FACULTY OF MEDICINE AND HEALTH SCIENCES

A surgeon's guide to the scientific literature

Wim Ceelen MSc MD PhD

GI Surgery - Ghent University Hospital

wim.ceelen@ugent.be @WimCeelen

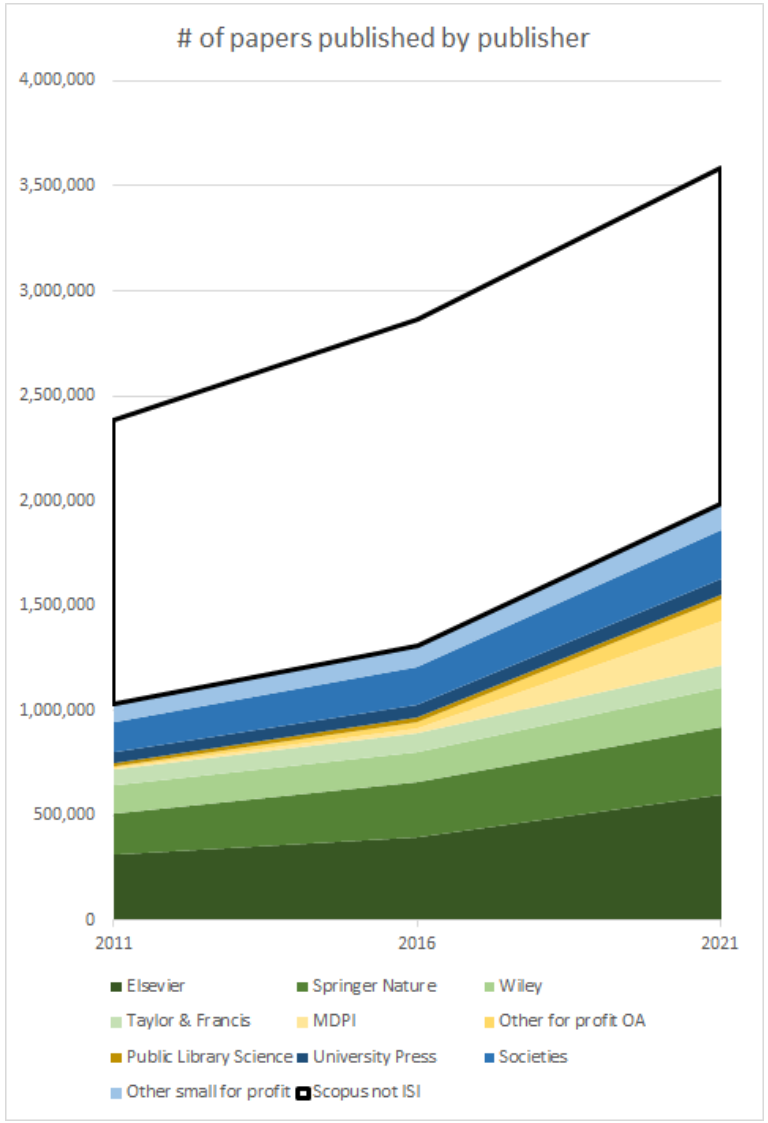
GHENT UNIVERSITY

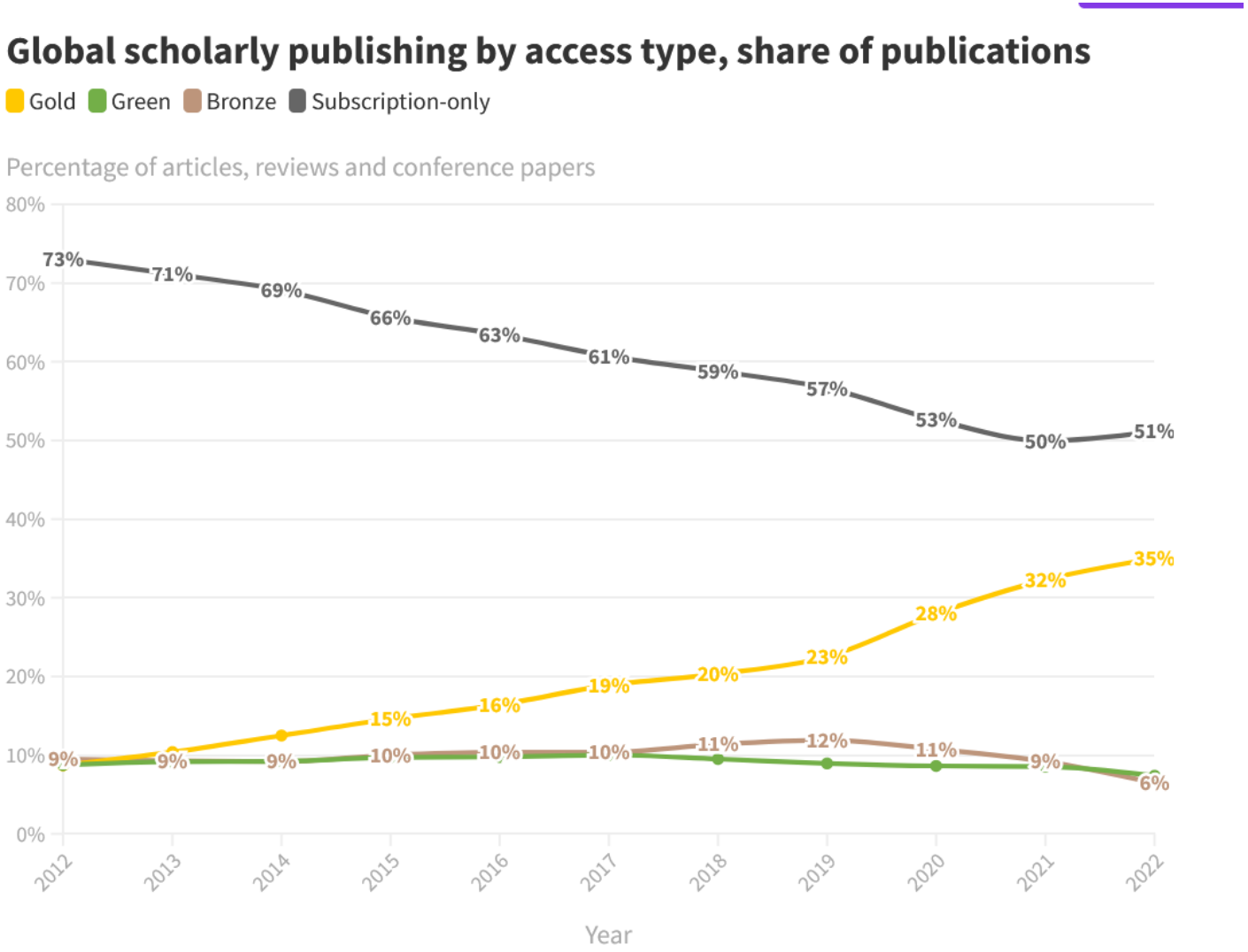
SURGX EXPERIMENTAL SURGERY LAB

CANCER RESEARCH INSTITUTE GHENT CRIG

UZ GENT

The publishing landscape





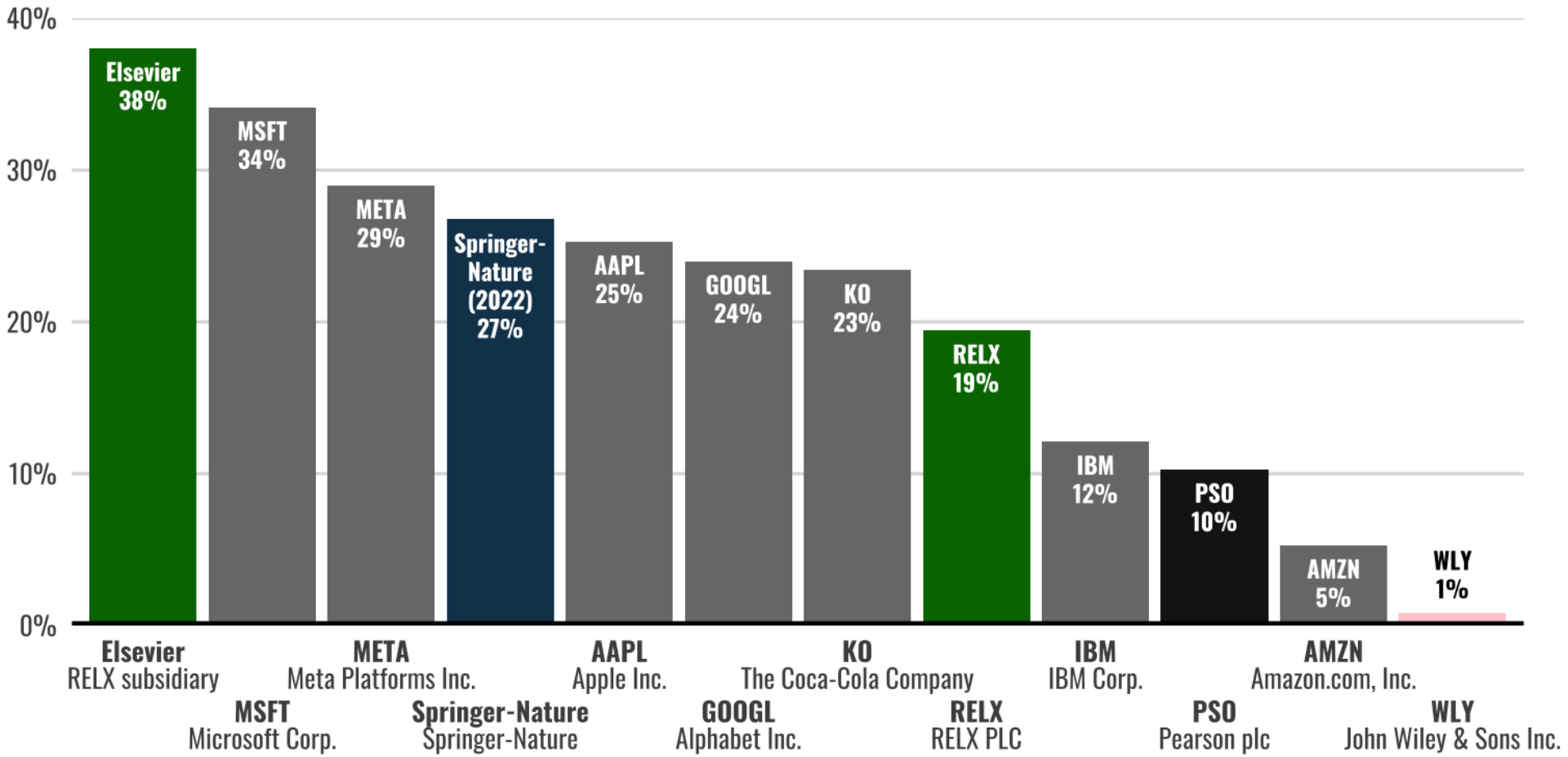
Open access publication fees for some of the biggest publishers

Fees range all the way from $200 to $12 290

Source: Elsevier, Springer-Nature publishing group

Elsevier and Springer-Nature had profit margins in 2023 that rival those of large tech companies

Springer-Nature financials are from 2022 since they haven't released their 2023 results yet



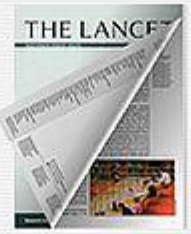
Source: Yahoo Finance, RELX Investor Presentations, Springer Nature Annual Reports

Why critical appraisal is an essential skill

- Appreciate validity (internal and external) of published research
- EBM – better healthcare
- Fight against disinformation
- Detect fraud: falsification, fabrication,…

THE LANCET

A new Lancet Series



increasing value
reducing waste
in research

The Observer
Peer review and scientific publishing

‘The situation has become appalling’: fake scientific papers push research credibility to crisis point

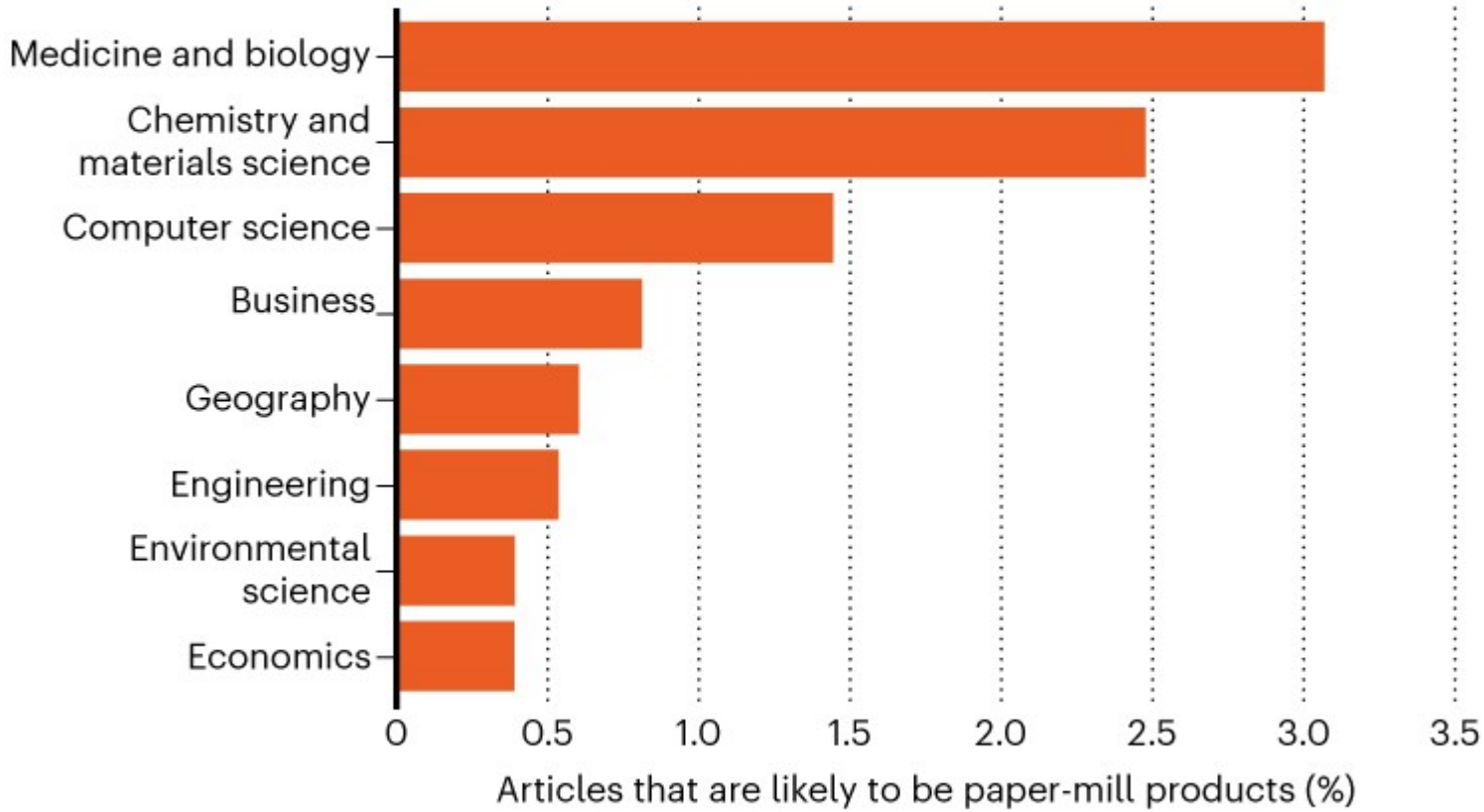
Last year, 10,000 sham papers had to be retracted by academic journals, but experts think this is just the tip of the iceberg

Robin McKie

Sat 3 Feb 2024 17.00 CET

SUBJECT BREAKDOWN

The scientific disciplines with the highest proportions of paper-mill articles are biology and medicine, and chemistry and materials science, the analysis suggests.



Subject fields from analysis of 'concepts' associated with some research articles in OpenAlex database.

ChatGPT

Contents lists available at ScienceDirect

Surfaces and Interfaces

journal homepage: www.sciencedirect.com/journal/surfaces-and-interfaces

The three-dimensional porous mesh structure of Cu-based metal-organic-framework - aramid cellulose separator enhances the electrochemical performance of lithium metal anode batteries

Manshu Zhang [a,1], Liming Wu [a,1], Tao Yang [b], Bing Zhu [a], Yangai Liu [a,*]

[a] Beijing Key Laboratory of Materials Utilization of Nonmetallic Minerals and Solid Wastes, National Laboratory of Mineral Materials, School of Materials Technology, China University of Geosciences, Beijing100083, China
[b] College of Materials & Environmental Engineering, Hangzhou Dianzi University, Hangzhou 310036, China

ARTICLE INFO

Keywords:
Lithium metal battery
Lithium dendrites
CuMOF-ANFs separator

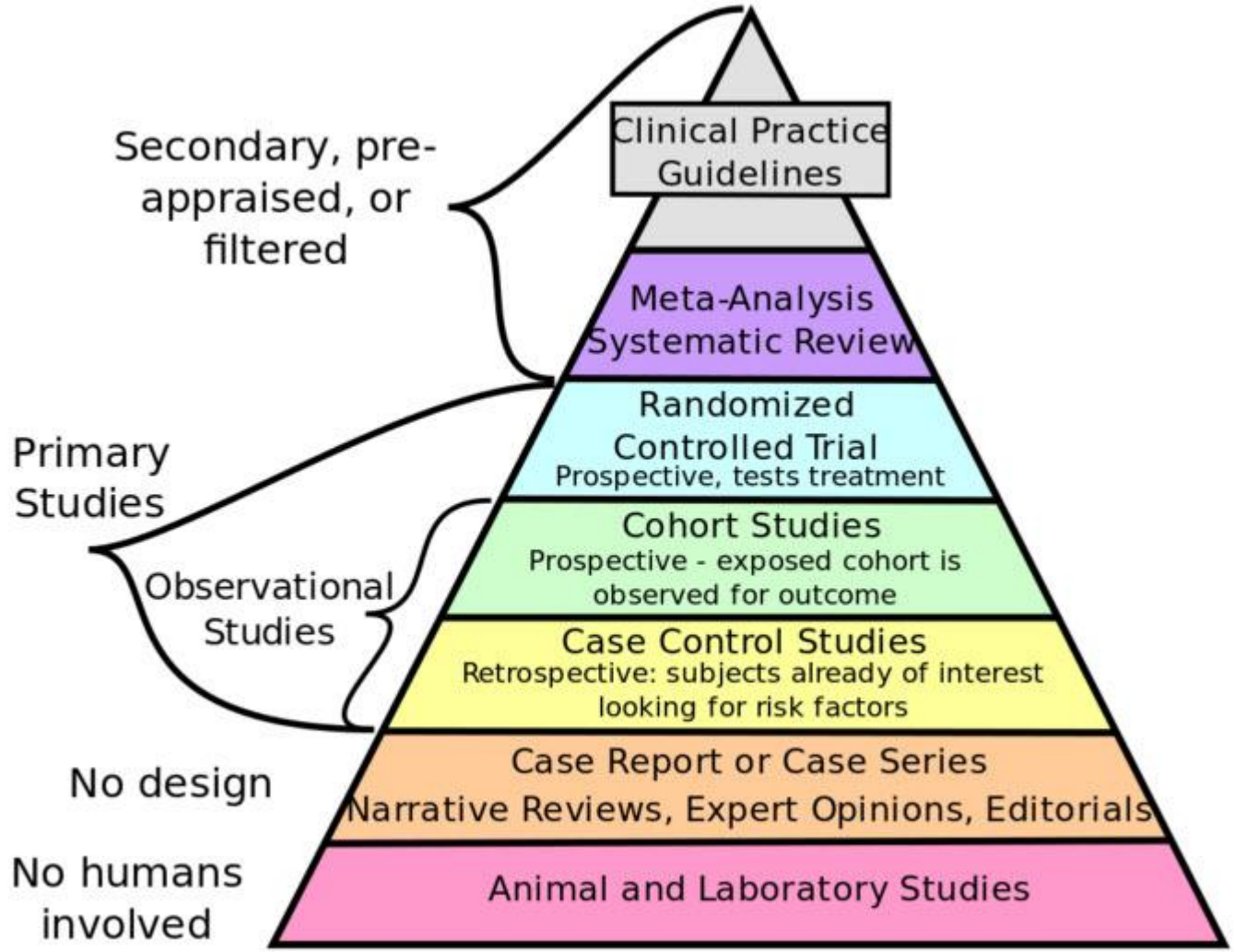
ABSTRACT

Lithium metal, due to its advantages of high theoretical capacity, low density potential, is used as a negative electrode material for batteries and brings grea of energy storage systems. However, the production of lithium metal dendrit poor safety, so lithium dendrites have been the biggest problem of lithium me the larger specific surface area and more pore structure of Cu-based metal-org (CuMOF-ANFs) composite separator can help to inhibit the formation of lithiu mA/cm^2, the discharge capacity retention rate of the Li-Cu battery using the C %. Li-Li batteries can continue to maintain low hysteresis for 2000 h at the show that CuMOF-ANFs composite membrane can inhibit the generation of l cycle stability and cycle life of the battery. The three-dimensional (3D) poro separator provides a new perspective for the practical application of lithium

1. Introduction

Certainly, here is a possible introduction for your topic:Lithium-metal batteries are promising candidates for high-energy-density rechargeable batteries due to their low electrode potentials and high theoretical capacities [1,2]. However, during the cycle, dendrites forming on the lithium metal anode can cause a short circuit, which can

chemical stability of the separator is equa the separator remains intact and does n ence of the electrolyte or other battery c separator helps to prevent the formatio further promote dendrite growth. Rese different materials and designs for sep chanical strength and chemical stabilit



BRITISH MEDICAL JOURNAL

LONDON SATURDAY OCTOBER 30 1948

STREPTOMYCIN TREATMENT OF PULMONARY TUBERCULOSIS

A MEDICAL RESEARCH COUNCIL INVESTIGATION

Results at End of Six Months

Four of the 55 S patients (7%) and 14 of the 52 C patients (27%) died before the end of six months. The difference between the two series is statistically significant; the probability of it occurring by chance is less than one in a hundred.

