell'ipotesi di ricerca in una più formale *ipotesi statistica*, tipicamente (ma non necessariamente) formulata come *ipotesi nulla*

L'IPOTESI NULLA (H_0)

- L'ipotesi nulla asserisce che *non* vi sono differenze dell'outcome all'interno del predittore/ dei predittori di interesse

H_0 NEL NOSTRO STUDIO

- “A 5 anni *non* ci attendiamo una differenza nel numero di fratture (outcome = O) nelle donne supplementate col calcio rispetto a quelle supplementate col placebo (predittore = P)”

H_0 NEL NOSTRO STUDIO

- Nel caso di un trial randomizzato controllato c'è un solo predittore: il trattamento
- La randomizzazione dovrebbe infatti proteggere dai confondenti noti e ignoti

QUATTRO POSSIBILITÀ

		“Verità” nella popolazione	
		Associazione PO presente	Associazione PO assente
Risultati nel campione	Rifiuto $H_0 \Rightarrow$ associazione PO presente		
	Accetto $H_0 \Rightarrow$ associazione PO assente		

QUATTRO POSSIBILITÀ

		“Verità” nella popolazione	
		Associazione PO presente	Associazione PO assente
Risultati nel campione	Rifiuto $H_0 \Rightarrow$ associazione PO presente	CORRETTO	
	Accetto $H_0 \Rightarrow$ associazione PO assente		

QUATTRO POSSIBILITÀ

		“Verità” nella popolazione	
		Associazione PO presente	Associazione PO assente
Risultati nel campione	Rifiuto $H_0 \Rightarrow$ associazione PO presente		
	Accetto $H_0 \Rightarrow$ associazione PO assente		CORRETTO

QUATTRO POSSIBILITÀ

		“Verità” nella popolazione	
		Associazione PO presente	Associazione PO assente
Risultati nel campione	Rifiuto $H_0 \Rightarrow$ associazione PO presente		ERRORE TIPO 1
	Accetto $H_0 \Rightarrow$ associazione PO assente		

QUATTRO POSSIBILITÀ

		“Verità” nella popolazione	
		Associazione PO presente	Associazione PO assente
Risultati nel campione	Rifiuto $H_0 \Rightarrow$ associazione PO presente		
	Accetto $H_0 \Rightarrow$ associazione PO assente	ERRORE TIPO 2	

ERRORE DI TIPO 1

- Rifiuto l'ipotesi nulla (= concludo che c'è associazione) quando non c'è associazione tra predittore e outcome
- La probabilità di commettere un errore di tipo 1 è detta “alpha” (α)
- (Metafora del “falso positivo”)

ERRORE DI TIPO 2

- Accetto l'ipotesi nulla (= concludo che non c'è associazione) quando c'è associazione tra predittore e outcome
- La probabilità di commettere un errore di tipo 2 è detta “beta” (β)
- (Metafora del “falso negativo”)

ERRORE DI TIPO 3

- Confondere l'errore di tipo 1 con l'errore di tipo 2 (scherzo!)

COSA È P ?

- Il valore p calcolato dai test di significatività statistica rappresenta la probabilità di osservare per caso un effetto uguale o maggiore a quello osservato se l'ipotesi nulla è vera

COSA È LA SIGNIFICATIVITÀ STATISTICA?

- p è detto “statisticamente significativo” se $< \alpha$
- α è 0.05 (5%) nelle scienze biologiche solo per convenzione

COSA È LA POTENZA ?

- β è 0.20 (20%) o 0.10 (10%) nelle scienze biologiche solo per convenzione
- Potenza = $1 - \beta$
 - $1 - 0.20 = 0.80 = 80\%$
 - $1 - 0.10 = 0.90 = 90\%$

L'IPOTESI ALTERNATIVA

- È la negazione dell'ipotesi nulla
- È accettata quando l'ipotesi nulla è rifiutata
- *Non può essere testata direttamente*

COSA È P ? (TORNIAMO INDIETRO)

- Il valore p calcolato dai test di significatività statistica rappresenta la probabilità di osservare per caso un effetto uguale o maggiore a quello osservato se l'ipotesi nulla è vera

DIREZIONE DELL'IPOTESI

- Ipotesi “a una direzione” (“one-sided”)
 - Specifica la direzione dell'associazione PO
 - Esempio: ci attendiamo *meno* fratture nel gruppo delle donne supplementate con calcio.
- Ipotesi “a due direzioni” (“two-sided”)
 - Non specifica la direzione dell'associazione PO
 - Esempio: ci attendiamo una frequenza *diversa* di fratture nei due gruppi.