Taxonomy of clinical trials

- Phase I-IV: pharmacological interventions
- IDEAL stage I-IV: medical devices and invasive procedures (surgery, endoscopic/radiological interventions)
- Number of groups
  - Single arm versus
  - >1 arm
    - Parallel groups
    - Crossover design
    - Factorial design
- Fixed versus adaptive design
- Allocation mechanism
  - Random assignment
  - Non random
- Blinding (masking)
  - Open label
  - Single blinded
  - Double blinded

GHENT UNIVERSITY

CLINICAL TRIALS AND OBSERVATIONS | NOVEMBER 25, 2021

A phase 3, open-label, randomized study of asciminib, a STAMP inhibitor, vs bosutinib in CML after 2 or more prior TKIs

Clinical Trials & Observations

Delphine Réa, Michael J. Mauro, Carla Boquimpani, Yosuke Minami, Elza Lomaia, Sergey Voloshin, Anna Turkina, Dong-Wook Kim, Jane F. Apperley, Andre Abdo, Laura Maria Fogliatto, Dennis Dong Hwan Kim, Philipp le Coutre, Susanne Saussele, Mario Annunziata, Timothy P. Hughes, Naeem Chaudhri, Koji Sasaki, Lynette Chee, Valentin García-Gutiérrez, Jorge E. Cortes, Paola Aimone, Alex Allepuz, Sara Quenet, Véronique Bédoucha, Andreas Hochhaus

Blood (2021) 138 (21): 2031–2041.

https://doi.org/10.1182/blood.2020009984

Article history

Ethical aspects of human experimentation

- Basis: Helsinki declaration, ICH GCP
- EC approval and written informed consent (when possible)
- Unethical or questionable designs:
  - Addressing questions already answered
  - Lacking full informed consent (e.g. Zelen design)
  - Placebo controlled surgical interventions
- When patient is *recognizable*: written informed consent after being shown the intended publication and being informed about dissemination channels

Special Communication

World Medical Association Declaration of Helsinki
Ethical Principles for Medical Research Involving Human Participants

World Medical Association

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

75th WMA General Assembly, Helsinki, Finland, October 2024

GHENT UNIVERSITY

Why register a study protocol?

→ To prevent HARKing (Hypothesizing After the Results are Known): presentation of a post hoc hypothesis as an a priori hypothesis

– 40–62% of publications had at least one primary outcome changed, newly introduced or omitted compared to protocol [Dwan et al, PLoS ONE 2008]

NIH U.S. National Library of Medicine
ClinicalTrials.gov

EudraCT

Welcome to PROSPERO
International prospective register of systematic reviews

GHENT UNIVERSITY

Interpretation of a P value

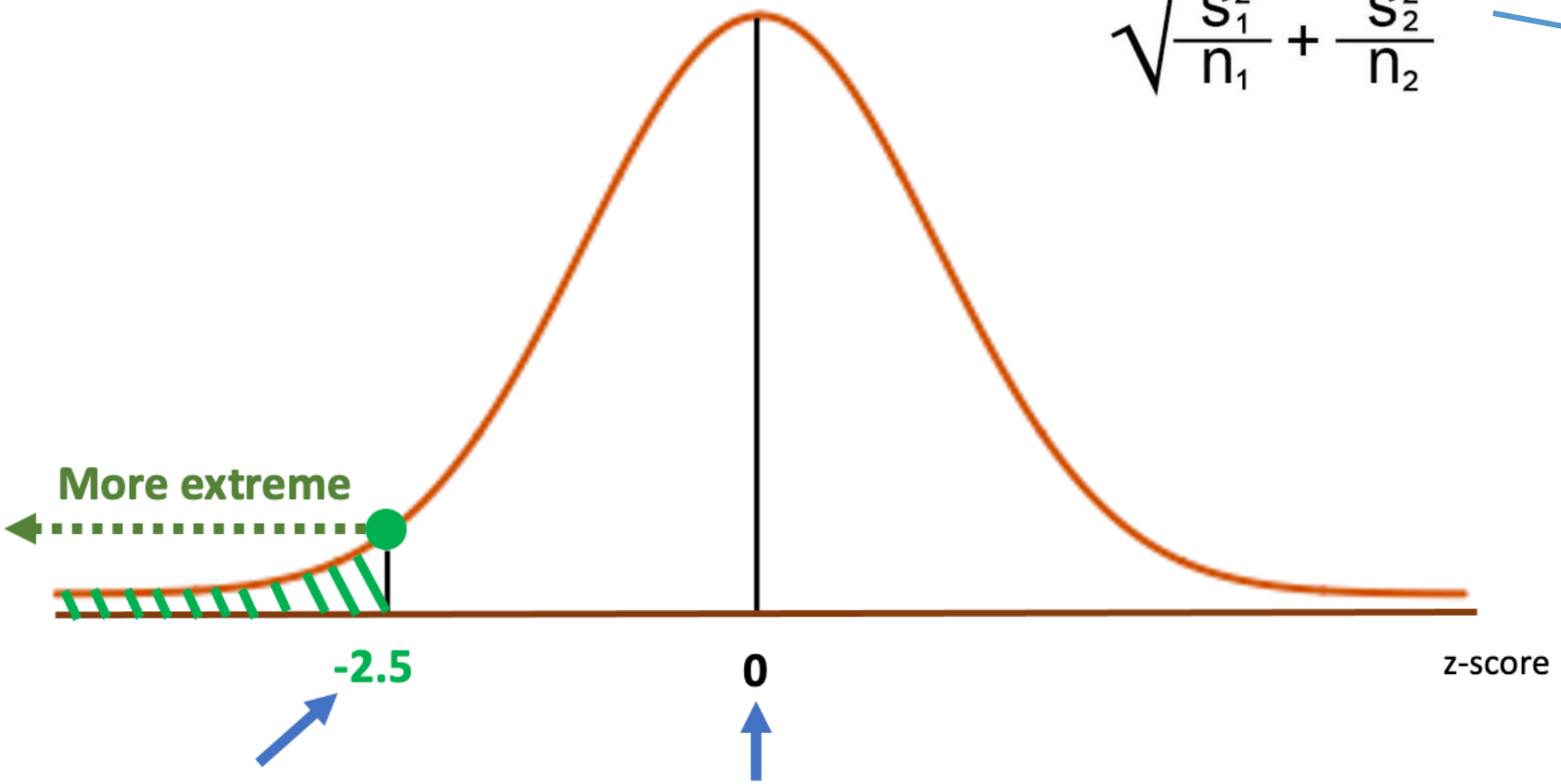
- A p value is the (conditional) probability to find a certain data distribution, given a certain hypothesis is true (usually: H_0 or hypothesis of a null effect)
- A p value is **NOT** the probability of a 'chance finding' (false positive)
- A p value does **NOT** inform about the size , importance, or direction of an effect → confidence intervals should be added
- P(D|H) ≠ P(H|D)! (*inverted conditional or prosecutor's fallacy)*
- Exact p values should be mentioned (and not p < 0.05 or p = NS)

The p value is a ‘surprise index’

$$t = \frac{(\bar{x}_1 - \bar{x}_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

Effect size

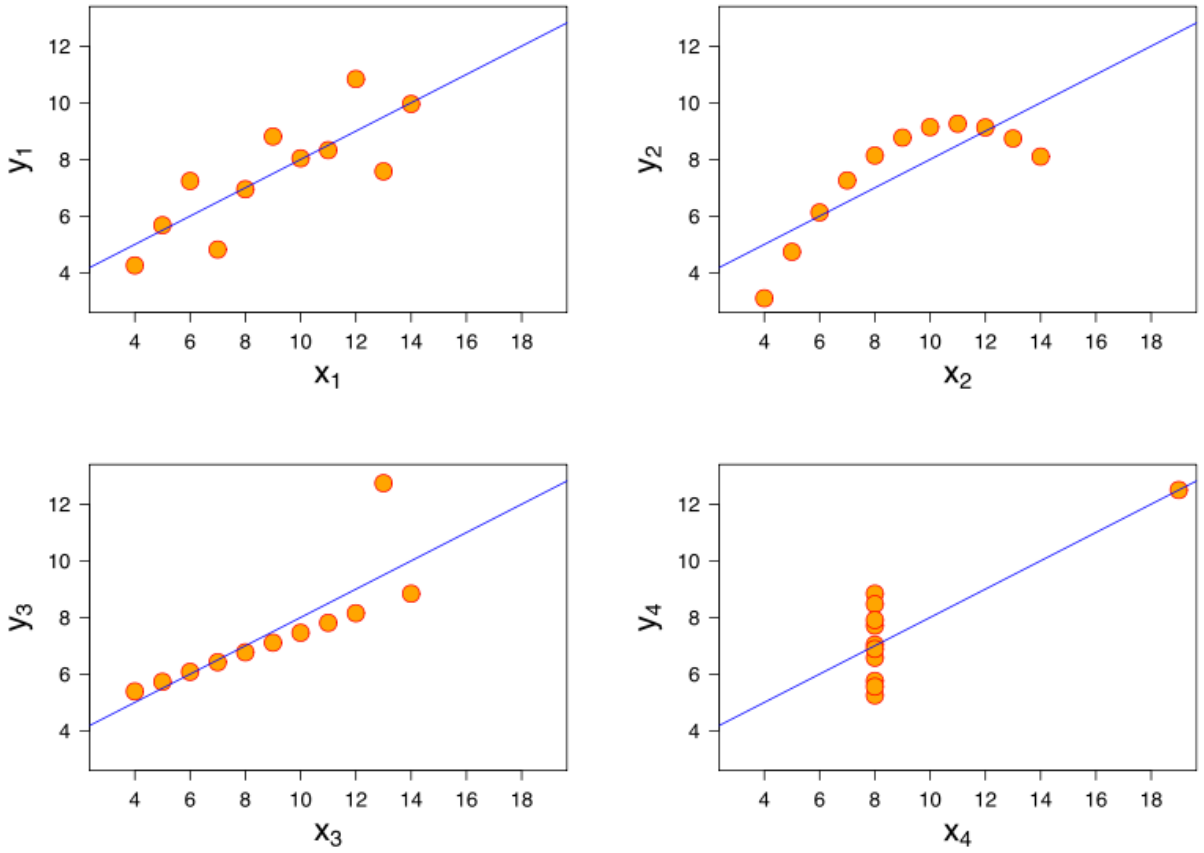
Uncertainty
Variability

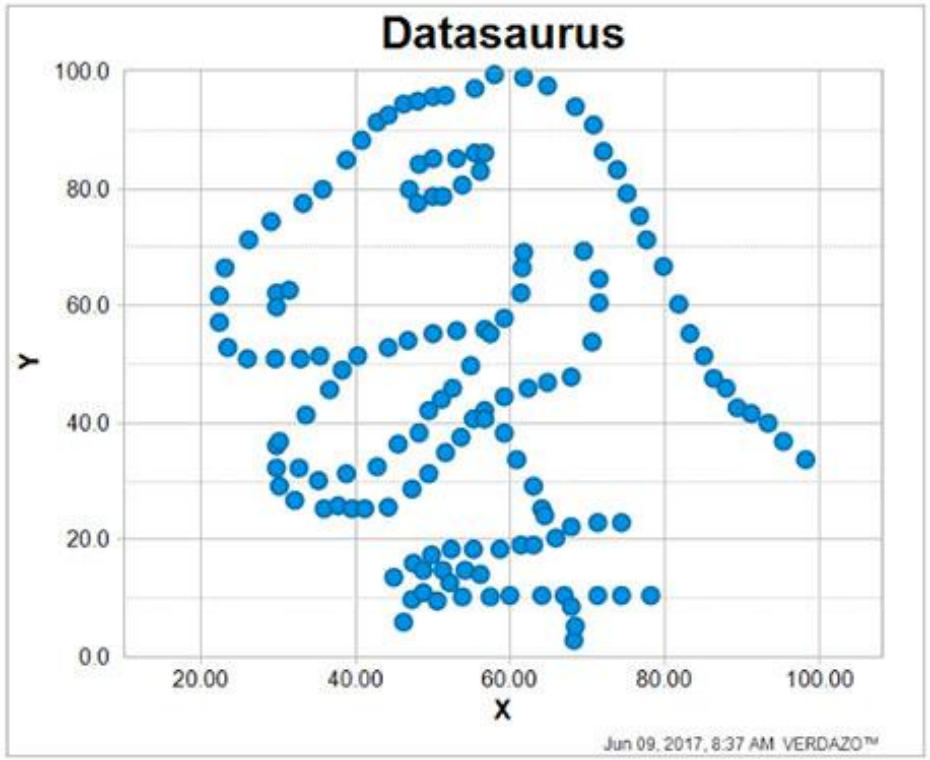


Use of statistical tests and models

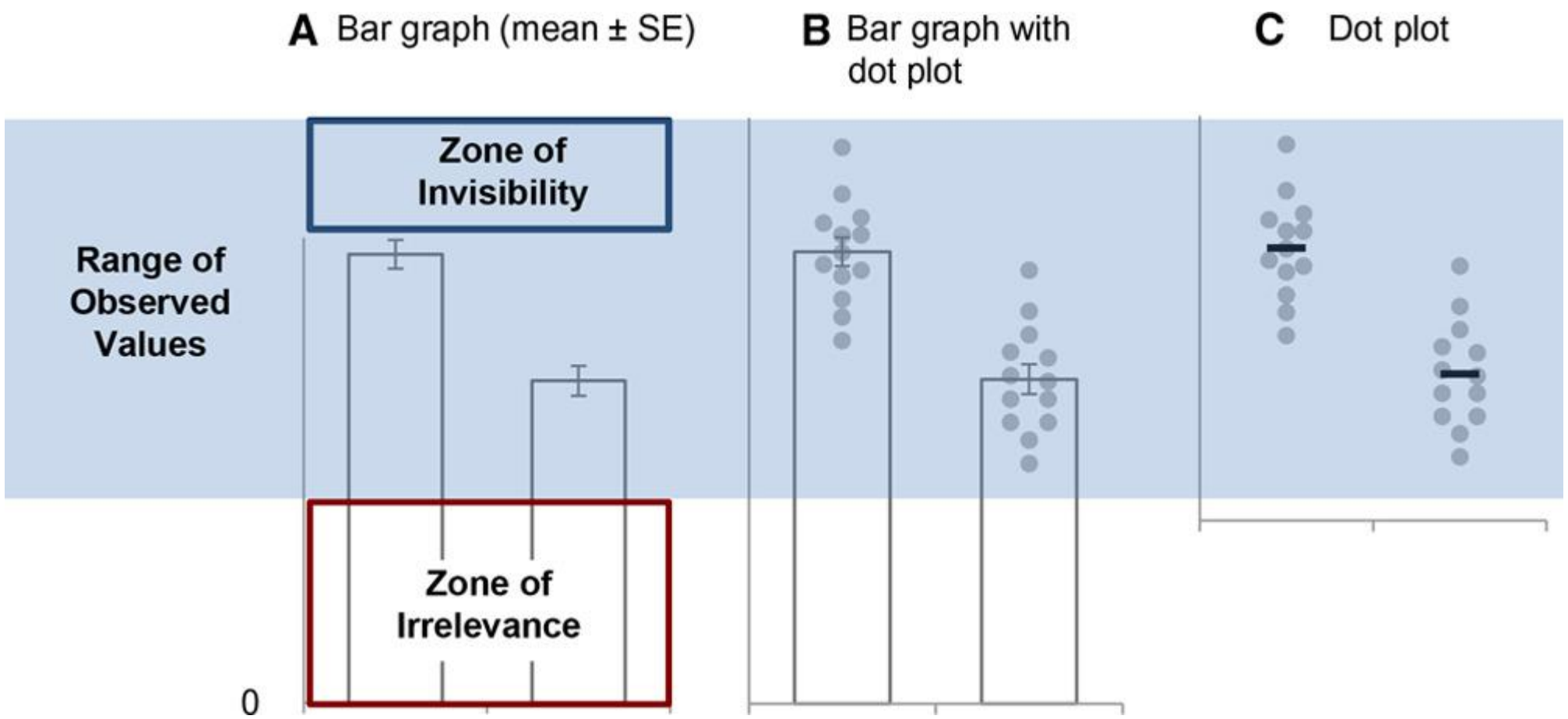
- Parametric tests: Gaussian distribution? (nb: CLT)
- Use of SE instead of SD: not a good measure of dispersion
- Data visualization: e.g. use data points, not bar charts
- Correlations tests (Pearson, Spearman): MUST show scatter plots
- How were missing data handled?
- Observational studies: baseline comparisons should not have p values calculated!
- Survival analysis: should be handled as time to event variable, not as binary (alive/dead)

Correlation tests: Importance of visualizing data!





$$\rho_{X,Y} = \frac{\operatorname{cov}(X,Y)}{\sigma_X \sigma_Y}$$



Weissgerber Circulation 2019

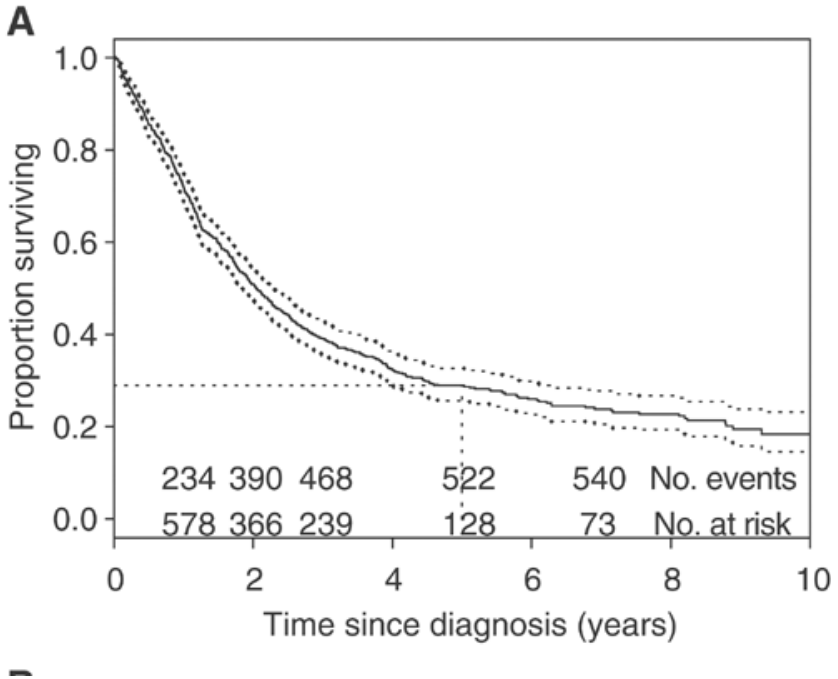
How to assess time to event curves (Kaplan Meier)

- Should state numbers at risk and (ideally) confidence intervals
- Cave: crossing survival curves
- Cave: informative censoring?

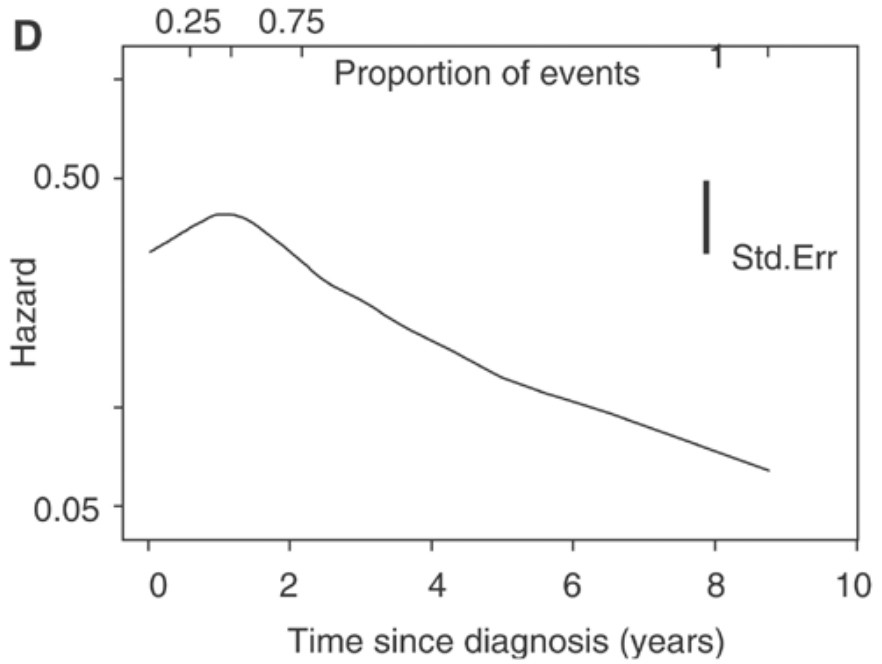
GHENT UNIVERSITY

What is a *hazard*?

Survival function

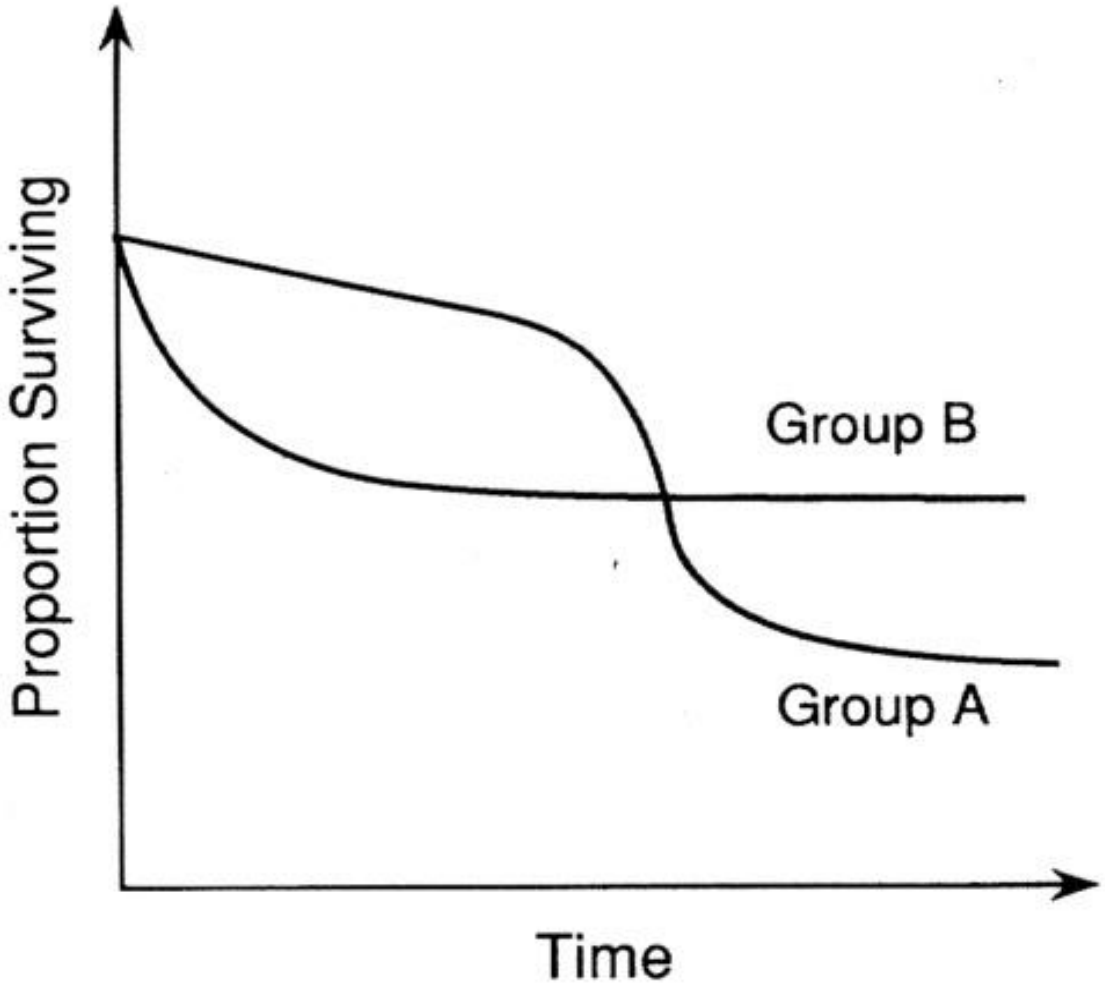


Hazard function



$$h(t) = -\frac{d}{dt}\ln\left(S(t)\right)$$

$$h(t) = \lim_{\Delta t \to 0} \frac{P(t \le T < t + \Delta t \mid T \ge t)}{\Delta t}$$



Non proportional Hazards

Appraisal of a randomized trial

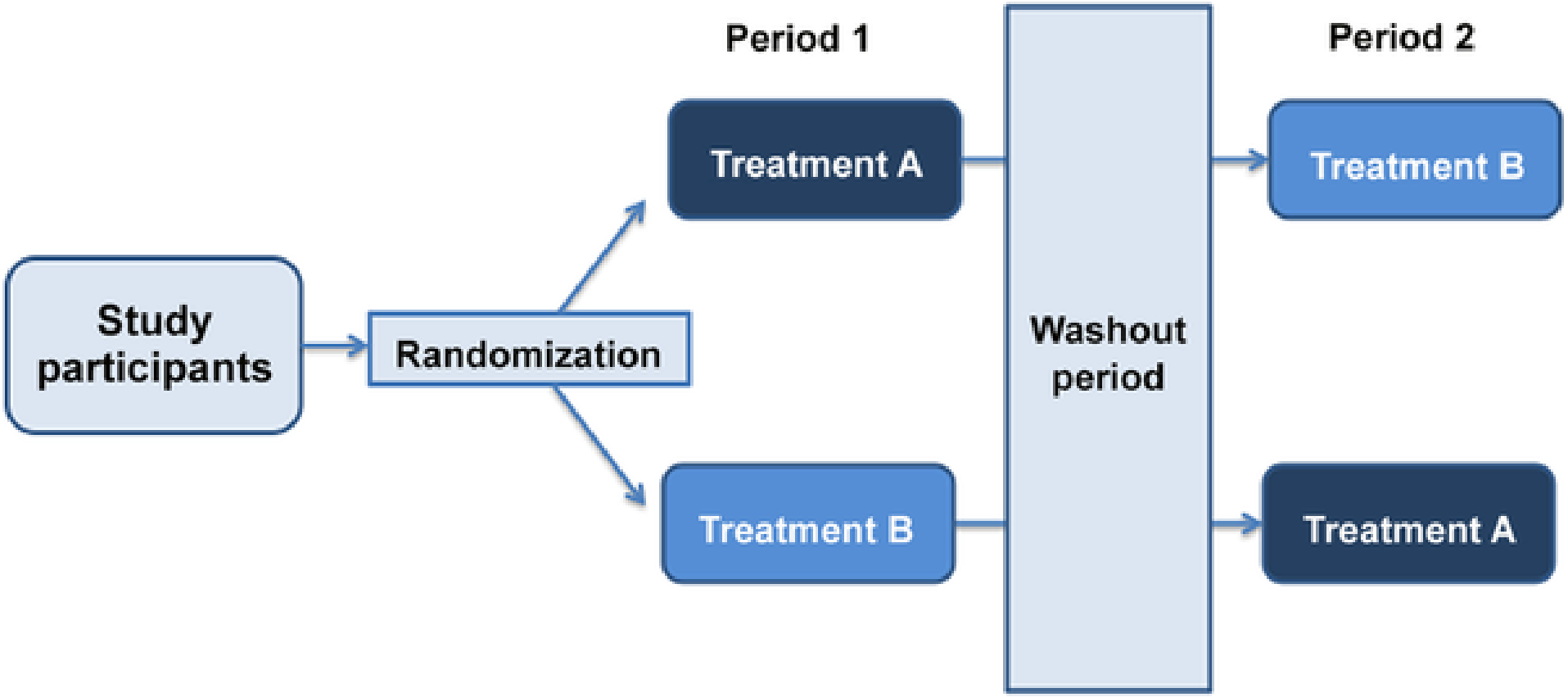
- Guidelines: CONSORT
  - Extensions: non pharmacological interventions; pragmatic trials

GHENT UNIVERSITY

GHENT UNIVERSITY

https://www.equator-network.org/

Crossover design



GHENT UNIVERSITY

Crossover design

- **Advantages**
  - Eliminates between-patient variability
  - Fewer patients needed for same number of observations
  - Fewer observations needed for same precision
  - All patients receive active treatment and may choose preferred treatment at the end

- **Disadvantages**
  - Drop-outs more problematic
  - Period by treatment interaction (e.g. carry-over) → only in stable conditions, e.g. diabetes
  - Several treatment periods may be inconvenient to patients
  - Difficult to analyze (mixed models)

Factorial design

- Tests >1 research question at once
- More efficient than multi-arm trial (= lower sample size for similar precision)
- Relies on assumption of no interactions → usually not realistic

		Carboplatin	
		No	Yes
Bevacizumab	No	Arm C: standard neoadjuvant chemotherapy*	Arm B: carboplatin + standard neoadjuvant chemotherapy
	Yes	Arm A: bevacizumab + standard neoadjuvant chemotherapy	Arm AB: bevacizumab + carboplatin + standard neoadjuvant chemotherapy

Outcomes (Endpoints)

- Primary
  - planned outcome that is most directly related to the primary objective of the trial
  - typically the outcome used in the sample size calculation
  - Usually one primary outcome, sometimes >1
- Secondary
  - Multiplicity → exploratory only

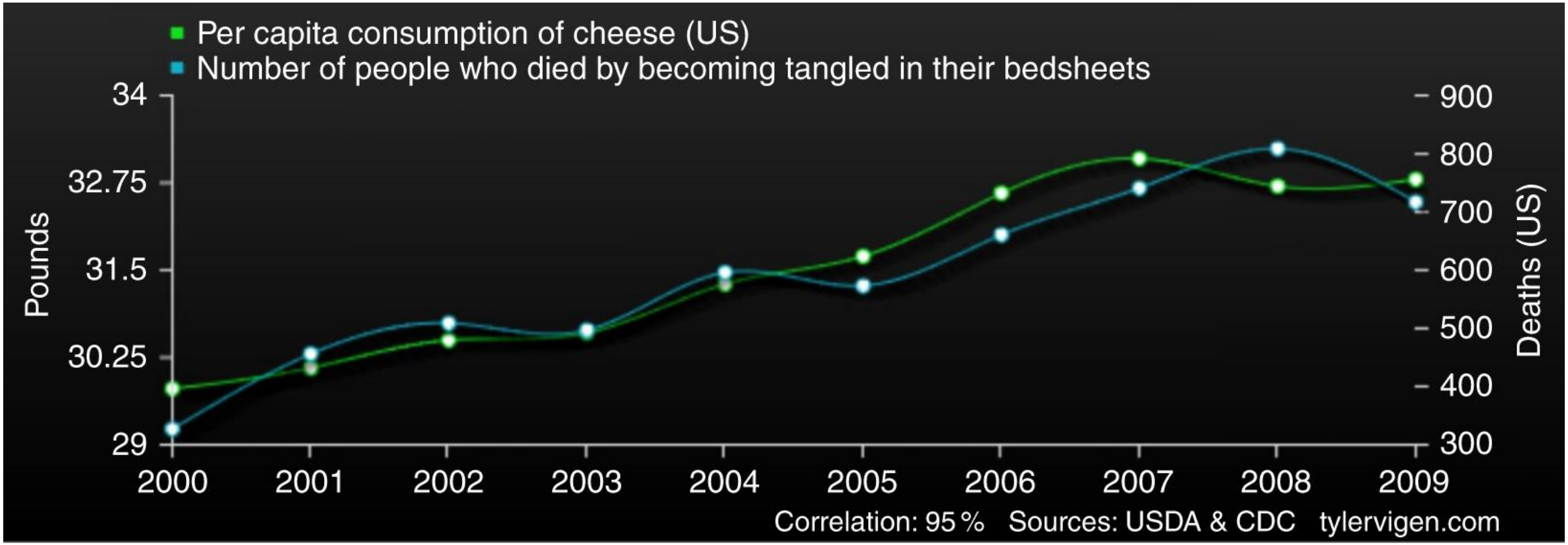
GHENT UNIVERSITY

What is a good primary endpoint?

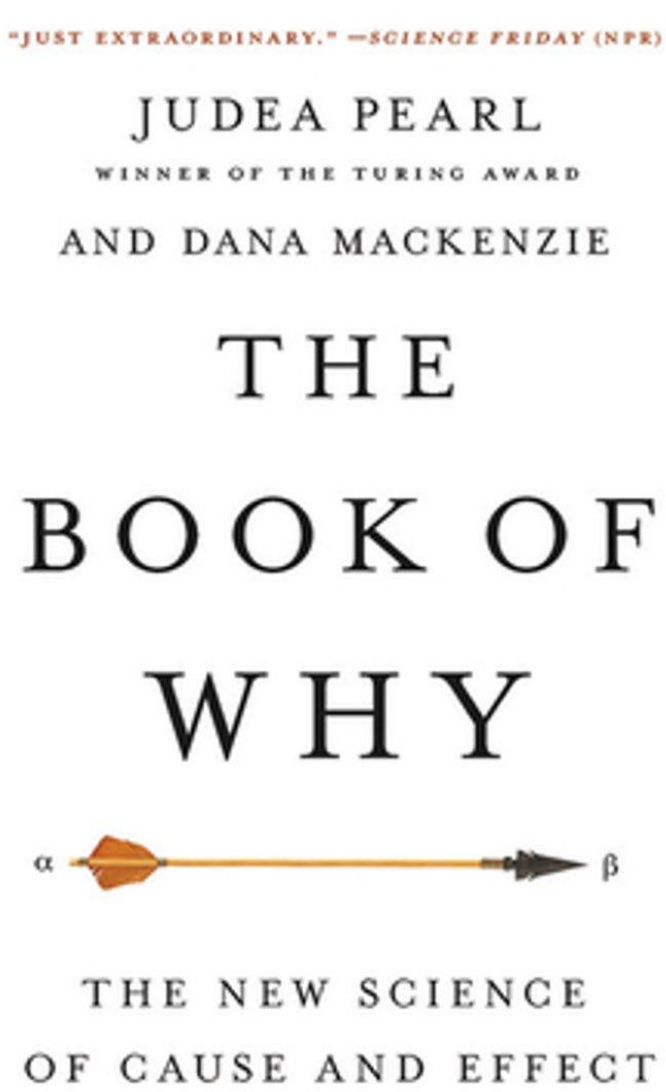
- Unique
  - Defined *a piori*
  - Multiple endpoints: more false positive results
  - RCT: sample size and power calculation based on SINGLE (**primary**) endpoint
- Clinically relevant
- Reliable and reproducible
- If surrogate endpoint: demonstrated validity?
- Available for all patients

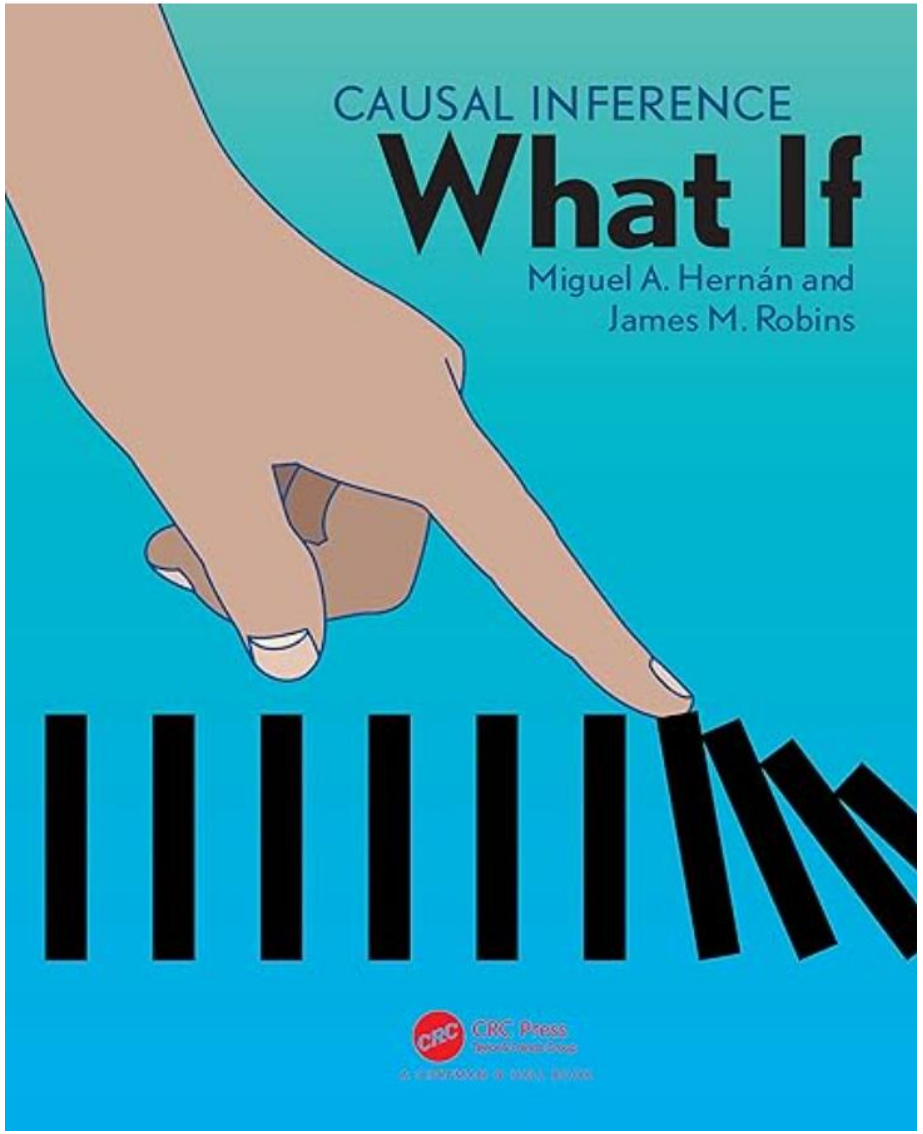
GHENT UNIVERSITY

Causal inference in RCTs



GHENT UNIVERSITY


"JUST EXTRAORDINARY." —SCIENCE FRIDAY (NPR)
JUDEA PEARL
WINNER OF THE TURING AWARD
AND DANA MACKENZIE
THE
BOOK OF
WHY
α
β
THE NEW SCIENCE
OF CAUSE AND EFFECT


CAUSAL INFERENCE
What If
Miguel A. Hernán and
James M. Robins
CRC Press

Types of outcomes

- Hard
  - Mortality
  - Quality of Life
  - Amputations, hearing loss, loss of vision
  - Pain reduction/increase
- Surrogate or intermediate
  - DFS, PFS, pCR as surrogate for OS
  - LN harvest or rectal amputation rate as surrogate for surgical quality in colorectal surgery
- Composite
  - ‘Overall complication rate’
  - MACE (major adverse cardiac events)
- Patient reported outcomes

GHENT UNIVERSITY

Surrogate Outcomes

- Valid:
  - the marker is **intermediate** on the causal pathway between treatment and hard outcome AND the association between treatment and surrogate endpoint is consistent
  - The association always has the same extent and sign as that between the treatment and the hard endpoint
- Invalid:
  - The surrogate marker is associated with the exposure, but there is **no causal association** between the surrogate marker and the hard endpoint

GHENT UNIVERSITY

Surrogate endpoints: examples

- Oncology trials: DFS, PFS, pCR as surrogate for OS
- Orthopedic trials: imaging data
- LN harvest or amputation rate as surrogate for surgical quality in colorectal surgery
- Prognostic indicators are not always surrogate endpoints!

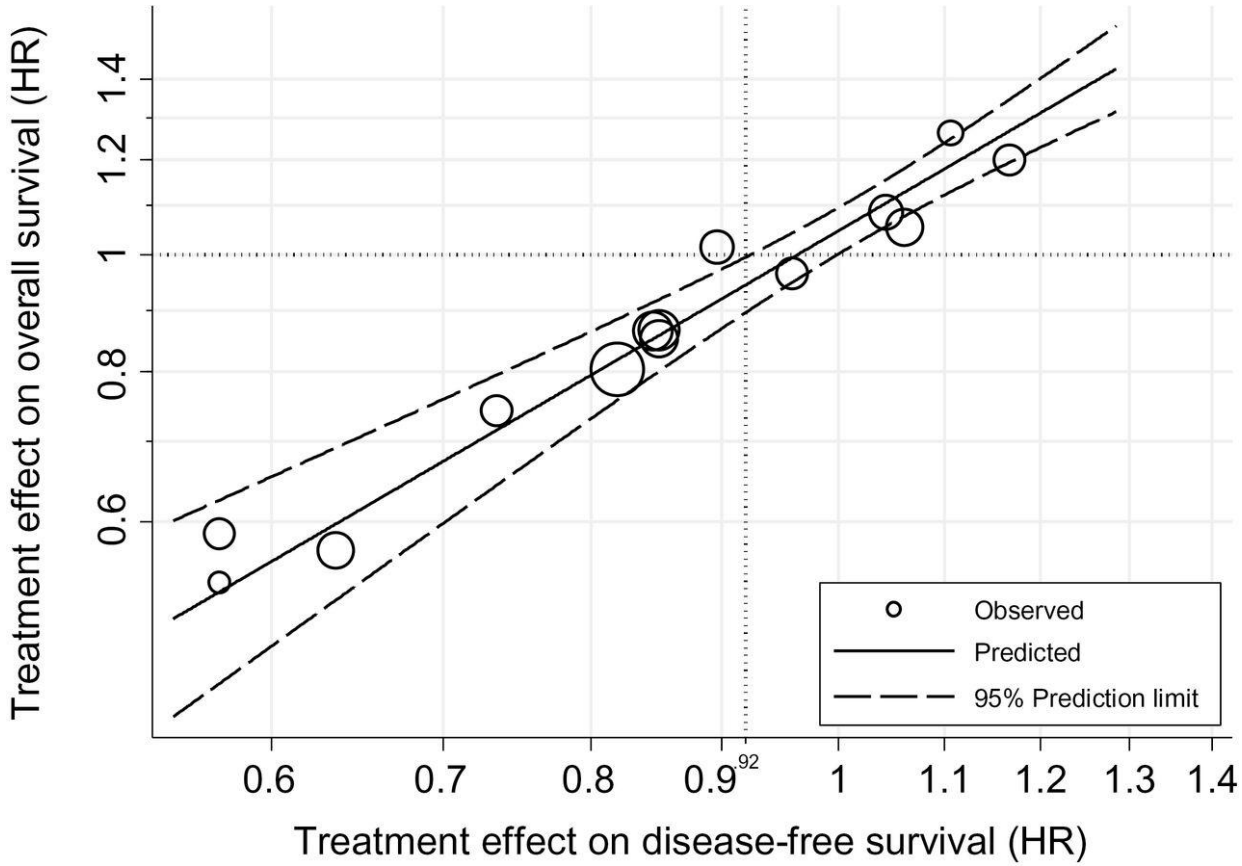
GHENT UNIVERSITY

TABLE 1. Reasons Why PFS Is an Inappropriate Primary End Point in Most Trials Evaluating Anticancer Drugs