DIREZIONE DELL'IPOTESI

- L'ipotesi uni-direzionale aumenta la potenza dello studio a parità di soggetti ma *riduce* la possibilità di testare la differenza nei due sensi:
 - E se le donne supplementate col calcio avessero più fratture?
 - Ricordate che *non* sappiamo ancora se il calcio protegge, altrimenti non avremmo fatto lo studio !

DIREZIONE DELL'IPOTESI

- Nella grande maggioranza dei casi, vi troverete ad utilizzare *ipotesi bi-direzionali*

STIMA DELLA NUMEROSITÀ CAMPIONARIA

○ Procedura

- Definisci l'ipotesi nulla (una o due direzioni)
- Scegli il test statistico appropriato in base alla variabile di outcome e al predittore/ai predittori (ricorda la critica fatta a Stevens!)
- Scegli una dimensione e una variabilità dell'effetto
- (Adesso e solo adesso) fissa α e β

STIMA DELLA NUMEROSITÀ CAMPIONARIA

- Non partire* col calcolo della numerosità campionaria finchè non hai risposto a questa domanda:

“Qual è la differenza minima clinicamente (o biologicamente) rilevante”?

L'imperativo è appropriato perché questo è un comandamento.

STIMA DELLA NUMEROSITÀ CAMPIONARIA

- “Qual è la differenza minima clinicamente (o biologicamente) rilevante”?

Ad esempio: *potrei* pensare che per dichiarare il calcio efficace nella prevenzione delle fratture, ci debba essere una riduzione della loro incidenza relativa di almeno il 20% nel gruppo trattato con calcio rispetto a quello trattato con placebo.

STIMA DELLA NUMEROSITÀ CAMPIONARIA

- Minore è il valore di α , maggiore è il numero di soggetti da arruolare
- Maggiore è il valore della potenza (= minore è il valore di β), maggiore è il numero di soggetti da arruolare
- Minore è la dimensione dell'effetto, maggiore è il numero di soggetti da arruolare
- Minore è la variabilità attorno all'effetto, maggiore è il numero di soggetti da arruolare

STIMA DELLA NUMEROSITÀ CAMPIONARIA

- A questo punto, il vostro compito è finito e inizia quello dello statistico (fuorché in casi molto semplici per i quali *potreste* ritenervi soddisfatti)
- Se non avete la padronanza di questi concetti di base, non potrete farvi aiutare dallo statistico, che vi porrà la domanda sulla dimensione dell'effetto che voi ve la siate posta o meno ...

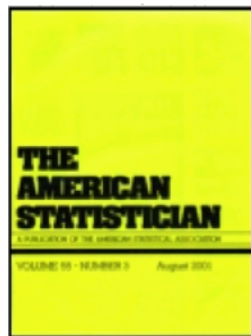
STIMA DELLA NUMEROSITÀ CAMPIONARIA

“To consult a statistician after an experiment is finished is often merely to ask him to conduct a post mortem examination. He can perhaps say what the experiment died of”

Sir Ronald Aylmer Fisher

<http://www.stat.tamu.edu/~lennox/other/humor.html>

P-VALUES - ASA POSITION PAPER



The American Statistician



ISSN: 0003-1305 (Print) 1537-2731 (Online) Journal homepage: <http://amstat.tandfonline.com/loi/utas20>

The ASA's Statement on p-Values: Context, Process, and Purpose

Ronald L. Wasserstein & Nicole A. Lazar

To cite this article: Ronald L. Wasserstein & Nicole A. Lazar (2016) The ASA's Statement on p-Values: Context, Process, and Purpose, The American Statistician, 70:2, 129-133, DOI: [10.1080/00031305.2016.1154108](https://doi.org/10.1080/00031305.2016.1154108)

To link to this article: <http://dx.doi.org/10.1080/00031305.2016.1154108>

P-VALUES - ASA POSITION PAPER

- “Q: Why do so many colleges and grad schools teach $p = 0.05$?