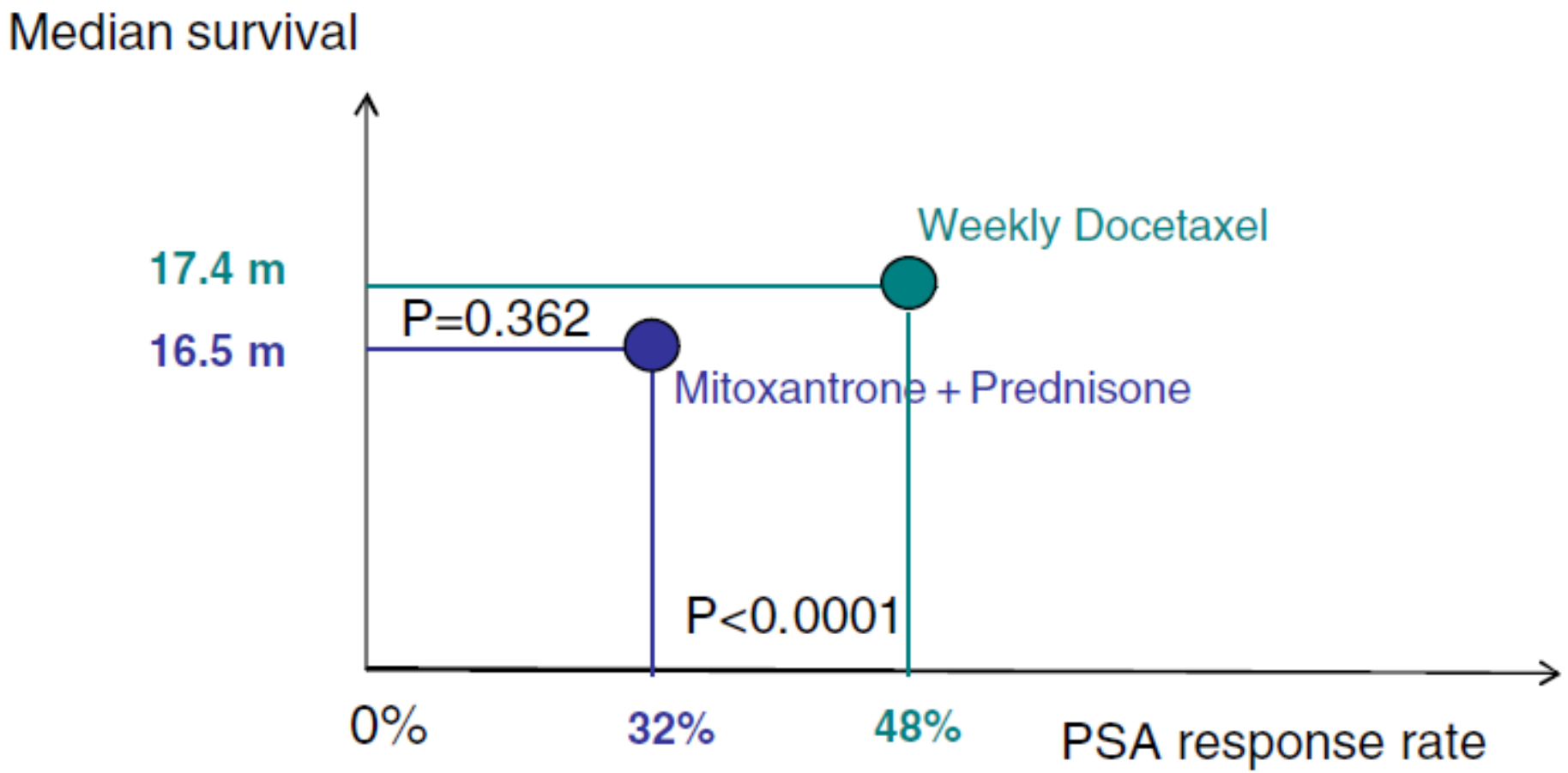
Improvement in PFS is seldom a surrogate for, nor reliably predictive of, improvement in OS
Improvement in PFS is not a surrogate for, nor predictive of, improvement in QoL
PFS does not recognize that the balance between benefit and harm depends not only on changes in tumor size but also on toxicity
PFS measurement and comparisons are subject to error and bias because of
Timing of assessment
Measurement error in assessing tumor progression
Informative censoring because of uneven dropout between groups in an RCT
Improvement in PFS is widely misunderstood by patients and the public to imply improvement in survival

Abbreviations: OS, overall survival; PFS, progression-free survival; QoL, quality of life; RCT, randomized controlled trial.

Booth JCO 2023



Oba Disease-free survival as a surrogate for overall survival in adjuvant trials of gastric cancer: a meta-analysis. JNCI 2013



Collette Eur J Cancer 2006

Patient Group Engagement Across the Clinical Trial Continuum*

Patient groups have potential to enhance the quality and efficiency of clinical trials by providing:

Discovery & Pre-Clinical†

- Financial support for research
- Natural history data
- Input on relevance of research to patients
- Access to translational tools
- Help defining eligibility criteria
- Input on meaningful endpoints & PROs
- Advocacy for policy & funding issues**
- Education to patient community**

Phase 1 - 3

- Benefit-risk & patient-preference studies
- Protocol design & study feasibility input
- Study recruitment & retention strategy input
- Increased awareness about trials
- Participant feedback on trial experience
- Input on informed consent content & processes
- Peer advocates for participants**
- Clinical trial networks**
- Data Safety Monitoring Board members**

Regulatory Review

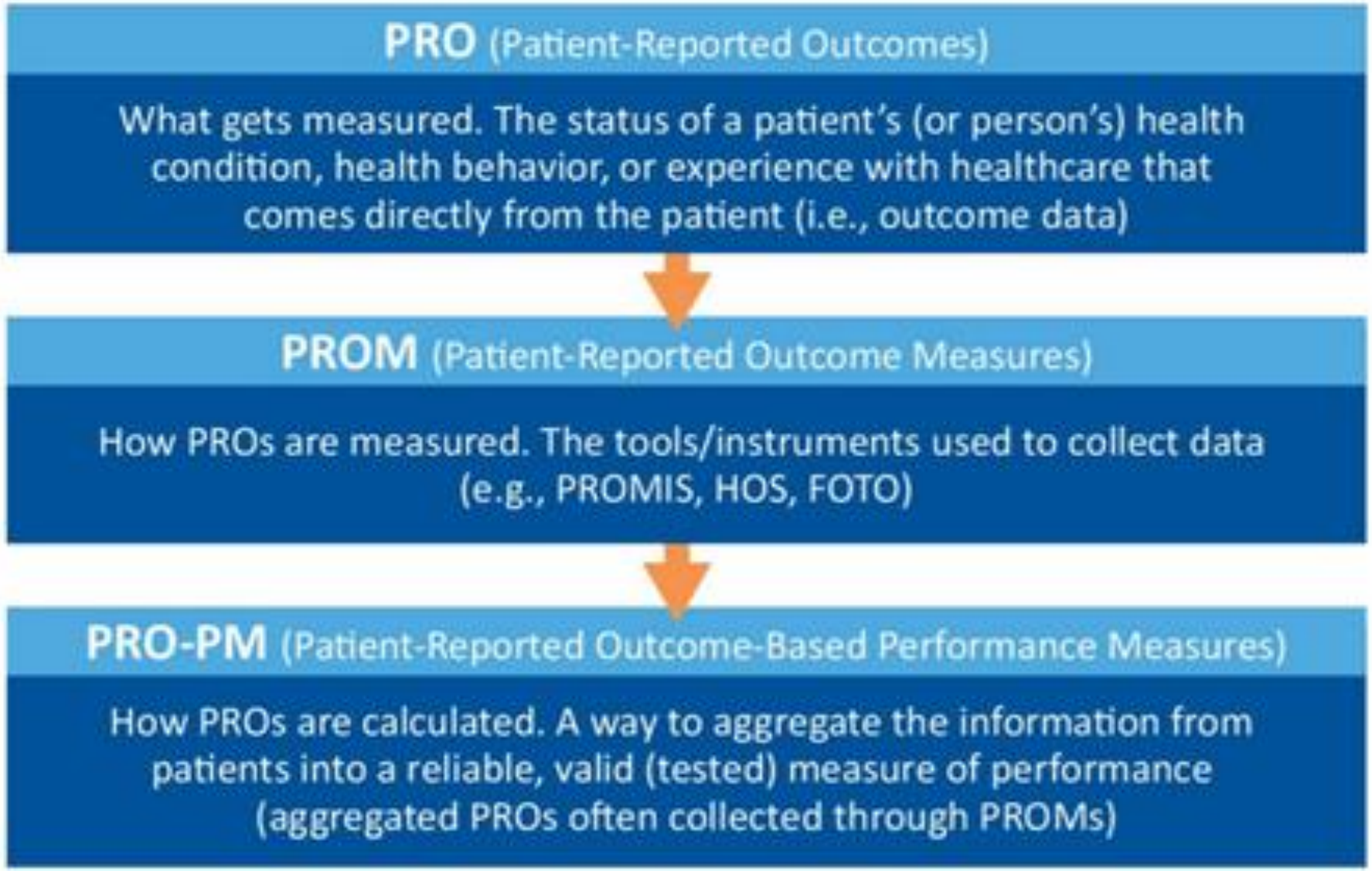
- Support to sponsors around key regulatory meetings
- Support preparing submissions for newborn screening for rare diseases
- Informing regulators on benefit-risk**
- Public testimony at regulatory meetings**

Post-Approval

- *Phase 1-3 activities and…*
- Support interpreting & disseminating study results
- Collaboration on post-marketing studies & surveillance initiatives
- Support developing access strategy & preparing for value or health technology review

**Updated 2018; adapted from Parkinson's Foundation materials*
***Patient group activities typically undertaken independently or with partners other than sponsors*
†Includes early planning for trials

CTTI


PRO (Patient-Reported Outcomes)
What gets measured. The status of a patient's (or person's) health condition, health behavior, or experience with healthcare that comes directly from the patient (i.e., outcome data)
PROM (Patient-Reported Outcome Measures)
How PROs are measured. The tools/instruments used to collect data (e.g., PROMIS, HOS, FOTO)
PRO-PM (Patient-Reported Outcome-Based Performance Measures)
How PROs are calculated. A way to aggregate the information from patients into a reliable, valid (tested) measure of performance (aggregated PROs often collected through PROMs)

Core Outcome Measures in Effectiveness Trials

www.comet-initiative.org

Are the results **clinically significant (important)**?

- Large sample size → even small effect magnitude becomes clinically significant
- Examples
  - Tx of hypertension: mean decrease of 2 mm in RR
  - OS in lung cancer: 5 weeks improvement
- Efficacy versus value

GHENT UNIVERSITY

How was the desired effect size chosen?

- Literature review
- Pilot study
- Consultation with stakeholders (patients, funders,…)

Clinical Review & Education

JAMA Guide to Statistics and Methods

Minimal Clinically Important Difference
Defining What Really Matters to Patients

Anna E. McGlothlin, PhD; Roger J. Lewis, MD, PhD

GHENT UNIVERSITY

DELTA^2 guidance on choosing the target difference and undertaking and reporting the sample size calculation for a randomised controlled trial

Jonathan A Cook,[1] Steven A Julious,[2] William Sones,[1] Lisa V Hampson,[3,4] Catherine Hewitt,[5] Jesse A Berlin,[6] Deborah Ashby,[7] Richard Emsley,[8] Dean A Fergusson,[9] Stephen J Walters,[2] Edward C F Wilson,[10] Graeme MacLennan,[11] Nigel Stallard,[12] Joanne C Rothwell,[2] Martin Bland,[5] Louise Brown,[13] Craig R Ramsay,[14] Andrew Cook,[15] David Armstrong,[16] Doug Altman,[1] Luke D Vale[17]

For numbered affiliations see end of article.

Correspondence to: J Cook
jonathan.cook@ndorms.ox.ac.uk

Additional material is published online only. To view please visit the journal online.

Cite this as: *BMJ* 2018;363:k3750
http://dx.doi.org/10.1136/bmj.k3750

Accepted: 9 August 2018

Randomised controlled trials are considered to be the best method to assess comparative clinical efficacy and effectiveness, and can be a key source of data for estimating cost effectiveness. Central to the design of a randomised controlled trial is an a

treatments that is considered realistic or important by one or more key stakeholder groups. The sample size calculation ensures that the trial will have the required statistical power to identify whether a difference of a particular magnitude exists. In this

ESMO clinical benefit scale

If median OS with the standard treatment is ≤12 months

GRADE 4	HR ≤0.65 AND gain ≥3 months	○
	Increase in 2 year survival ≥10%	○
GRADE 3	HR ≤0.65 AND gain ≥2.0-<3 months	○
GRADE 2	HR ≤0.65 AND gain ≥1.5-<2.0	○
	HR >0.65-0.70 AND gain ≥1.5 months	○
GRADE 1	HR >0.70 OR gain <1.5 months	○

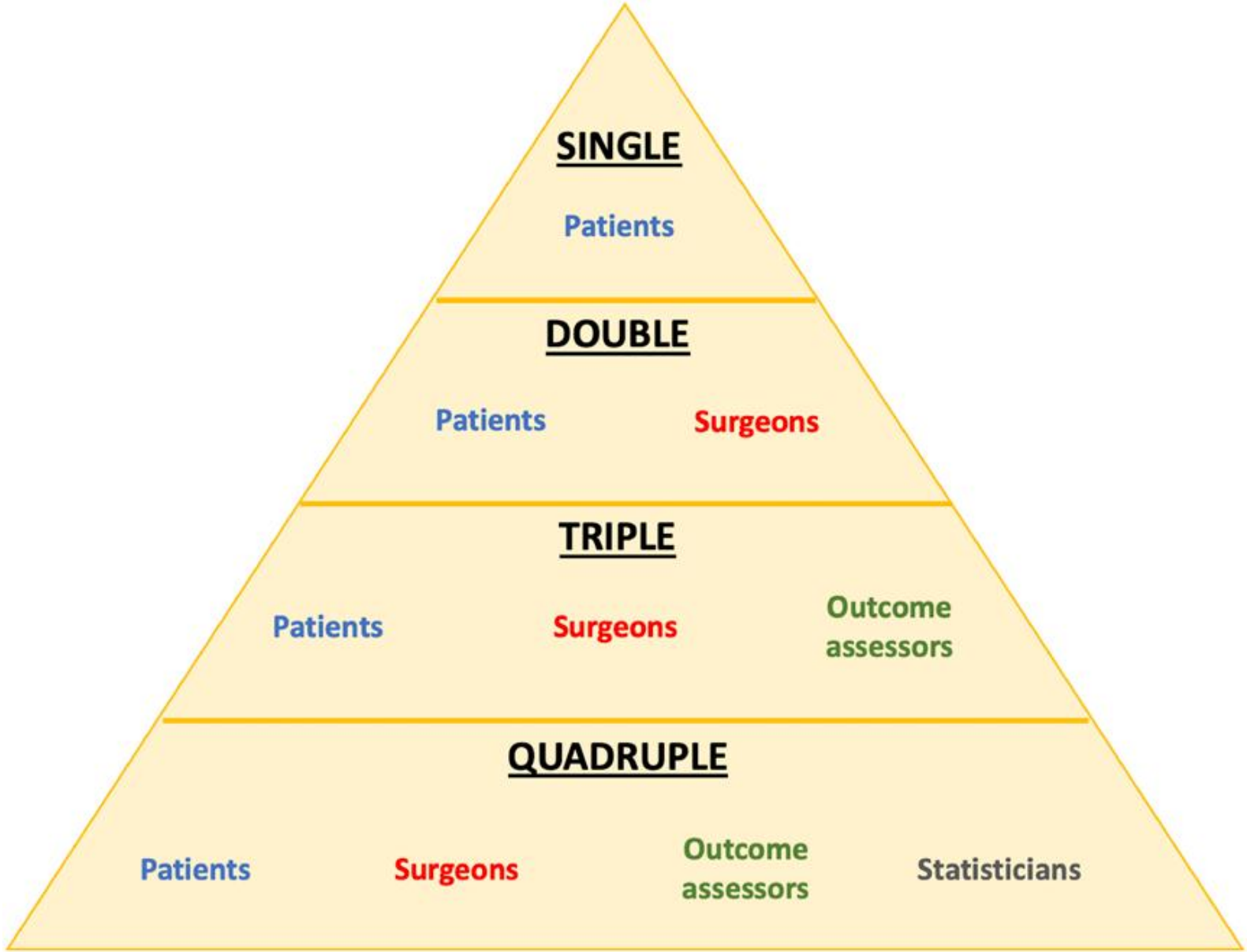
Mark with √ if relevant

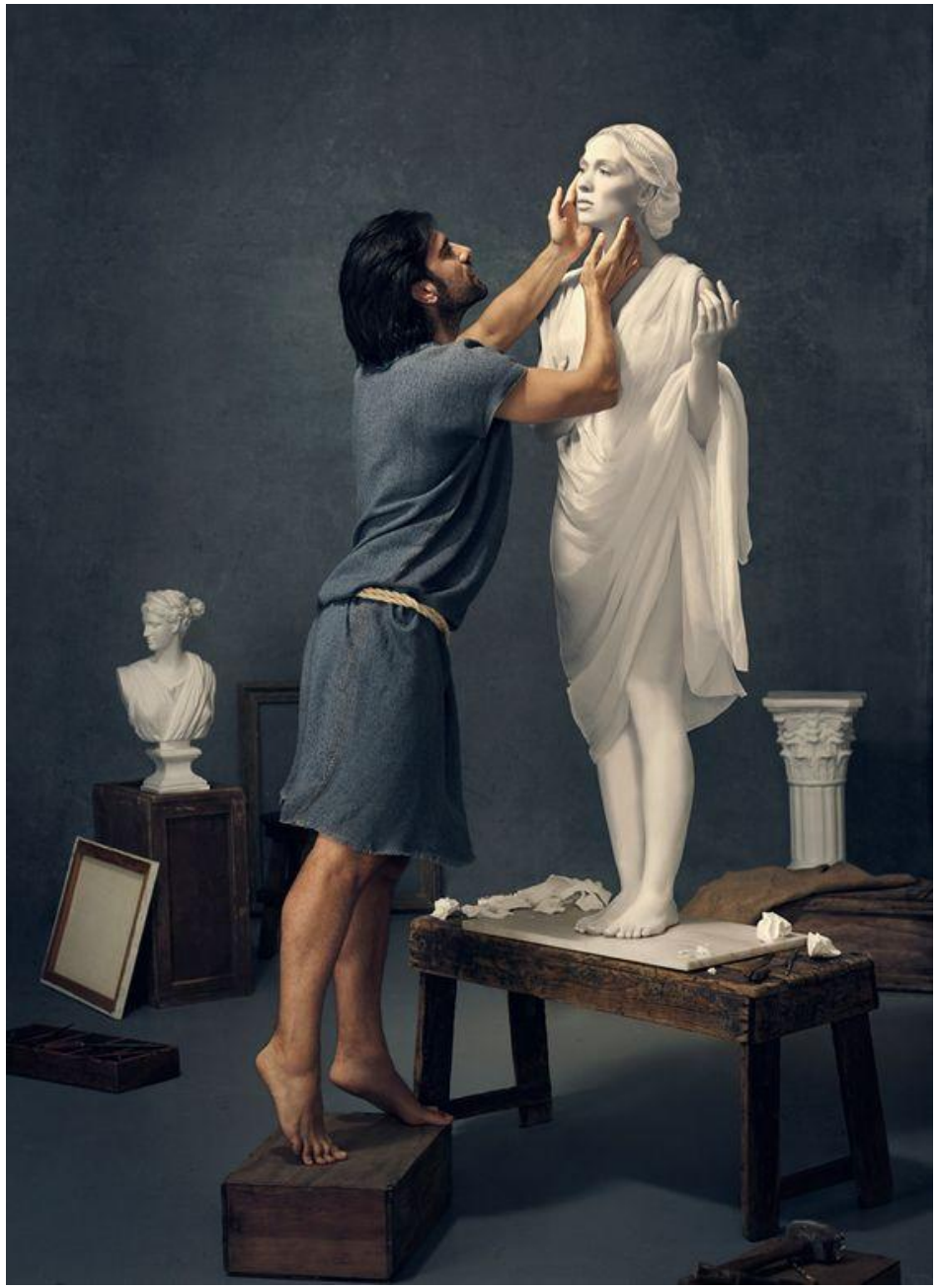
GHENT UNIVERSITY

Blinding (masking)

- Single, double, triple
- Aim: avoid bias
  - Participants: can alter expectations, assessment of efficacy, treatment seeking behavior
  - Trial staff: Differential treatment, attention, or attitudes (Pygmalion effect)

GHENT UNIVERSITY


SINGLE
Patients
DOUBLE
Patients
Surgeons
TRIPLE
Patients
Surgeons
Outcome assessors
QUADRUPLE
Patients
Surgeons
Outcome assessors
Statisticians



Treatment allocation

- Should be blinded
  - Improper: assignment according to date of visit, etc
- **Randomised** allocation
  - eliminates all sources of bias except accidental bias
  - tends to ensure balance among treatments with respect to known and unknown prognostic factors
  - guarantees the distributional assumptions of the test statistics and estimators
  - Ratio: usually 1:1 (most efficient)

Trial participant selection

- Based on strict inclusion/exclusion criteria
- = Convenience sampling
  - Not a random sample
  - Not a representative sample
- → Limited external validity

GHENT UNIVERSITY

Analysis sets of participants

- **Intention-to-treat** (ITT): all randomized patients, according to the randomization outcome
- **Full-analysis set**: ICH E9 – the set as close as possible to the ideal implied by ITT
- **Per-protocol**: subset of full-analysis, compliant with the protocol
- **As-treated**: patients included according to treatment they actually received
- **Safety:** as-treated + minimum dose requirements

GHENT UNIVERSITY

Intention to treat analysis

- Aim: prevent attrition bias
- Analyze patients according to treatment randomized to, regardless of whether treatment was actually received or not
  - Dropout due to toxicity, competing event,…
  - Crossover
  - Lost to FU
  - Withdrawal of consent

GHENT UNIVERSITY

Full-Analysis Set (ICH E9)

- Possible exclusions from ITT
  - eligibility violations
  - failures to take at least one dose of trial medication
  - the lack of post-randomisation data
- Should always be justified
- Potential bias due to exclusions has to be addressed using sensitivity analysis

GHENT UNIVERSITY

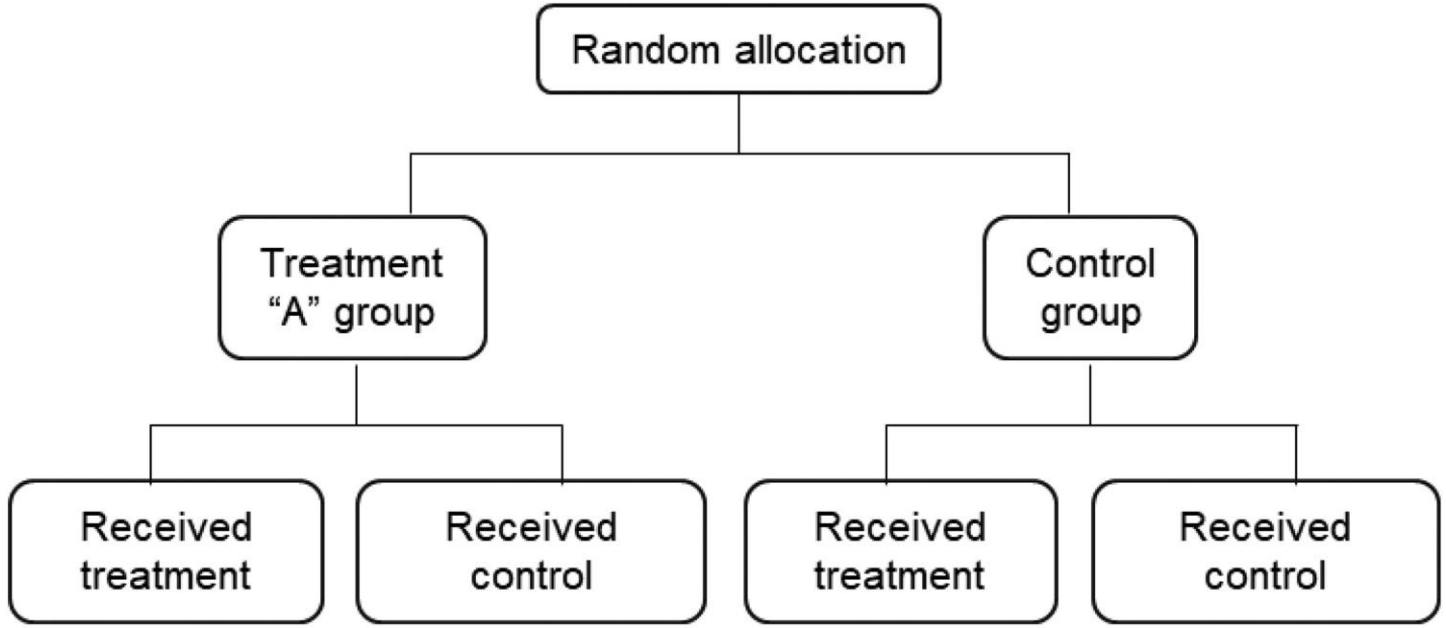
Per-Protocol set

- A subset of the Full-Analysis Set
- Subjects compliant with the protocol, e.g.:
  - Completion of a certain prespecified minimal exposure to the treatment regimen;
  - Availability of measurements of the primary variable(s);
  - Absence of any major protocol violations.
- May be severely biased if adherence to the study protocol related to treatment and outcome

As treated set

- Subjects included according to the treatment *actually received* (generally, at least one dose of drug taken)
- Patients who do not take the drug are almost certainly not a random sample of all patients
- Should be considered mainly for safety analyses

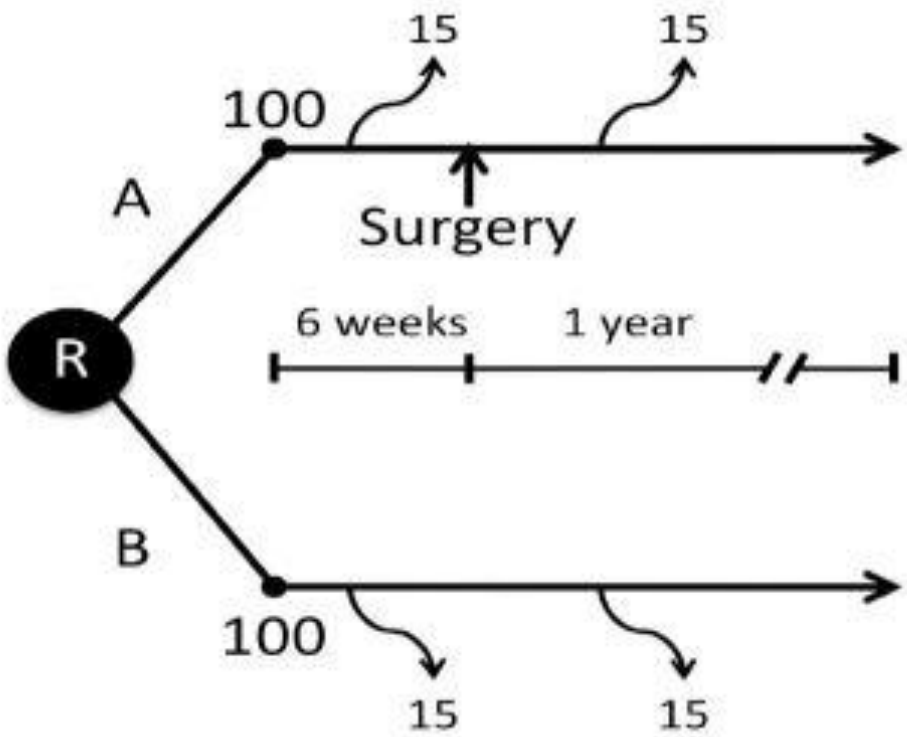
GHENT UNIVERSITY



GROUP INCLUDED IN ANALYSIS				
Intention-to-treat	Treatment group	Treatment group	Control group	Control group
Per protocol	Treatment group	Ignore	Ignore	Control group
As treated	Treatment group	Control group	Treatment group	Control group

GHENT UNIVERSITY

	Analytical Approach		
	Intention to Treat	**Per Protocol**	**As Treated**
Basic principle of comparing participants	As randomized, ignoring actual treatment	As randomized, conditional to protocol compliance	As per treatment actually received, ignoring randomization
Scope	Effectiveness of treatment offer	Efficacy of treatment under ideal circumstances (compliance)	Efficacy of treatment; safety
Properties	Typically underestimating superiority effect	Ideal effect estimated in superiority assessment (anti-conservative in superiority assessment)	Estimated effect subject to self-selection bias, typically anti-conservative in superiority assessment; unbiased effect estimation may require conditioning (e.g., adjustment)
Strengths	Randomization-protected from bias due to imbalance of baseline characteristics; simple	Proof of therapeutic concept	Data set comparable to that of safety analysis and to observational studies; allows for analysis under high treatment crossover rates
Limitations	Imputation of missing data required; generalizability depending on correspondence of in-study and real-life compliance	Reduced power depending on non-compliance; risk of ignorability assumption being violated (selection bias); bias possible in any direction, most likely anti-conservative	Reduced power depending on dropout-rate; bias possible in any direction, most likely anti-conservative



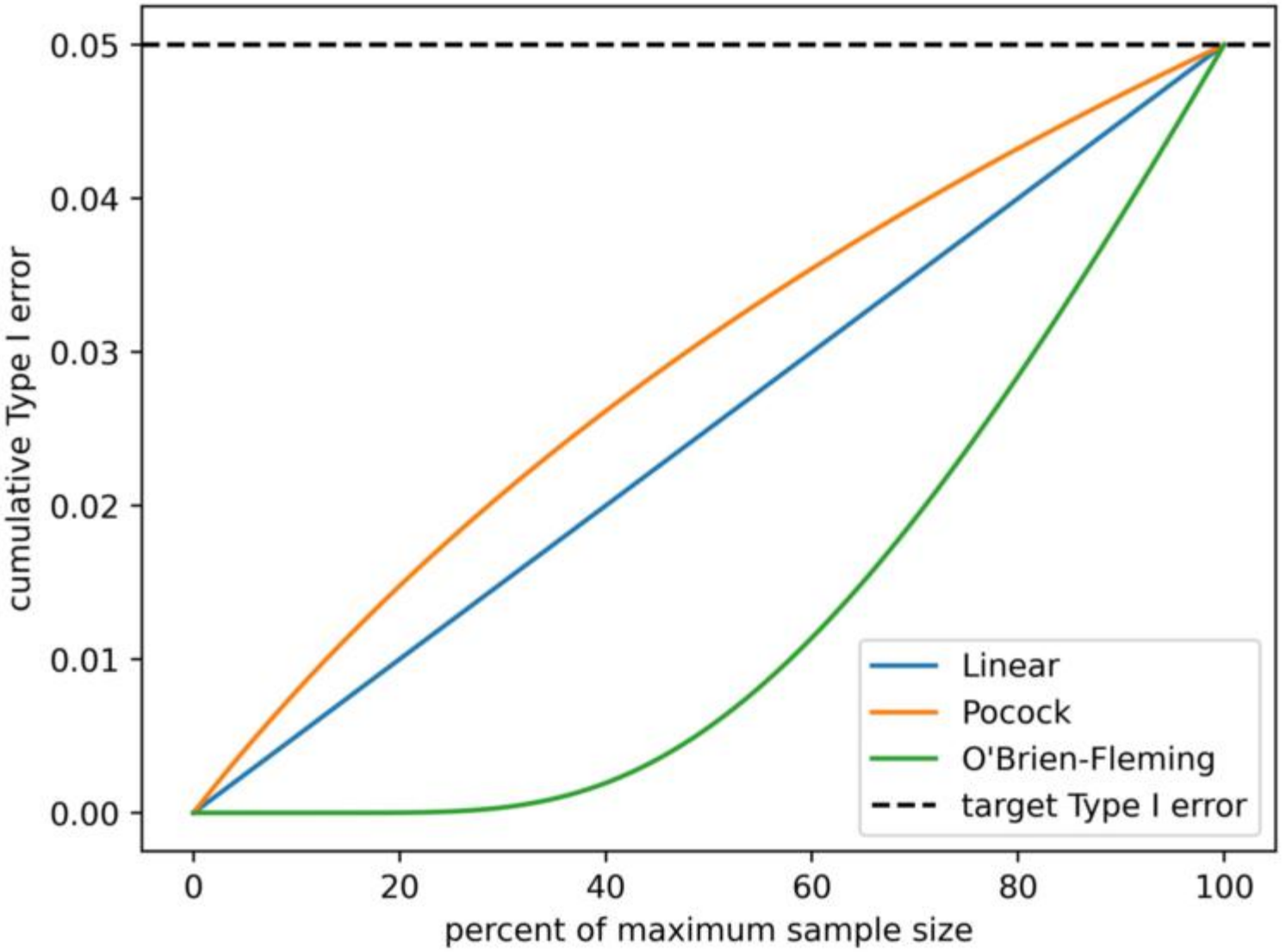
	Intention-to-treat	Per-protocol
	30/100 = 0.3	15/85 = 0.18
	RR = 1	RR = 0.59
	RRR = 0	RRR = 0.41
	30/100 = 0.3	30/100 = 0.3
	No effect (true)	Apparent effect (untrue)

A = medical management + surgery
B = medical management only

Group sequential designs

- Motivation:
  - Fixed sample size may be unethical
  - Sequential designs impracticable + inflation of type I error
- Solution: group sequential design
  - Planned interim analysis (usually 2, can be >2)
  - Should control alpha value
  - Role of independent data monitoring committee (IDMC)
  - Possibility to close early if
    - Larger than expected toxicity/side effects
    - Futility
    - Larger than expected efficacy

Alpha spending functions



GHENT UNIVERSITY

Adaptive trial design

- An adaptive design clinical study is a study that includes a prospectively planned opportunity for modification of one or more of the study design features based on analysis of interim data from subjects in the study

GHENT UNIVERSITY

Design Features	Ancillary Features
Population / Eligibility criteria	Accrual rate
Treatment regimen / Dose	Follow-up time
Endpoint / Timing of endpoint	Overall event rate
Target treatment effect size	Schedule of evaluations
Sample size	
Primary statistical test	
Type I / Type II error rates	
Randomization ratio	

Clinical trials using devices or implants

The road to progress

Get a novel drug approved

Get a novel device approved

Get a novel surgical procedure approved

Benefits

Innovation

Risks

Enable disruptive treatments/methods

Reduce invasiveness, preserve QoL

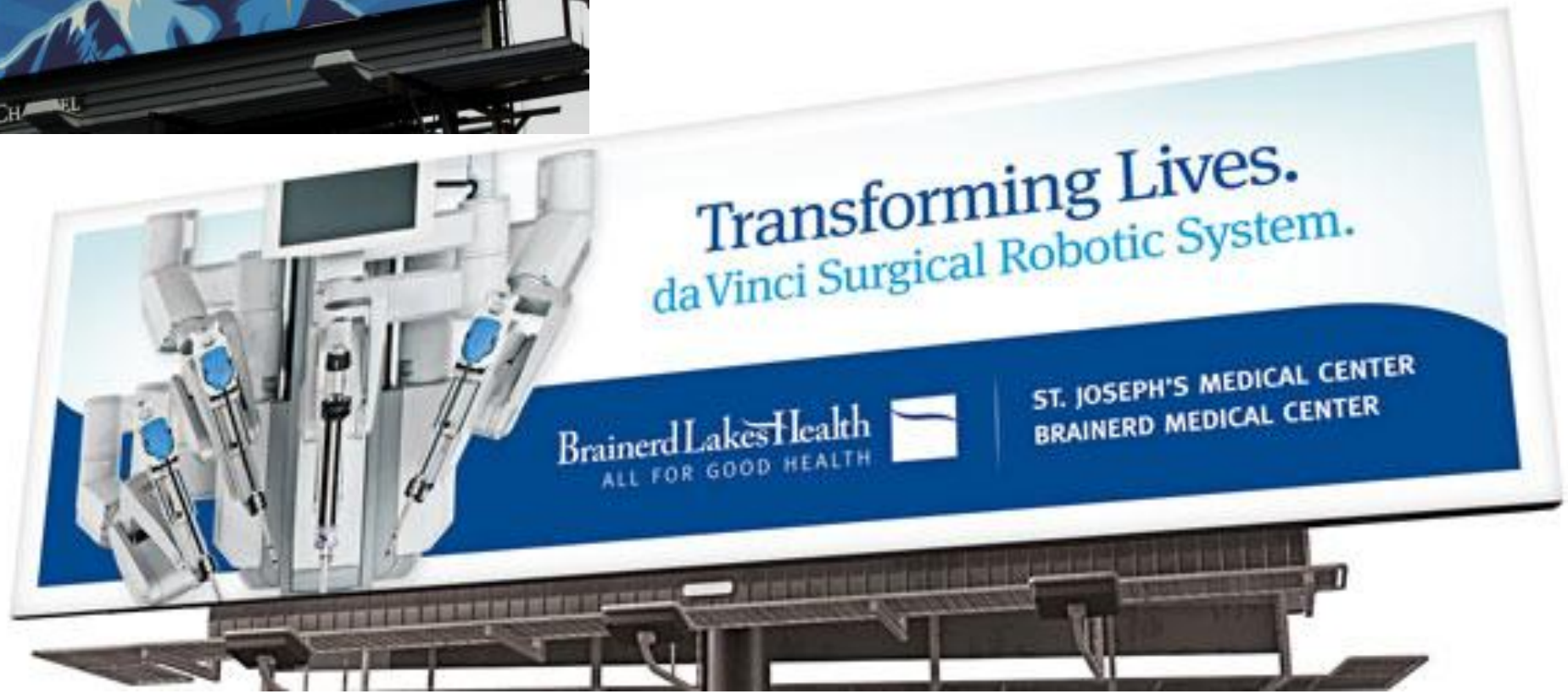
Risk of failed innovations

May foster unreasonable optimism about potential

Runaway diffusion – Buxton's law

Conflicts of interest: financial, prestige


ROBOTIC-ASSISTED SURGERY
MADE ME FAMOUS
BRINGING IT TO BRIGHTON
MAKES ME PROUD
DR. INGOLF TUERK
CHIEF OF UROLOGY
St. Elizabeth's
Medical Center


Transforming Lives.
da Vinci Surgical Robotic System.
Brainerd Lakes Health
ALL FOR GOOD HEALTH
ST. JOSEPH'S MEDICAL CENTER
BRAINERD MEDICAL CENTER

UPDATE: Caution with Robotically-Assisted Surgical Devices in Mastectomy: FDA Safety Communication

Safety Communications

2021 Safety Communications

2020 Safety Communications

2019 Safety Communications

Date Issued: August 20, 2021

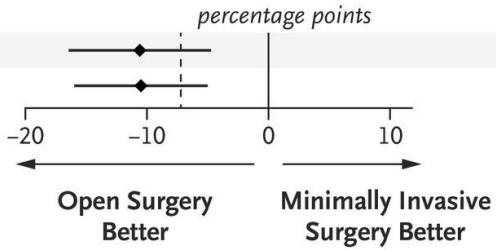
The U.S. Food and Drug Administration (FDA) is reminding patients and health care providers that the safety and effectiveness of robotically-assisted surgical (RAS) devices for use in mastectomy procedures or in the prevention or treatment of breast cancer have not been established. In addition, the FDA is aware of allegations that clinical studies are being conducted using RAS devices to perform mastectomies for the prevention or treatment of cancer without the FDA oversight required for such significant risk studies.

Content current as of:
08/20/2021

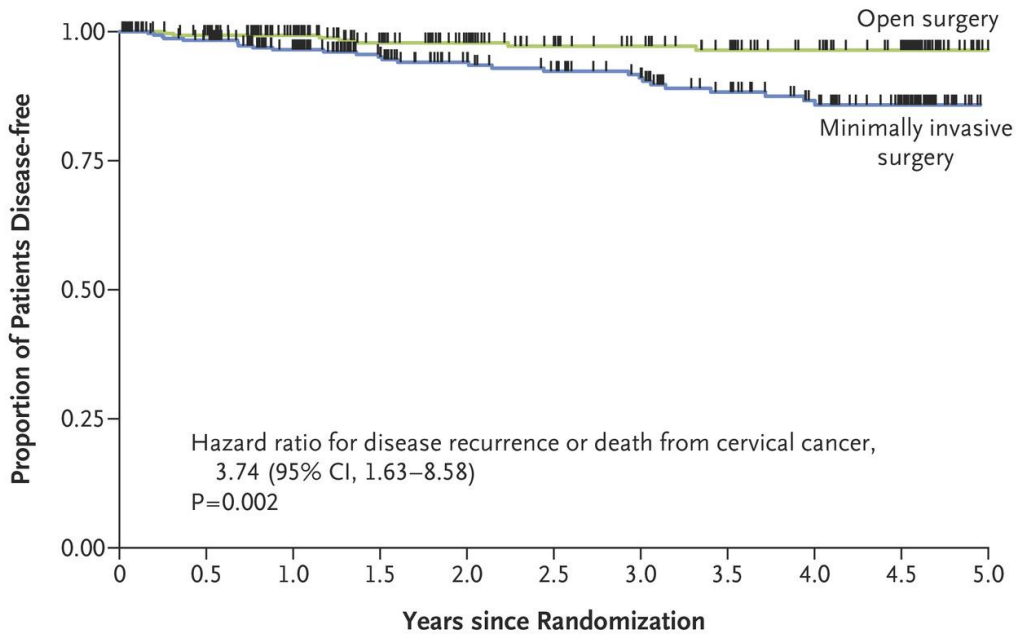
Regulated Product(s)
Medical Devices

A

Population	Disease-free Survival Rate at 4.5 Yr (95% CI)		Difference (95% CI)	P Value for Noninferiority
	Minimally Invasive Surgery	Open Surgery		
	percent		*percentage points*	
Intention-to-treat population	86.0 (79.7–90.4)	96.5 (92.7–98.4)		0.87
Per-protocol population	87.1 (81.0–91.3)	97.6 (94.1–99.0)		0.88



B



No. at Risk

Open surgery	312	280	236	187	163	144	134	123	104	90	7
Minimally invasive surgery	319	292	244	192	167	155	142	121	102	80	5

Ramirez NEJM 2018

BOOKS APRIL 20, 2020 ISSUE

DO SOME SURGICAL IMPLANTS DO MORE HARM THAN GOOD?

Many are clearly lifesaving, but others have proved to be life-threatening, and dangerous implants are marketed with scant oversight.

By Jerome Groopman

April 13, 2020

‘It is more likely for Toyota to know about faulty exhaust pipes in a Prius than DePuy to understand how a new hip implant is performing in the United States’

The New Medical Device Regulation (MDR)

What is the MDR?

The Medical Device Regulation (MDR) is a new regulation that replaces the Medical Device Directive (MDD) 93/42/EEC and Active Implantable Medical Devices (AIMD) Directive 90/385/EEC. It applies to all medical device manufacturers who intend to place their products in the European Union (EU).