A: Because that’s still what the scientific community and journal editors use.

Q: Why do so many people still use $p = 0.05$?

A: Because that’s what they were taught in college or grad school.”

(George Cobb, 2014, on the ASA discussion forum)

P-VALUES - ASA POSITION PAPER

- “The ASA Board was also stimulated by highly visible discussion over the last few years”
- “The ASA has *not* previously taken positions on specific matters of statistical practice”

P-VALUES - ASA POSITION PAPER

- “That turned out to be relatively easy to do [to find points of agreement], but *it was just as easy* to find points of intense disagreement”.

DÉJÀ VU?

Editorial Commentary

Surrogate Science: The Idol of a Universal Method for Scientific Inference

Gerd Gigerenzer

Max Planck Institute for Human Development

Julian N. Marewski

University of Lausanne

Journal of Management

Vol. 41 No. 2, February 2015 421–440

DOI: 10.1177/0149206314547522

© The Author(s) 2014

The application of statistics to science is not a neutral act. Statistical tools have shaped and were also shaped by its objects. In the social sciences, statistical methods fundamentally changed research practice, making statistical inference its centerpiece. At the same time, textbook writers in the social sciences have transformed rivaling statistical systems into an apparently monolithic method that could be used mechanically. The idol of a universal method for scientific inference has been worshipped since the “inference revolution” of the 1950s. Because no such method has ever been found, surrogates have been created, most notably the quest for significant p values. This form of surrogate science fosters delusions and borderline cheating and has done much harm, creating, for one, a flood of irreproducible results. Proponents of the “Bayesian revolution” should be wary of chasing yet another chimera: an apparently universal inference procedure. A better path would be to promote both an understanding of the various devices in the “statistical toolbox” and informed judgment to select among these.

P-VALUES - ASA POSITION PAPER

- Ci torneremo nel modulo M3
- Oltre al position paper, leggete i commenti degli *Invited Discussants* disponibili nell'appendice elettronica.
- Consiglio in particolare i commenti di: Gelman A (“bayesiano [atipico]”), Mayo D (“statistica dell'errore”), Senn S (“frequentista [atipico]”). Tutti concordano sull'*abuso* dei p -values, non sulla loro inutilità a priori.

P-VALUES – ASA POSITION PAPER – ALLEGATO

Statistical Tests, *P*-values, Confidence Intervals, and Power: A Guide to Misinterpretations

Sander GREENLAND, Stephen J. SENN, Kenneth J. ROTHMAN, John B. CARLIN, Charles POOLE,
Steven N. GOODMAN, and Douglas G. ALTMAN

PROGRAMMA

- Una visione d'insieme
- La domanda dello studio
- I soggetti dello studio
- Le variabili e i metodi dello studio
- L'ipotesi nulla e la stima della numerosità campionaria
- *I disegni di studio*
- Aspetti etici

DUE TIPI PRINCIPALI DI STUDI

- Studio osservazionale
 - “Osservo”
- Studio sperimentale
 - “Intervengo”

TRE TIPI PRINCIPALI DI STUDI OSSERVAZIONALI

- Studio di coorte (“cohort study”)
 - Un gruppo di soggetti seguito nel tempo
- Studio trasversale (“cross-sectional study”)
 - Un gruppo di soggetti valutati in un periodo di tempo determinato
- Studio caso-controllo (“case-control study”)
 - Due gruppi di soggetti definiti in base alla presenza *vs.* assenza dell’outcome di interesse

L'UNICO “VERO” STUDIO SPERIMENTALE

- Studio randomizzato controllato (“randomized controlled trial”)
 - Due gruppi di soggetti *scelti in maniera casuale*

STUDIO DI COORTE

- Un gruppo di soggetti (“coorte”) viene seguito per un periodo di tempo prefissato
 - Es. un gruppo di donne in post-menopausa seguito per 5 anni

STUDIO DI COORTE

- Qual è l'*incidenza* (= numero di nuovi casi) dell'outcome?
 - Quante donne hanno nuove fratture nel corso dei 5 anni dello studio?
- C'è associazione tra i predittori misurati durante (o prima) lo studio e l'outcome?
 - Es. l'introito di calcio durante lo (o prima dello) studio è associato con la comparsa di nuove fratture?