Key changes in MDR

- Wider scope of regulated medical devices
- More stringent clinical evidence and documentation
- Increased focus on identification and traceability
- Definition of common specifications
- Unannounced factory audits
- Increased Notified Body authority and/or involvement
- More rigorous vigilance and market surveillance
- At least one person responsible for regulatory compliance

Timeline

2012 – Poly Implant Prothèse (PIP) scandal sparks global health scare

2014 – EU Commission publishes proposal for MDR, seeking to strengthen existing MDD

2015 – EU Parliament reviews draft and proposes changes

EU Council reviews draft and proposes additional changes

May 2017 – Publication of MDR

Notified Bodies can request re-designation and manufacturers can place devices on the market under new MDR

2020

End of three-year transition period

Global medical devices market

£333 BILLION

Size of global medical devices market in 2020

R&D expenditure for medtech globally to grow 3.5% annually by 2020

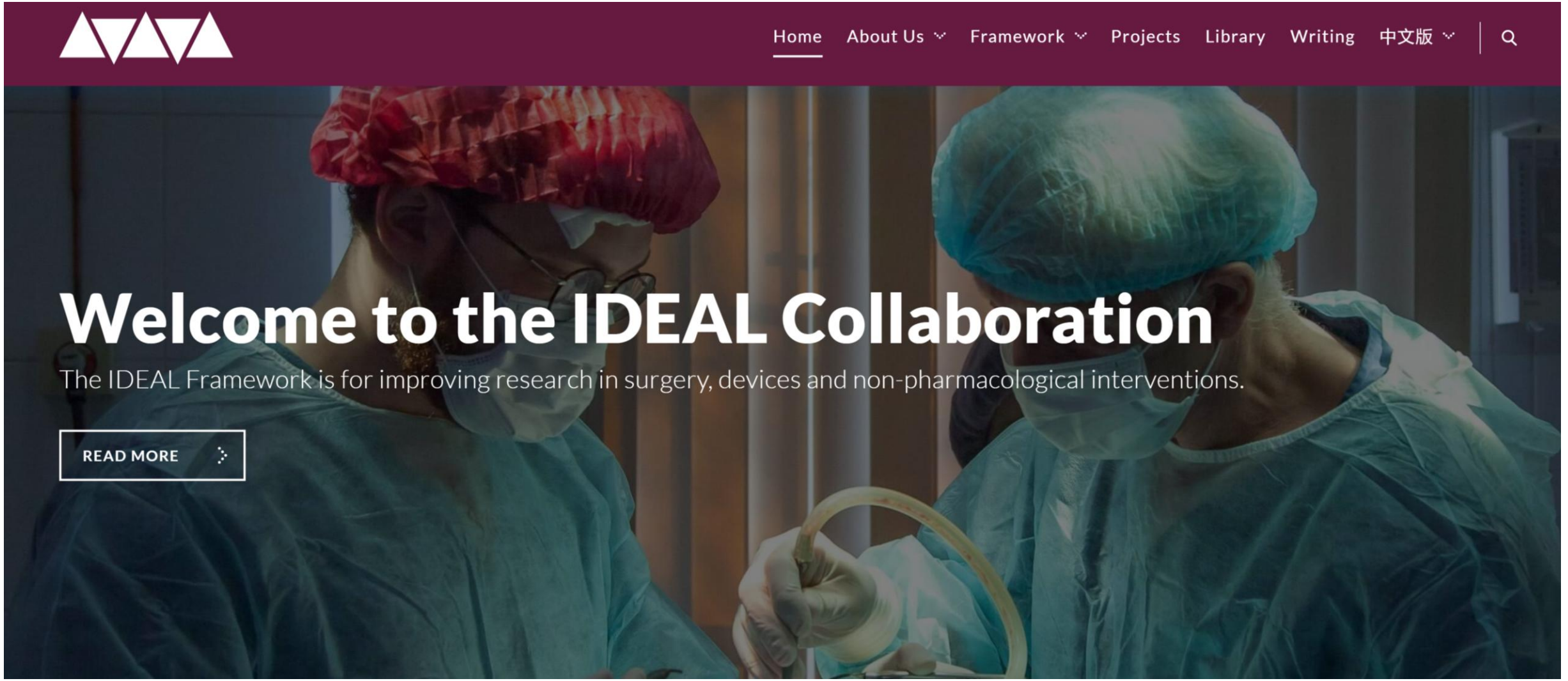
The EU accounts for one-third of the global medical device market

Sources: www.evaluategroup.com/public/reports/EvaluateMedTech-World-Preview-2015.aspx www.medtecheurope.org/

To find out more about how LRQA can help you with your requirements, visit **lrqa.co.uk/mdr**, email **enquiries@lrqa.co.uk** or call **0800 783 2179**

Lloyd's Register LRQA

Improving performance, reducing risk


Home
About Us
Framework
Projects
Library
Writing
中文版
Welcome to the IDEAL Collaboration
The IDEAL Framework is for improving research in surgery, devices and non-pharmacological interventions.
READ MORE

	1 Idea	2a Development	2b Exploration	3 Assessment	4 Long-term study
Purpose	Proof of concept	Development	Learning	Assessment	Surveillance
Number and types of patients	Single digit; highly selected	Few; selected	Many; may expand to mixed; broadening indication	Many; expanded indications (well defined)	All eligible
Number and types of surgeons	Very few; innovators	Few; innovators and some early adopters	Many; innovators, early adopters, early majority	Many; early majority	All eligible
Output	Description	Description	Measurement; comparison	Comparison; complete information for non-RCT participants	Description; audit, regional variation; quality assurance; risk adjustment
Intervention	Evolving; procedure inception	Evolving; procedure development	Evolving; procedure refinement; community learning	Stable	Stable
Method	Structured case reports	Prospective development studies	Research database; explanatory or feasibility RCT (efficacy trial); diseased based (diagnostic)	RCT with or without additions/ modifications; alternative designs	Registry; routine database (eg, SCOAP, STS, NSQIP); rare-case reports
Outcomes	Proof of concept; technical achievement; disasters; dramatic successes	Mainly safety; technical and procedural success	Safety; clinical outcomes (specific and graded); short-term outcomes; patient-centred (reported) outcomes; feasibility outcomes	Clinical outcomes (specific and graded); middle-term and long-term outcomes; patient-centred (reported) outcomes; cost-effectiveness	Rare events; long-term outcomes; quality assurance
Ethical approval	Sometimes	Yes	Yes	Yes	No
Examples	NOTES video[6]	Tissue engineered vessels[7]	Italian D2 gastrectomy study[8]	Swedish obese patients study[9]	UK national adult cardiac surgical database[10]

RCT=randomised controlled trial. SCOAP=Surgical Clinical Outcomes Assessment Programme. STS=Society of Thoracic Surgeons. NSQIP=National Surgical Quality Improvement Program. NOTES=natural orifice translumenal endoscopic surgery.

Table: **Stages of surgical innovation**

Methodological obstacles for RCTs with devices/procedures

- Lack of standardisation
  - Skill and preference dependence
  - Learning curve effects, Buxton's law
- Impossibility to blind (mask) patients
- Ethical challenges of 'sham' surgery

GHENT UNIVERSITY

When to evaluate a novel procedure?

- Too early: risk = evolving results → unfair evaluation
- Too late: risk = established procedure → difficult to dislodge from practice

***Buxton's law*: 'it's always too early until, unfortunately, it's suddenly too late!'**

Martin J Buxton. Problems in the economic appraisal of new health technology: the evolution of heart transplants in the U.K. in: M.F. Drummond (Ed.), Economic Appraisal of Health Technology in the European Community, Oxford University Press, Oxford (1987)

Alternative *prospective* study designs

- **Non randomized designs**
  - Cohort studies
  - Case-control studies
  - Interrupted time series
- **Modified randomized**
  - Cluster randomized trials: stepped wedge
  - pragmatic RCTs
  - Registry-based RCTs
  - Trials-within-cohorts (TwiCs)
  - Patient preference designs: Zelen, Wennberg, comprehensive cohort
  - Expertise based trials
  - Tracker or adaptive trials (Bayesian)

GHENT UNIVERSITY

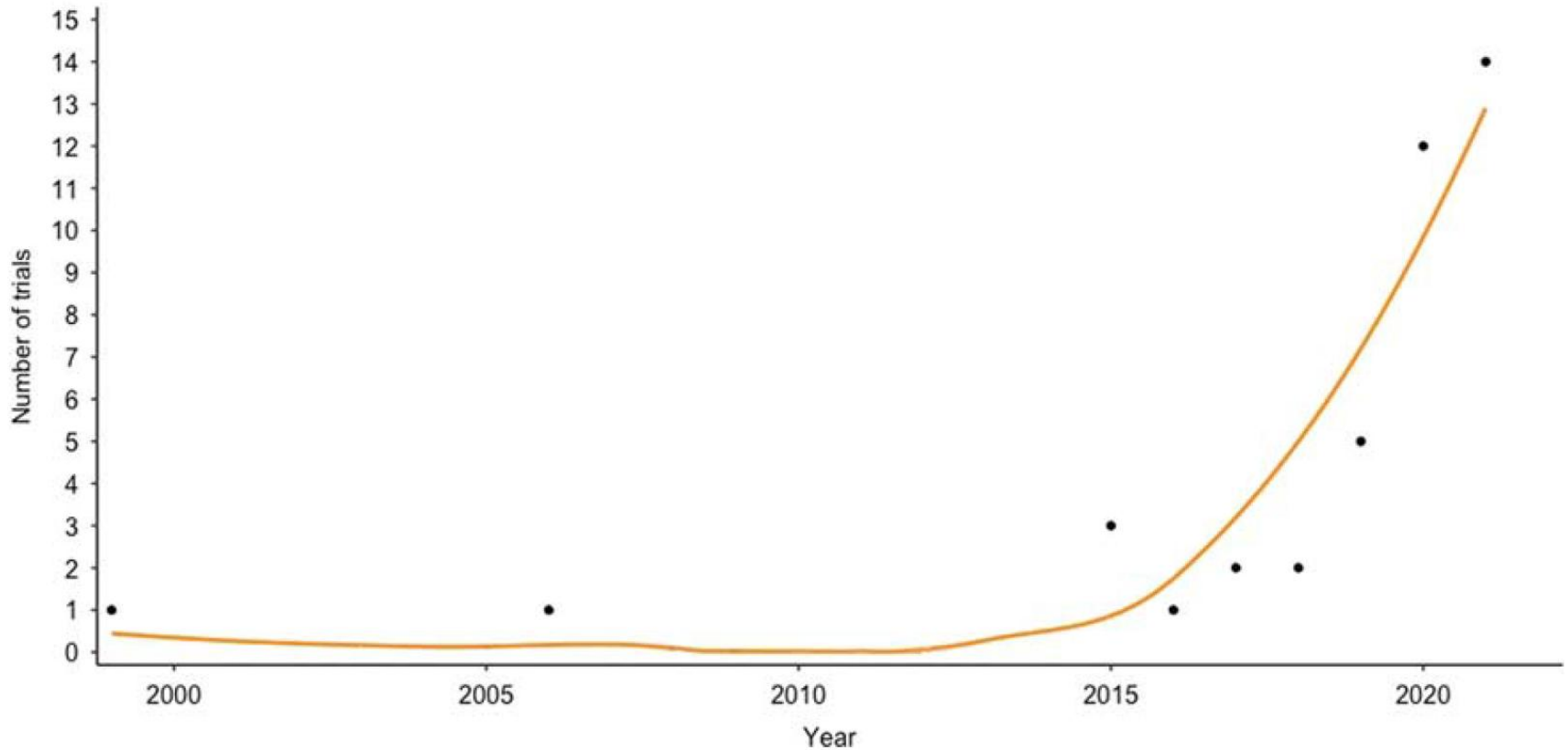
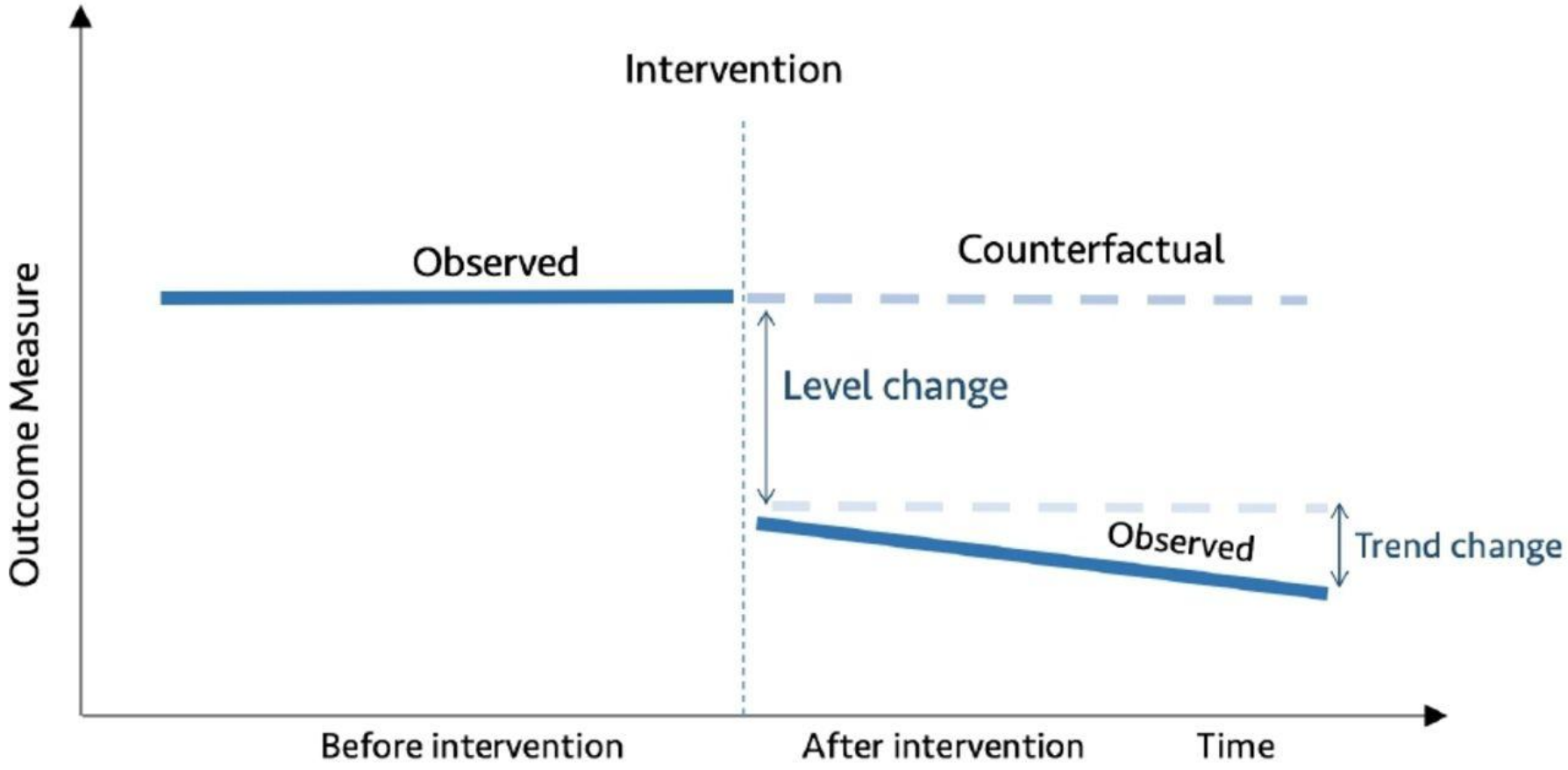
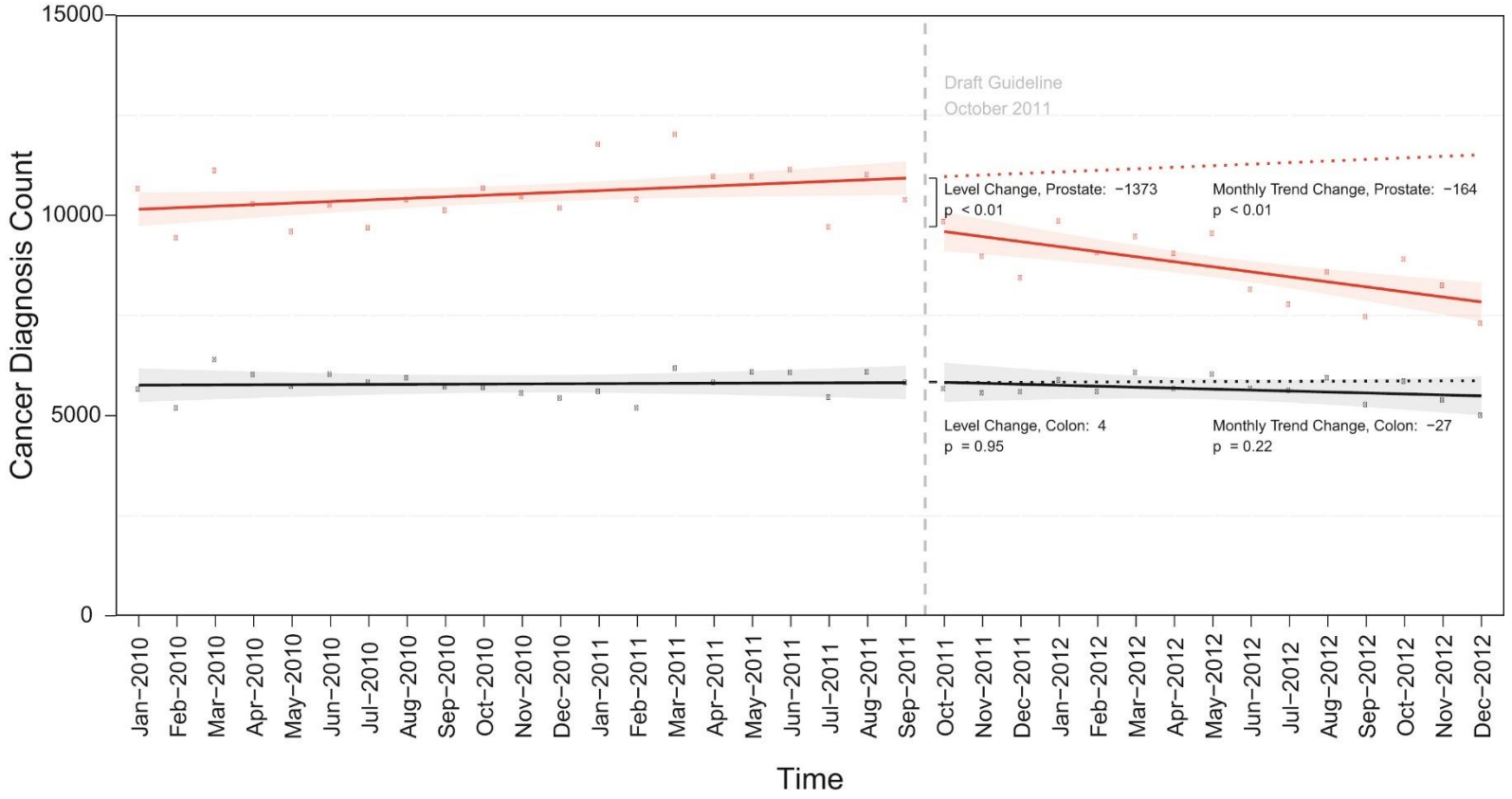


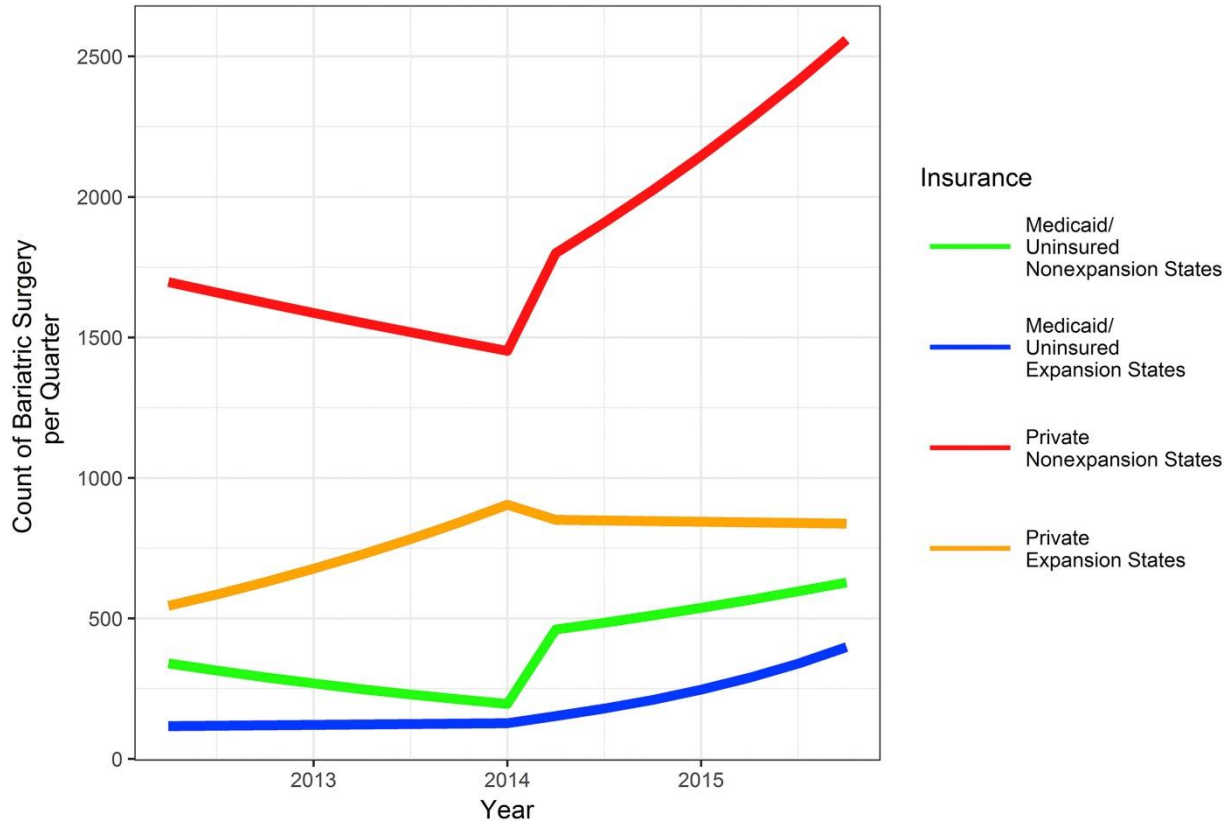
FIGURE 2. Time trend in published surgical RCTs and RCT protocols with alternative designs.