STUDIO DI COORTE

- Ha un vantaggio principale
 - È lo studio osservazionale che consente di valutare meglio il rapporto di causa-effetto (il predittore viene misurato *prima* dell'outcome)
- Ha due limiti importanti
 - Non può essere utilizzato per malattie rare
 - Richiede notevoli risorse

STUDIO DI COORTE

- Ha due “varietà” principali
 - Prospettiva (migliore): l'arruolamento dei soggetti è fatto sulla base di dati raccolti all'inizio dello studio
 - Retrospettiva* (peggiore): l'arruolamento dei soggetti è fatto sulla base di dati raccolti *in precedenza*

*Scherzosamente: TROHOC

STUDIO TRASVERSALE

- Un gruppo di soggetti viene studiato in un periodo di tempo determinato
 - Es. Quante donne hanno avuto una frattura dopo la menopausa in una determinata popolazione?

STUDIO TRASVERSALE

- Qual è la *prevalenza* (= numero di casi attuali) dell'outcome?
 - Quante donne hanno avuto una frattura dopo la menopausa nella popolazione, così come rilevato nel periodo in cui lo studio è stato eseguito?

STUDIO TRASVERSALE

- Ha un vantaggio principale
 - È il disegno di studio più semplice
- Ha un limite importante
 - Non consente di stabilire relazioni causa-effetto perché il predittore viene misurato *assieme* all'outcome

STUDIO CASO-CONTROLLO

- Un gruppo di soggetti con l'outcome di interesse (“casi”) viene confrontato con un gruppo senza di esso (“controlli”)
 - Es. confronto un gruppo di donne che hanno avuto fratture (“outcome”) dopo la menopausa (“casi”) con uno di donne che non ne hanno avute (“controlli”)

STUDIO CASO-CONTROLLO

- La caratteristica di interesse (predittore) è associata all'outcome?
 - Le donne con e senza fratture dopo la menopausa hanno (avevano in passato) un diverso introito di calcio?

STUDIO CASO-CONTROLLO

- Vantaggi principali
 - Ottimo per le malattie rare
 - Utile per identificare associazioni da valutare in studi successivi (studi di coorte o studi randomizzati controllati)
- Limiti principali
 - Non consente (*di regola*) di stabilire relazioni causa-effetto (il predittore viene misurato *assieme* all'outcome)

STUDIO RANDOMIZZATO CONTROLLATO

- *Randomizzato*: perché il trattamento (predittore) è distribuito *a caso* tra due (o più) gruppi
- *Controllato*: perché un gruppo ha ricevuto il trattamento di controllo
 - *Sperimentale* è il gruppo che ha ricevuto il trattamento sperimentale

STUDIO RANDOMIZZATO CONTROLLATO

- L'*outcome* è dunque valutato all'interno di due o più gruppi scelti a caso
- L'assegnazione casuale “protegge” lo studio dall'effetto di fattori confondenti noti e *ignoti*

STUDIO RANDOMIZZATO CONTROLLATO

- Ha un vantaggio unico
 - Offre la migliore protezione dall'effetto di fattori confondenti (noti e *ignoti*)
- Ha due limiti evidenti
 - Logistica complicata
 - Costo elevato

CAUSALITÀ E STUDI OSSERVAZIONALI

- Il miglior tipo di studio per stabilire un rapporto di *causa-effetto* è lo studio randomizzato controllato
- La relazione causa-effetto deve essere interpretata con cautela nel caso di uno studio osservazionale

MA... COSA È UNA CAUSA?

- I filosofi occidentali dibattono cosa sia una “causa” da almeno 2500 anni
- Il concetto di causa è estraneo al mondo della matematica*

Baber RL. *The Language of Mathematics*. Wiley, 2011.

Freedman D. *Statistical models: theory and practice*. CUP, 2009.

*Per un'autorevole parere contrario si veda:

https://en.wikipedia.org/wiki/Judea_Pearl

MA... COSA È UNA CAUSA?

- Una causa è definita *operativamente* attraverso un'associazione praticamente rilevante con un effetto di interesse

Vineis P. *Int J Epidemiol* 2009;38:675 (“The skeptical epidemiologist”).



MA... COSA È UNA CAUSA?

Emerging Themes in Epidemiology



Analytic perspective

Open Access

Causal criteria and counterfactuals; nothing more (or less) than scientific common sense

Carl V Phillips*¹ and Karen J Goodman^{1,2}

Address: ¹University of Alberta School of Public Health, Edmonton, Canada and ²Department of Medicine, Faculty of Medicine and Dentistry, University of Alberta, Edmonton, Canada

Email: Carl V Phillips* - carl.v.phillips@ualberta.ca; Karen J Goodman - karen.goodman@ualberta.ca

* Corresponding author

Published: 26 May 2006

Emerging Themes in Epidemiology 2006, **3**:5 doi:10.1186/1742-7622-3-5

Received: 22 November 2005

Accepted: 26 May 2006

MA... COSA È UNA CAUSA?

- “Two persistent myths in epidemiology are that we can use a list of ‘causal criteria’ to provide an algorithmic approach to inferring causation and that a modern ‘counterfactual model’ can assist in the same endeavor”

Phillips CV *et al. Emerg Themes Epidemiol* 2006; **3**:5.

MA... COSA È UNA CAUSA?

- “We argue that these are neither criteria nor a model, but that lists of causal *considerations* and formalizations of the counterfactual *definition* of causation are nevertheless useful tools for promoting scientific thinking”

Phillips CV *et al. Emerg Themes Epidemiol* 2006; 3:5.

MA... COSA È UNA CAUSA?

Statistics
in Medicine

Featured Article

Received: 23 October 2015, Accepted: 06 November 2015 Published online 8 December 2015 in Wiley Online Library

(wileyonlinelibrary.com) DOI: 10.1002/sim.6825

Exposure-wide epidemiology: revisiting Bradford Hill

John P. A. Ioannidis^{a,b,c,d*†}

Fifty years after **Bradford Hill** published his extremely influential criteria to offer some guides for separating causation from association, we have accumulated millions of papers and extensive data on observational research that depends on epidemiologic methods and principles. This allows us to re-examine the **accumulated empirical evidence for the nine criteria**, and to re-approach epidemiology through the lens of exposure-wide approaches. The lecture discusses the evolution of these exposure-wide approaches and tries to use the evidence from meta-epidemiologic assessments to reassess each of the nine criteria and whether they work well as guides for causation. I argue that of the nine criteria, experiment remains important and consistency (replication) is also very essential. Temporality also makes sense, but it is often difficult to document. Of the other six criteria, strength mostly does not work and may even have to be inverted: small and even tiny effects are more plausible than large effects; when large effects are seen, they are mostly transient and almost always represent biases and errors. There is little evidence for specificity in causation in nature. Biological gradient is often unclear how it should be modeled and thus difficult to prove. Coherence remains usually unclear how to operationalize. Finally, plausibility as well as analogy do not work well in most fields of investigation, and their invocation has been mostly detrimental, although exceptions may exist. Copyright © 2015 John Wiley & Sons, Ltd.

CAUSALITÀ E STUDI OSSERVAZIONALI

- Il nostro (ipotetico) studio trasversale ha “dimostrato” che un minor consumo di calcio “si associa” a un maggior numero di fratture nelle donne in post-menopausa
- Oltre a una associazione del tipo causa-effetto, ci sono altre quattro possibilità da considerare
- (Quello che diremo vale per tutti gli studi osservazionali)

ASSOCIAZIONI SPURIE

Spiegazione	Associazione	“Verità”
Errore casuale	Spuria	Il consumo pregresso di calcio e la frequenza di fratture non sono correlati
Errore sistematico	Spuria	Il consumo pregresso di calcio e la frequenza di fratture non sono correlati

ASSOCIAZIONI REALI

Spiegazione	Associazione	“Verità”
Effetto-causa	Reale	La frequenza di fratture ha determinato un maggior consumo di calcio
Confondimento	Reale	Il consumo di calcio è associato con un <i>terzo fattore</i> responsabile dell’associazione

FATTORI CONFONDENTI

- L'Epidemiologia potrebbe essere considerata la scienza (con una buona dose di arte) che cerca di controllare i fattori confondenti !