Interrupted time series





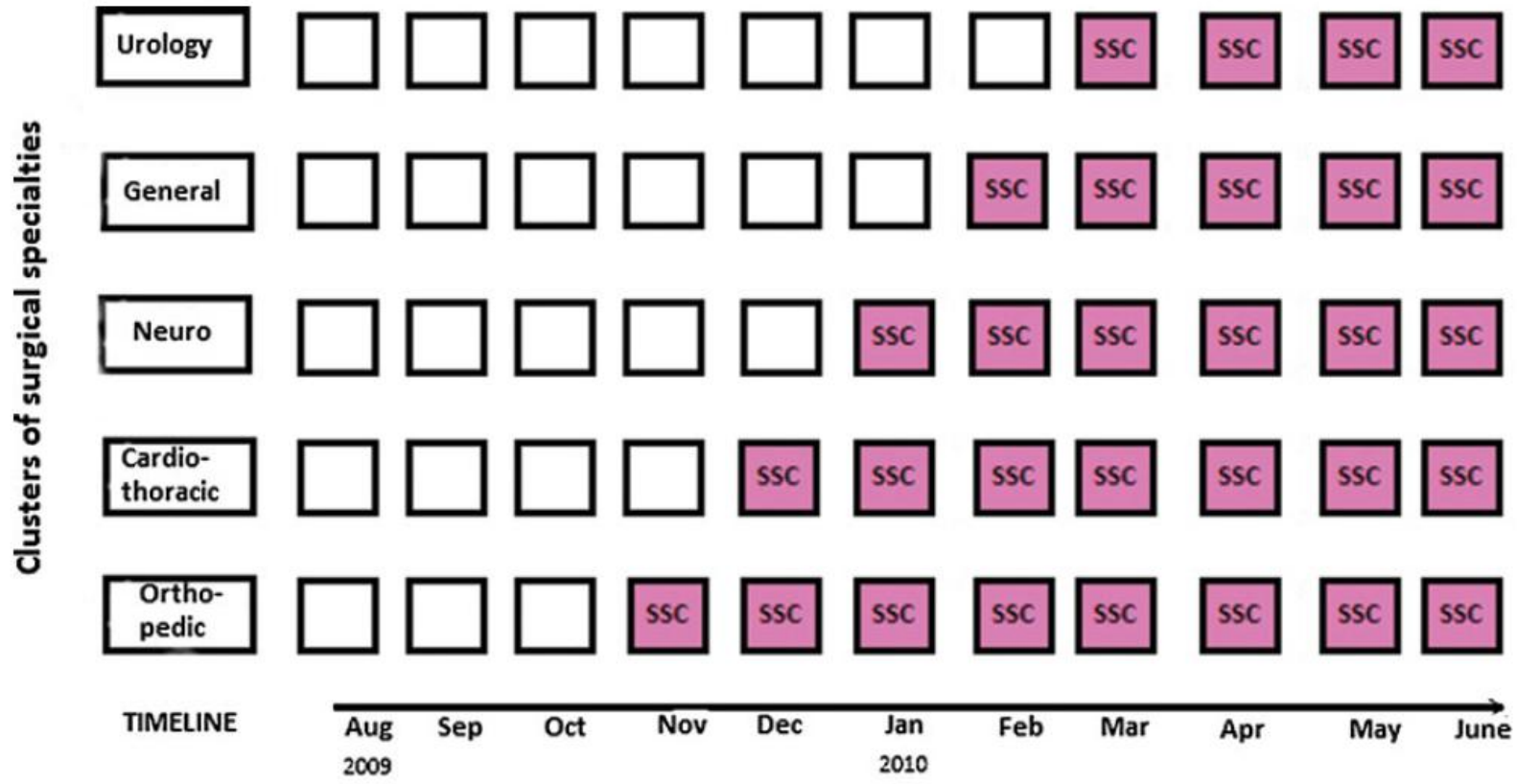
Barocas DA. Effect of the USPSTF Grade D Recommendation against Screening for Prostate Cancer on Incident Prostate Cancer Diagnoses in the United States. J Urology 2015



Bariatric surgery among vulnerable populations: The effect of the Affordable Care Act's Medicaid expansion Gould, Surgery 2019

Stepped wedge design

- Unit of randomization = cluster (hospital,…)
- Sequential roll-out to all clusters over time
- By the end of the study, all clusters will have received experimental intervention
- Used mainly when P(success) perceived as high



Lynch N. Effect of the World Health Organization Checklist on Patient Outcomes: A Stepped Wedge Cluster Randomized Controlled Trial. Ann Surg 2016

Pragmatic RCTs

Test effectiveness

Focus on external validity

Loose inclusion criteria

Reflects ‘real world’ efficacy

Table 2 Outcome of explanatory and pragmatic trials[12]

	Intervention better than control	Intervention equal to or worse than control
Explanatory trial	Equivocal—Will the intervention work in my patients?	Clear—Do not implement this intervention.
Pragmatic trial	Clear—Implement this intervention.	Equivocal—Why did the intervention not work?

Explanatory RCTs

Test efficacy

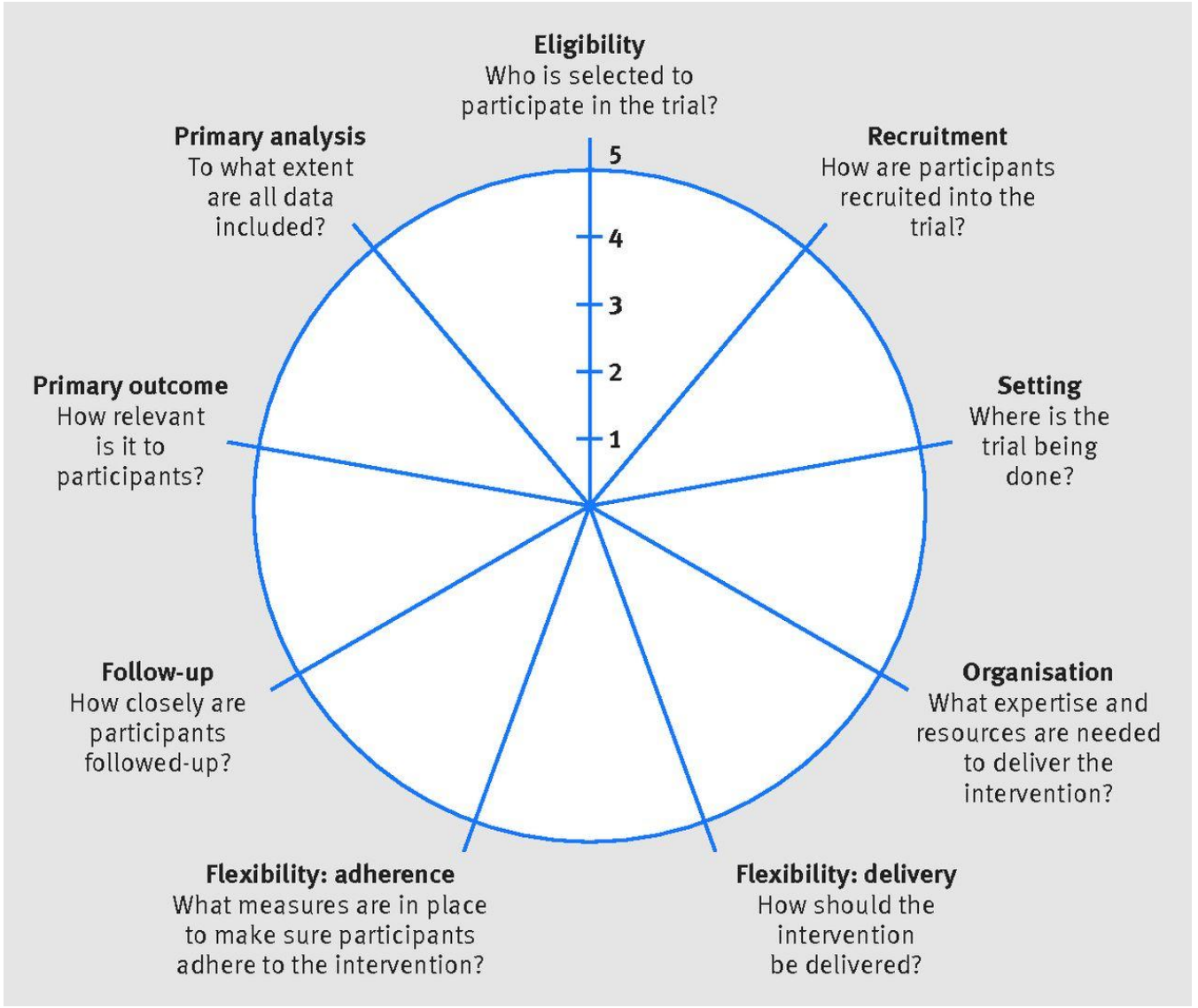
Focus on internal validity

Strict inclusion criteria

Tends to overestimate ‘real world’ efficacy

The PRagmatic-Explanatory Continuum Indicator Summary 2 (PRECIS-2) wheel

ranges from 9 points (indicating a very explanatory study) to 45 points (indicating a very pragmatic study)

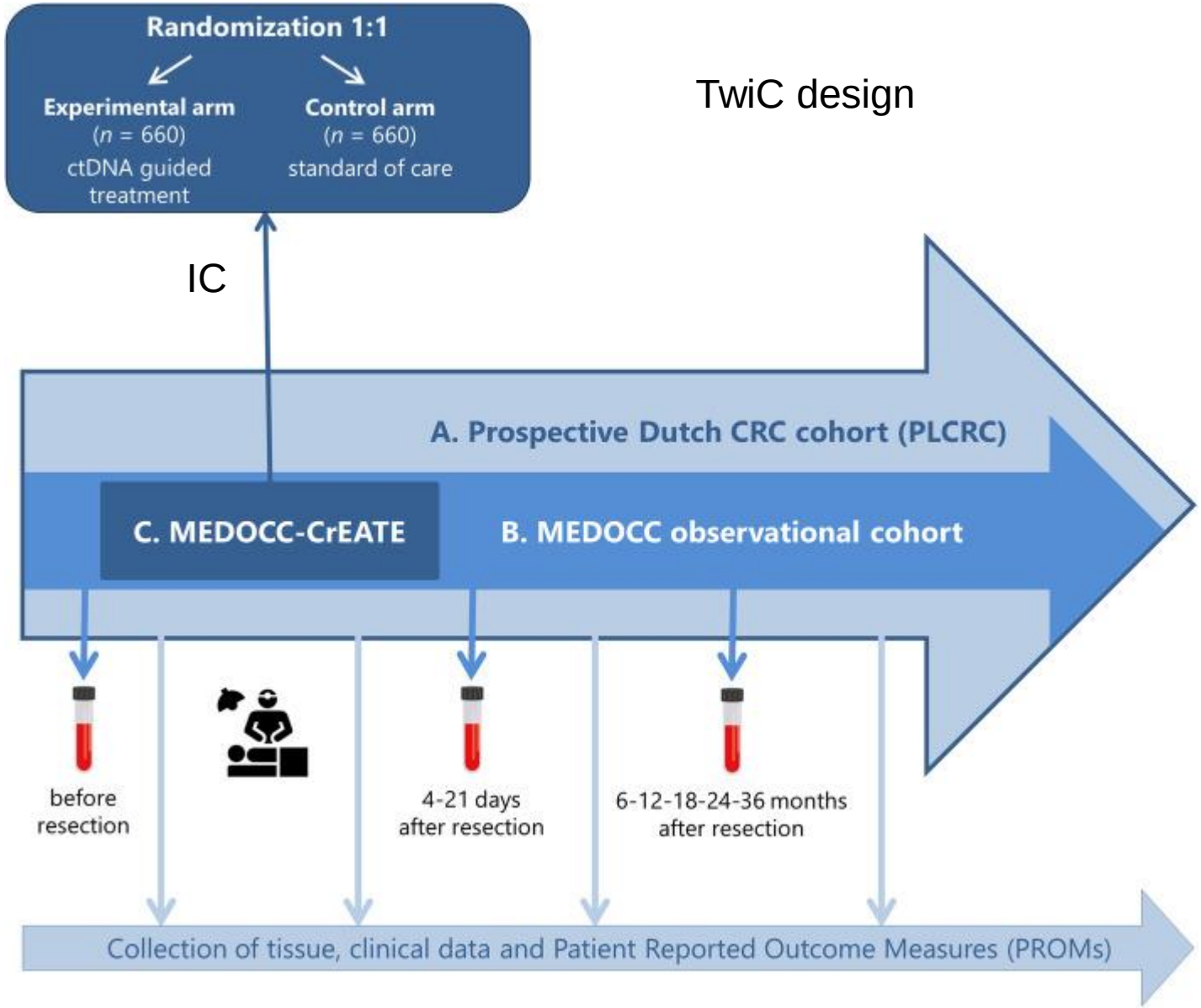


Kirsty Loudon BMJ 2015

TwiC design

IC

IC



Schraa BMC Cancer 2020

Zelen's design

- Patients are randomised **before** they give consent to participate in the trial.
- Standard treatment group: not told that they are part of the trial
- Interventional group: are told that they are part of the trial; if they refuse to participate in the trial, they are given the standard treatment but analysed as if they had received the experimental intervention
- Avoids bias in control group when patient blinding impossible
- Controversial ethics

GHENT UNIVERSITY

Research Article

EASL EUROPEAN ASSOCIATION FOR THE STUDY OF THE LIVER | JOURNAL OF HEPATOLOGY

A randomized controlled trial of radiofrequency ablation and surgical resection in the treatment of small hepatocellular carcinoma

Kai Feng[1,2,3,†], Jun Yan[1,†], Xiaowu Li[1], Feng Xia[1], Kuansheng Ma[1,*], Shuguang Wang[1], Ping Bie[1], Jiahong Dong[1]

Since this trial was a comparative study between a new and a standard therapy, the double-blind technique was considered impractical. In addition, many patients and physicians in China strongly believe that surgical resection is a more definitive treatment for HCC. Therefore, some patients may refuse to participate or refuse randomization. Therefore, we used the Zelen method [22] to randomly divide groups and to include as many eligible patients as possible, while simulta-

JAMA Internal Medicine | Original Investigation

Effectiveness of a Nurse-Led Multidisciplinary Intervention vs Usual Care on Advance Care Planning for Vulnerable Older Adults in an Accountable Care Organization
A Randomized Clinical Trial

Jennifer Gabbard, MD; Nicholas M. Pajewski, PhD; Kathryn E. Callahan, MD, MS; Ajay Dharod, MD; Kristie L. Foley, PhD; Keren Ferris, MPH; Adam Moses, MHA; James Willard, MAS; Jeff D. Williamson, MD, MPH

IMPORTANCE Advance care planning (ACP), especially among vulnerable older adults, remains underused in primary care. Additionally, many ACP initiatives fail to integrate directly into the electronic health record (EHR), resulting in infrequent and disorganized documentation.

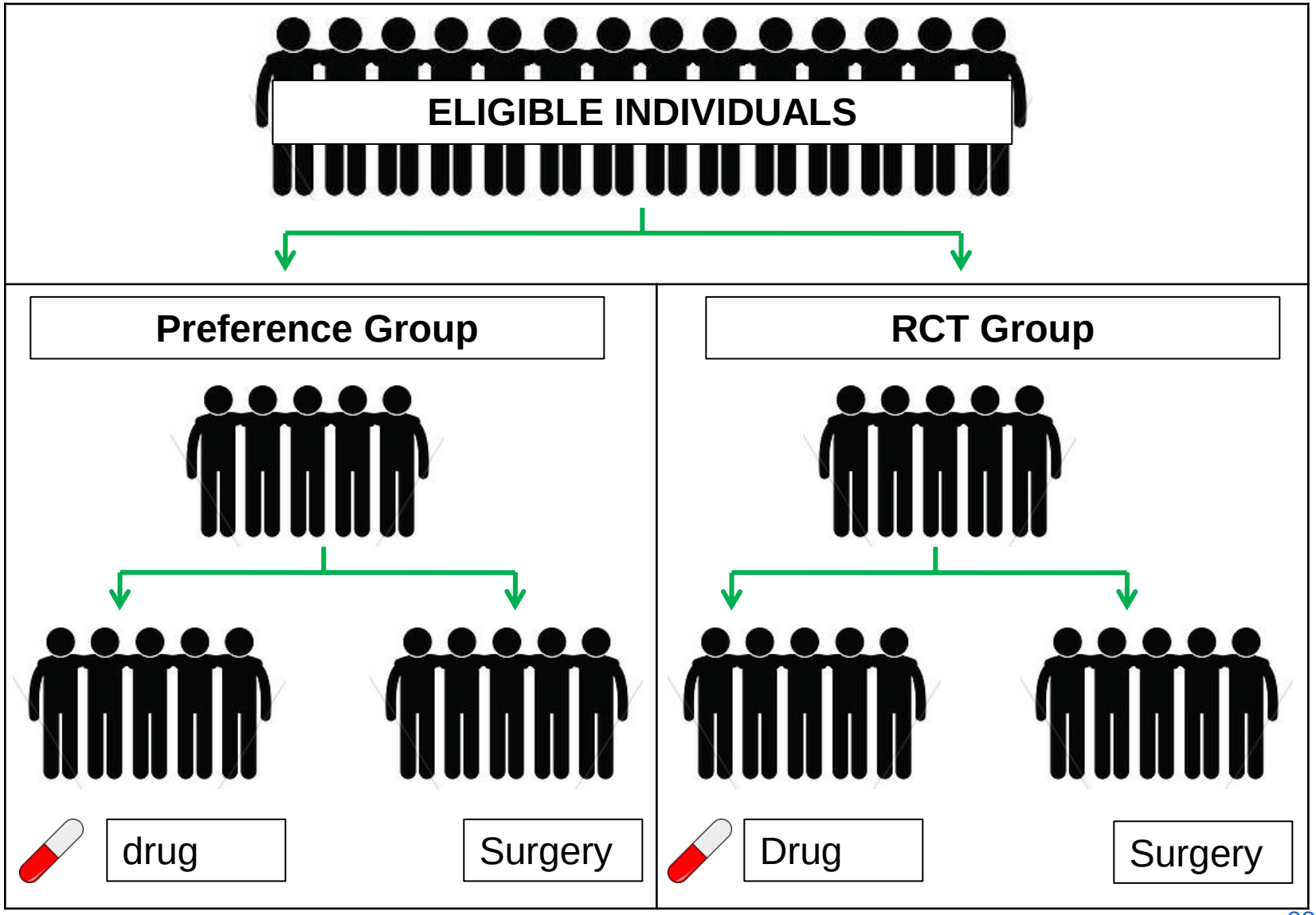
OBJECTIVE To determine whether a nurse navigator–led ACP pathway combined with a health care professional–facing EHR interface improves the occurrence of ACP discussions and their documentation within the EHR.

DESIGN, SETTING, AND PARTICIPANTS This was a randomized effectiveness trial using the Zelen design, in which patients are randomized prior to informed consent, with only those randomized to the intervention subsequently approached to provide informed consent.

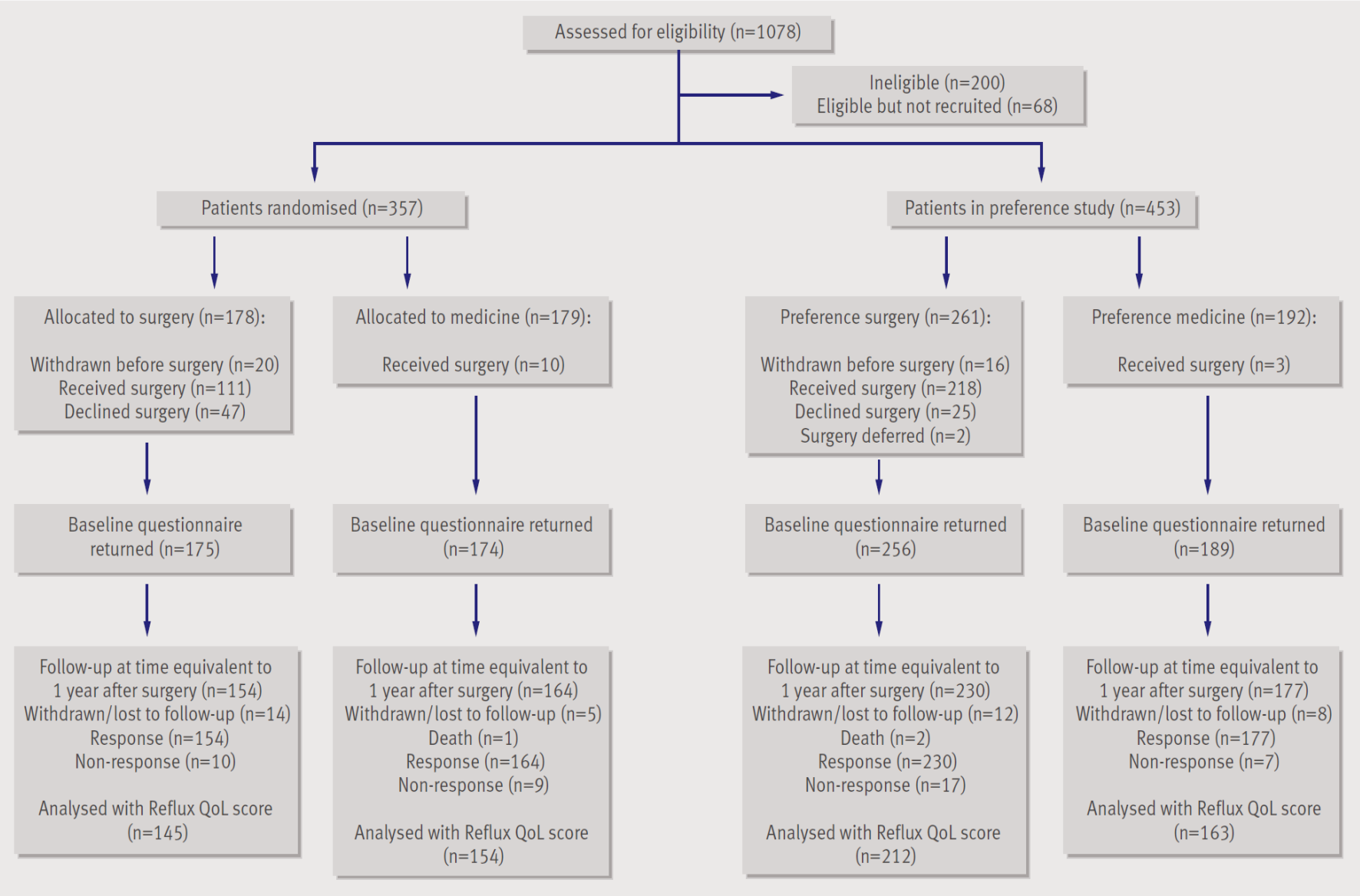
Visual Abstract

Editorial page 309

Supplemental content

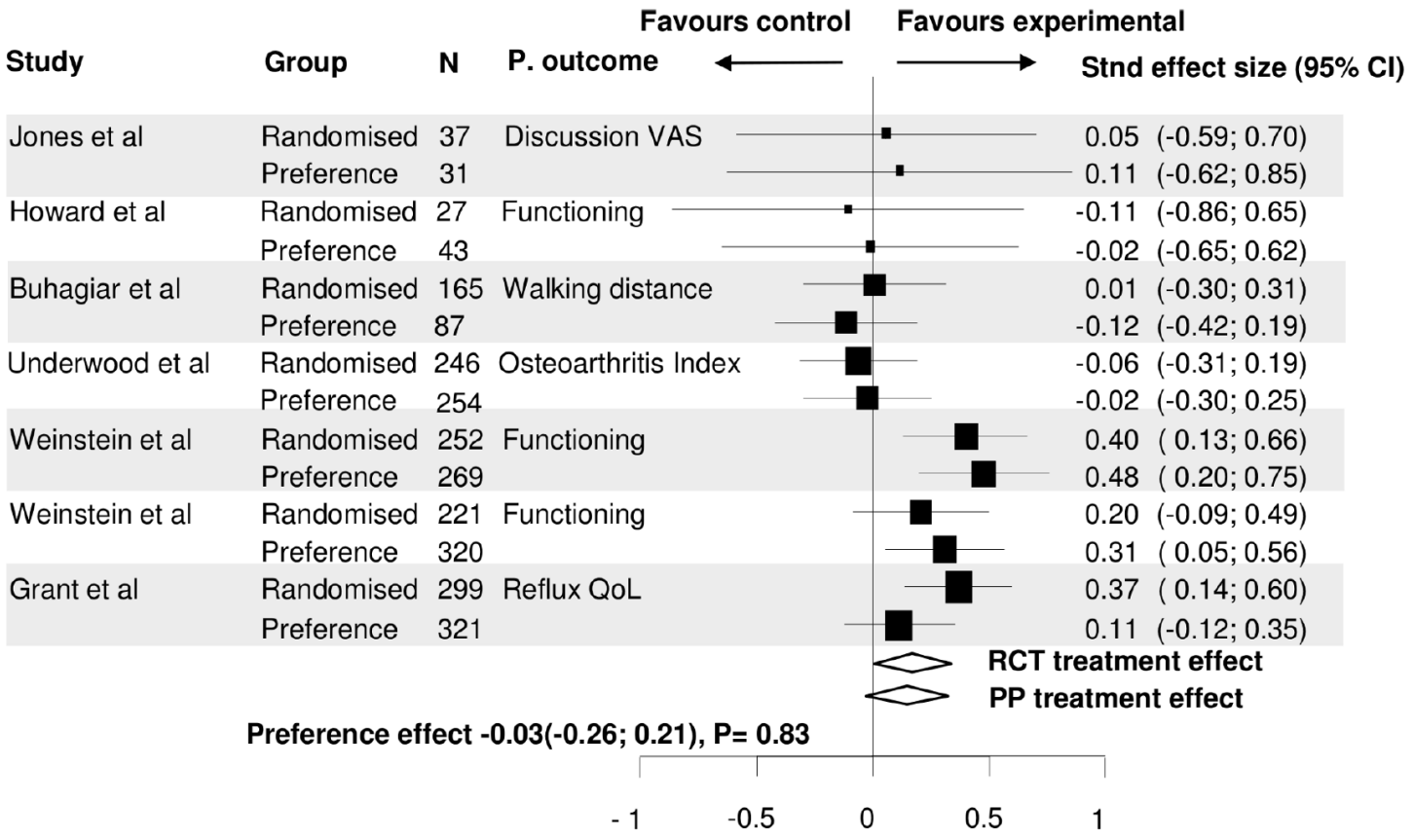


Wennberg's preferential design



Grant BMJ 2008

Partially randomized patient preference trials do not influence the primary outcome