FATTORI CONFONDENTI E STUDI OSSERVAZIONALI

- Fase di disegno
 - Specificare le caratteristiche dei soggetti in base ai fattori confondenti (*oggi noti*)
 - Appaiamento
 - Stratificazione
 - (...)
- Fase di analisi (è tardi !)
 - Analisi multivariata *
 - (...)

* Nota pedante: il termine corretto è “multivariabile”

PROGRAMMA

- Una visione d'insieme
- La domanda dello studio
- I soggetti dello studio
- Le variabili e i metodi dello studio
- L'ipotesi nulla e la stima della numerosità campionaria
- I disegni di studio
- *Aspetti etici*

IL CODICE DI NORIMBERGA, 1947



LA DICHIARAZIONE DI HELSINKI, 1964



WMA Declaration of Helsinki - Ethical Principles for Medical Research Involving Human Subjects

Adopted by the 18th WMA General Assembly, Helsinki, Finland, June 1964
and amended by the:

29th WMA General Assembly, Tokyo, Japan, October 1975

35th WMA General Assembly, Venice, Italy, October 1983

41st WMA General Assembly, Hong Kong, September 1989

48th WMA General Assembly, Somerset West, Republic of South Africa, October 1996

52nd WMA General Assembly, Edinburgh, Scotland, October 2000

53rd WMA General Assembly, Washington DC, USA, October 2002 (Note of
Clarification added)

55th WMA General Assembly, Tokyo, Japan, October 2004 (Note of Clarification added)

59th WMA General Assembly, Seoul, Republic of Korea, October 2008

64th WMA General Assembly, Fortaleza, Brazil, October 2013

“TUSKEGEE STUDY OF UNTREATED SYPHILIS IN THE NEGRO MALE” – U.S. PUBLIC HEALTH SERVICE (1932-1972)

The New York Times

WORLD	U.S.	N.Y. / REGION	BUSINESS	TECHNOLOGY	SCIENCE	HEALTH	SPORTS	OPINION
-------	------	---------------	----------	------------	---------	--------	--------	---------

Article Preview

Syphilis Victims in U.S. Study Went Untreated for 40 Years; SYPHILIS VICTIMS GOT NO THERAPY

By JEAN HELLERThe Associated Press ();
July 26, 1972,
, Section , Page 1, Column , words

 PERMISSIONS

[DISPLAYING ABSTRACT]

WASHINGTON, July 25 -- For 40 years the United States Public Health Service has conducted a study in which human beings with syphilis, who were induced to serve as guinea pigs, have gone without medical treatment for the disease and a few have died of its late effects, even though an effective therapy was eventually discovered.

“TUSKEGEE STUDY OF UNTREATED SYPHYSILIS IN THE NEGRO MALE” – U.S. PUBLIC HEALTH SERVICE (1932-1972)

December 1964

The Tuskegee Study of Untreated Syphilis **The 30th Year of Observation**

DONALD H. ROCKWELL, MD; ANNE ROOF YOBBS, MD; M. BRITTAIN MOORE JR., MD

» Author Affiliations

Arch Intern Med. 1964;114(6):792-798. doi:10.1001/archinte.1964.03860120104011

Abstract

The year 1963 marks the 30th year of the long-term evaluation of the effect of untreated syphilis in the male Negro conducted by the Venereal Disease Branch, Communicable Disease Center, United States Public Health Service. This paper summarizes the information obtained in this study—well known as the "Tuskegee Study"—from earlier publications,¹⁻¹¹ reviews the status of the original study group, and reports the clinical and laboratory findings on those remaining participants who were examined in the 1963 evaluation.

TUSKEGEE STUDY OF UNTREATED SYPHYSISILIS IN THE NEGRO MALE – U.S. PUBLIC HEALTH SERVICE (1932-1972)

- 399 *uomini* (non “soggetti dello studio”) mai trattati
 - 28 deceduti per sifilide
 - (100 deceduti per cause imputabili alla sifilide)
 - +40 moglie infette
 - +19 bambini con sifilide congenita

https://en.wikipedia.org/wiki/Tuskegee_syphilis_experiment

TUSKEGEE STUDY OF UNTREATED SYPHILIS IN THE NEGRO MALE – U.S. PUBLIC HEALTH SERVICE (1932-1972)



IL BELMONT REPORT, 1979

The Belmont Report

**Office of the Secretary
Ethical Principles and Guidelines for the Protection
of Human Subjects of Research
The National Commission for the Protection of
Human Subjects of Biomedical and Behavioral
Research**

April 18, 1979

I TRE “PRINCIPI” DEL BELMONT REPORT

- Rispetto per le persone
- Beneficenza
- Giustizia

I TRE “PRINCIPI” DEL BELMONT REPORT

- *Rispetto per le persone*
- Beneficenza
- Giustizia

RISPETTO PER LE PERSONE

- Ottenere il consenso informato
- Proteggere le persone con capacità decisionale ridotta
- Mantenere l'anonimato

I TRE “PRINCIPI” DEL BELMONT REPORT

- Rispetto per le persone
- *Beneficenza*
- Giustizia

BENEFICENZA

- Disegno scientificamente valido
- Accettabile distribuzione dei rischi rispetto ai benefici

I TRE “PRINCIPI” DEL BELMONT REPORT

- Rispetto per le persone
- Beneficenza
- *Giustizia*

GIUSTIZIA

- Benefici e svantaggi devono essere distribuiti equamente

UN'IMPORTANTISSIMA ASIMMETRIA

- I soggetti arruolati nei trial clinici sono spesso gli stessi pazienti che si rivolgono al medico per avere una cura per la loro malattia
- Ciò li rende estremamente *vulnerabili*

DIFFERENZE TRA TERAPIA E RICERCA

- Scopo principale
 - Terapia: migliorare la salute del paziente
 - Ricerca: produrre conoscenza affidabile
- Beneficiario
 - Terapia: paziente (oggi)
 - Ricerca: la comunità dei pazienti (domani?)
- Presupposto
 - Terapia: impiego del trattamento migliore
 - Ricerca: incertezza sulla superiorità di un trattamento rispetto a un altro

RUOLO DEI COMITATI ETICI

- Accertarsi che
 - i rischi siano ridotti al minimo
 - i rischi siano “ragionevoli” in relazione ai benefici attesi e alla crescita attesa della conoscenza
 - la selezione dei partecipanti rispetti criteri di equità
 - il consenso informato sia ottenuto da tutti i partecipanti
 - venga mantenuto l’anonimato

CONDOTTE NON ETICHE

- Fabbricazione dei dati
- Falsificazione dei dati
- Plagiarismo

CONFLITTO DI INTERESSI

- Due tipi principali
 - Doppio ruolo ricercatore / curante del paziente
 - Economico

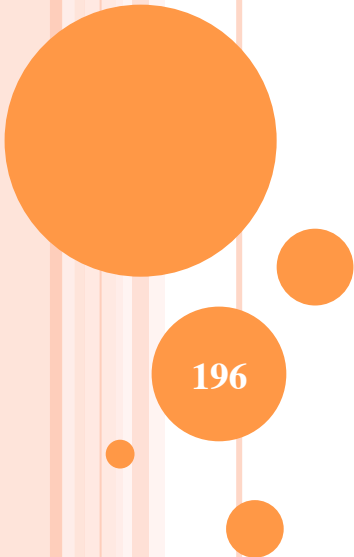
RISORSE WEB UTILI

- <http://www.giorgiobedogni.it/links/plinks.html>

M1-SESSIONE PRATICA 21-03-2017

○ Idee

1. Rapido retraining sui concetti del modulo M1
2. Lavoro a piccoli gruppi
 - Richiede aula attrezzata per gruppi e computer portatili
 - Meglio gruppi scelti in maniera casuale (ci pensa GB e poi ci spiega come ha fatto...)
 - Disegniamo studi clinici sulla base di “domande” poste da GB
 - Discutiamo insieme il materiale elaborato dai gruppi
3. Lavoro collegiale
 - Disegniamo uno studio scelto da qualcuno di noi (per le parti che siamo in grado di fare col modulo M1-teoria; il resto lo aggiungeremo strada facendo o chiederemo a GB)



**GRAZIE E ARRIVEDERCI AL
21-03-2017!**