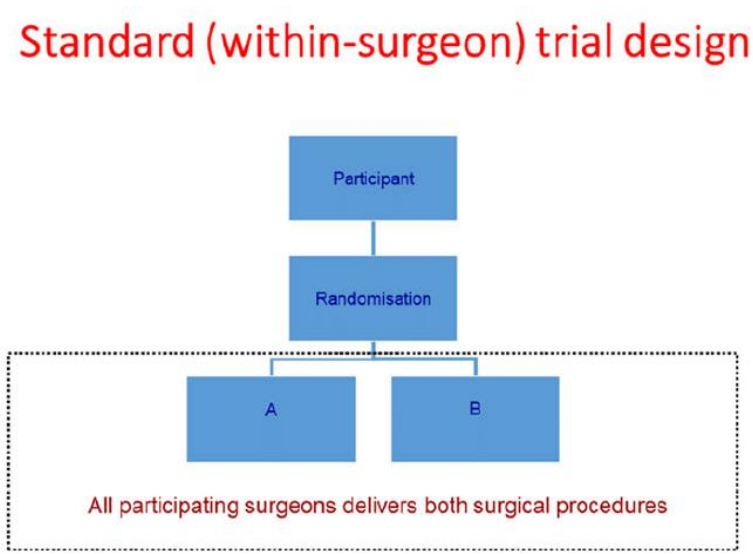


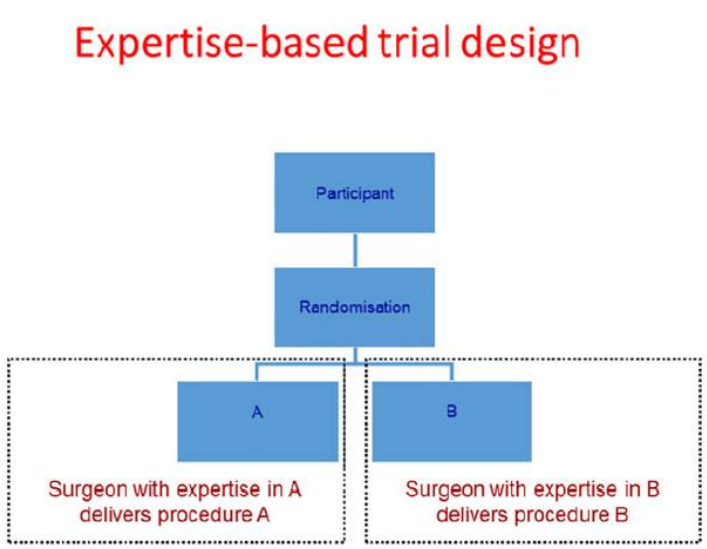
Wasmann BMJ Open 2019

Expertise based RCT

Table 1 Participants' perceived potential advantages and disadvantages of an expertise-based versus a standard trial design

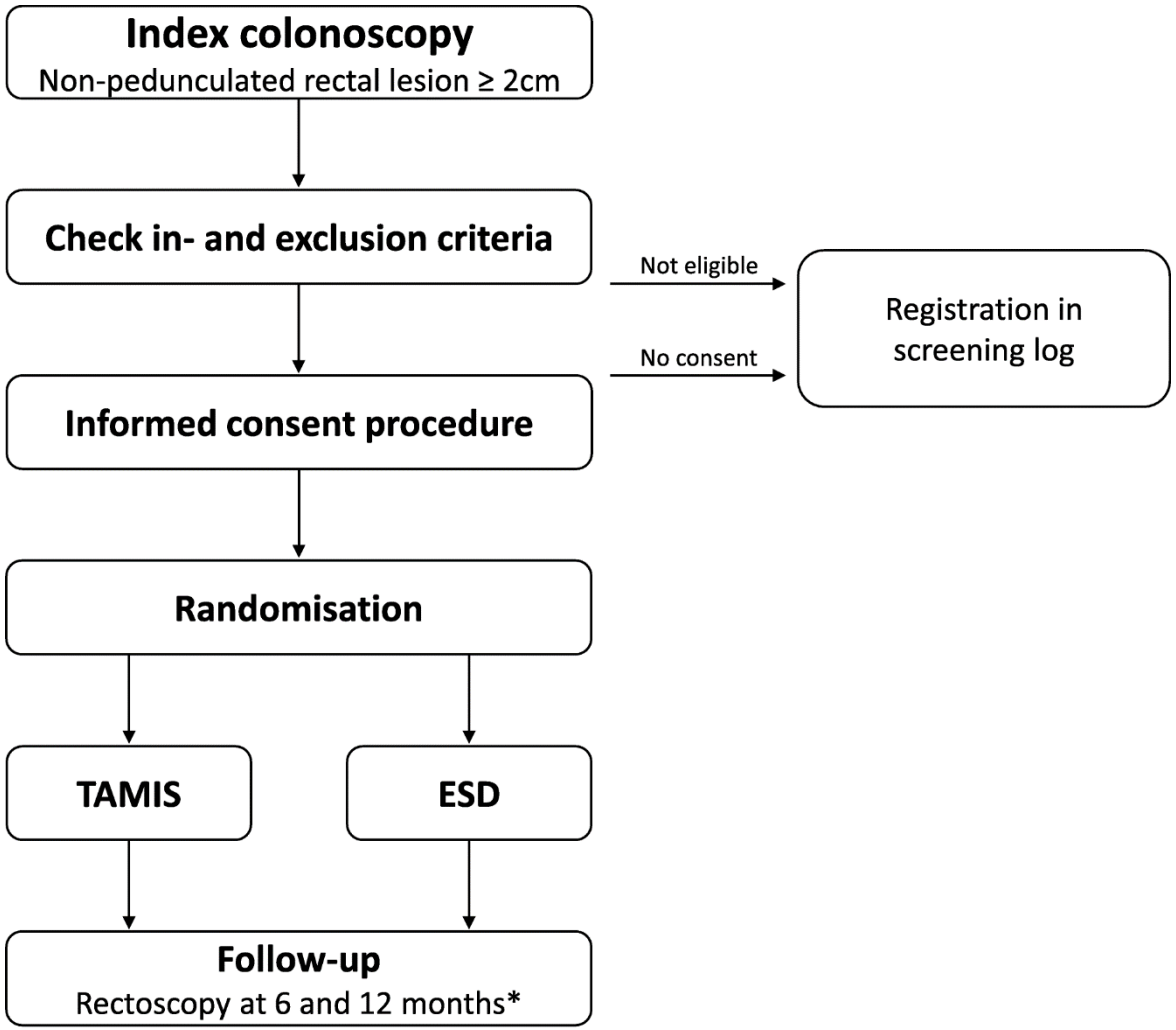
Expertise-based versus standard trial design
Advantages
Greater accommodation of surgeons' treatment preferences
Treatments performed in their 'best light'
More appealing to patients
Better suited to some clinical settings
Disadvantages
Added complexity in terms of site set and administration, including greater co-ordination between surgeons required
Design specific challenges which need to be addressed (e.g. defining an expert)
Impact upon the patient-surgeon relationship
Relation to clinical practice
Perception of stakeholders





Cook Trials 2018

TRIASSIC trial: Expertise based RCT

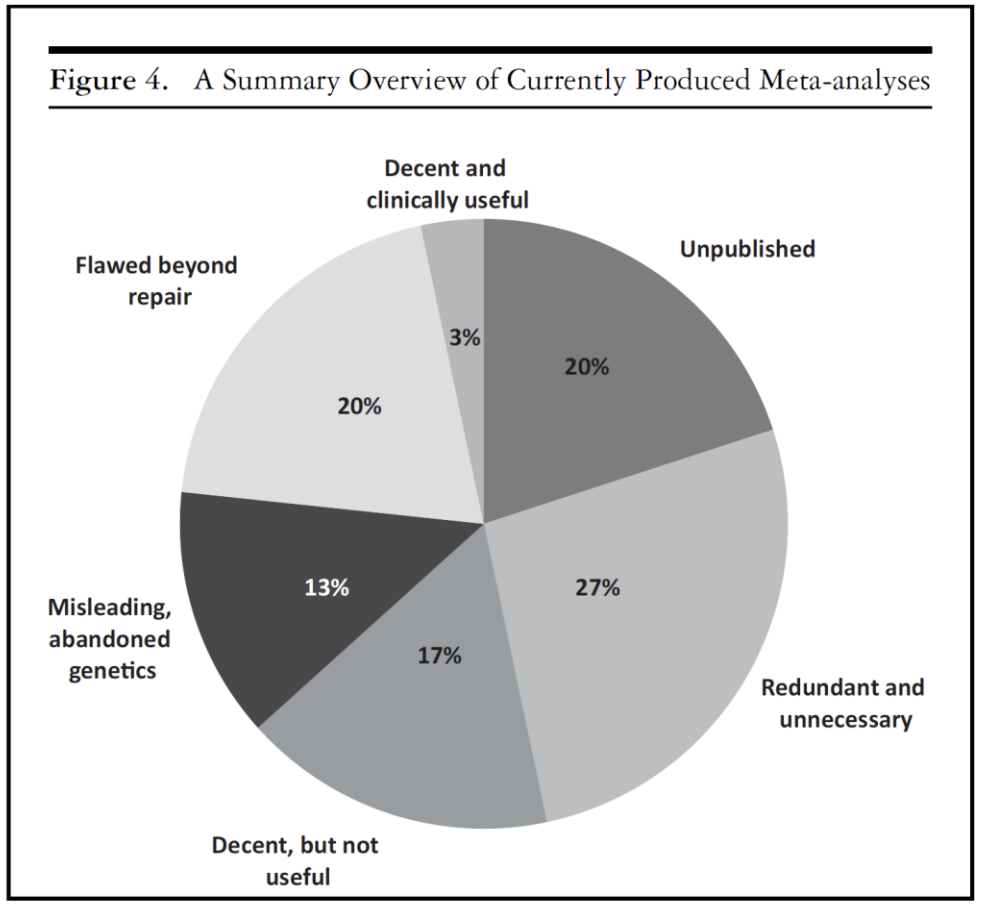


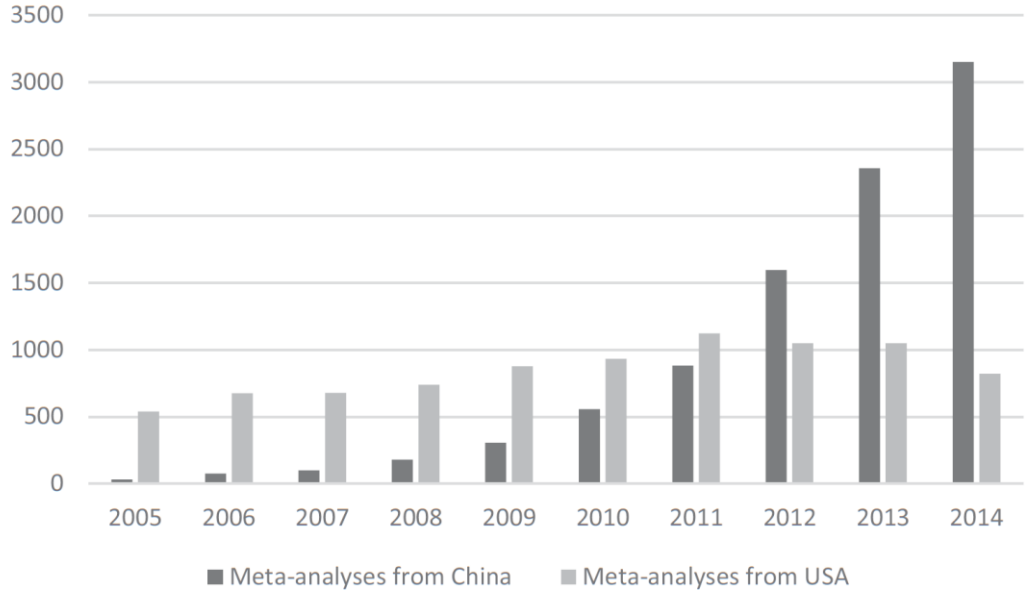
Dekkers BMC Gastroenterol 2020

How to appraise systematic reviews and meta-analyses

The Mass Production of Redundant, Misleading, and Conflicted Systematic Reviews and Meta-analyses

JOHN P.A. IOANNIDIS





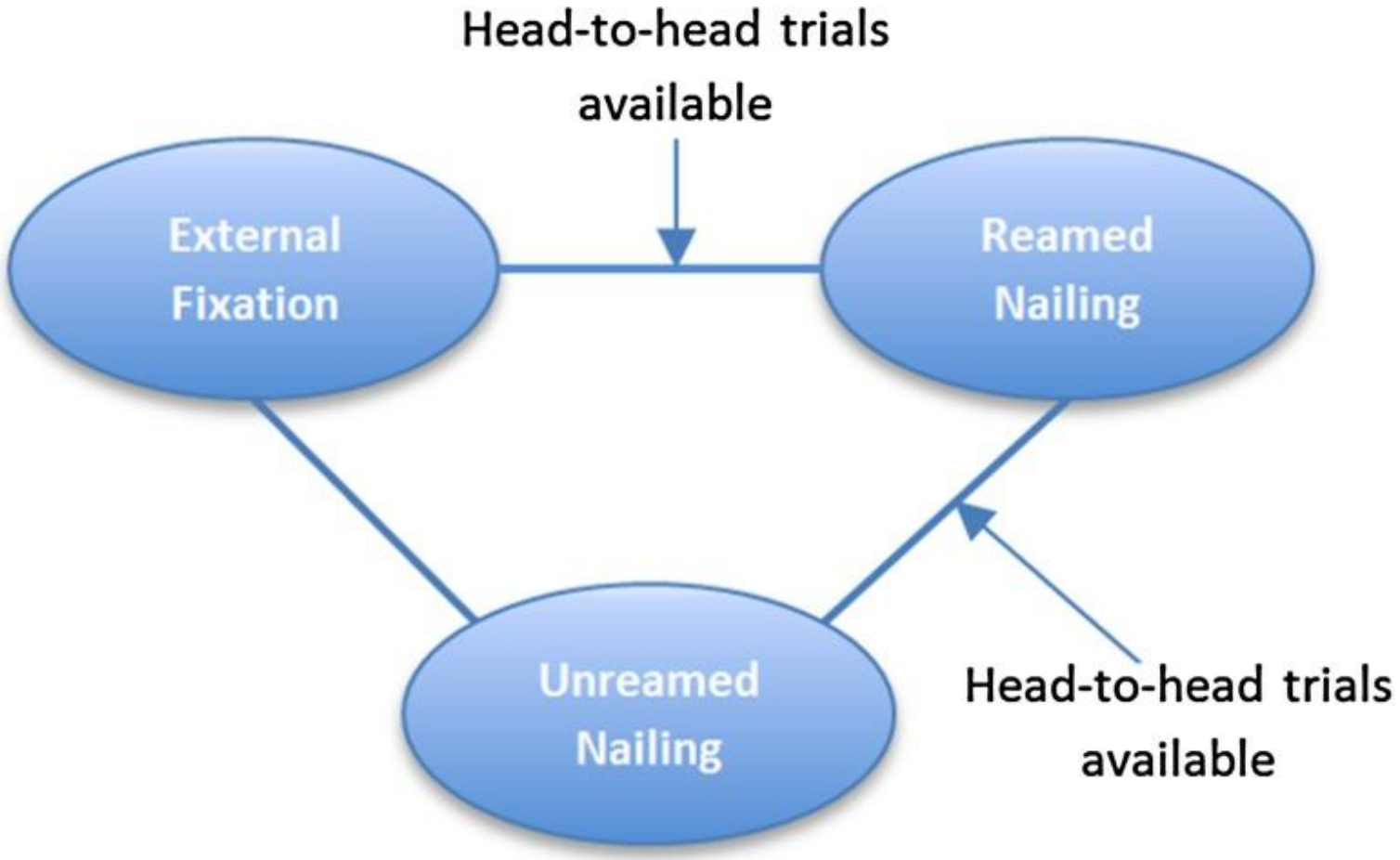
Milbank Q. 2016 Sep; 94(3): 485–514

Types of Review

- Narrative review
- Scoping review
- Systematic review (from comprehensive, systematic literature search)
  - Int. Register: //www.crd.york.ac.uk/PROSPERO/
- Meta-analysis: SR with calculation of **summary statistics**
- Meta-analysis based on individual patient data (IPD)
- Network meta-analysis

GHENT UNIVERSITY

Network meta-analysis



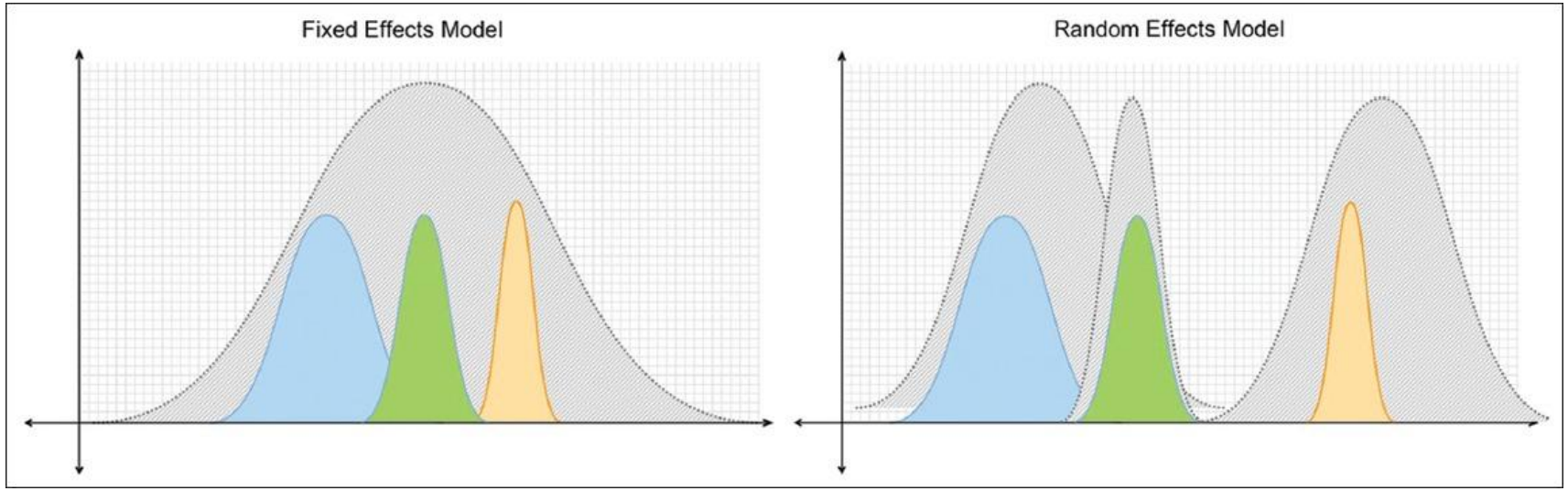
What to appraise in SR/MA

- Search strategy: encompassing?
- Inclusion/exclusion criteria; restrictions
- Statistical heterogeneity
- Fixed versus random effects meta-analysis
- Test for publication bias
- Sensitivity analyses

GHENT UNIVERSITY

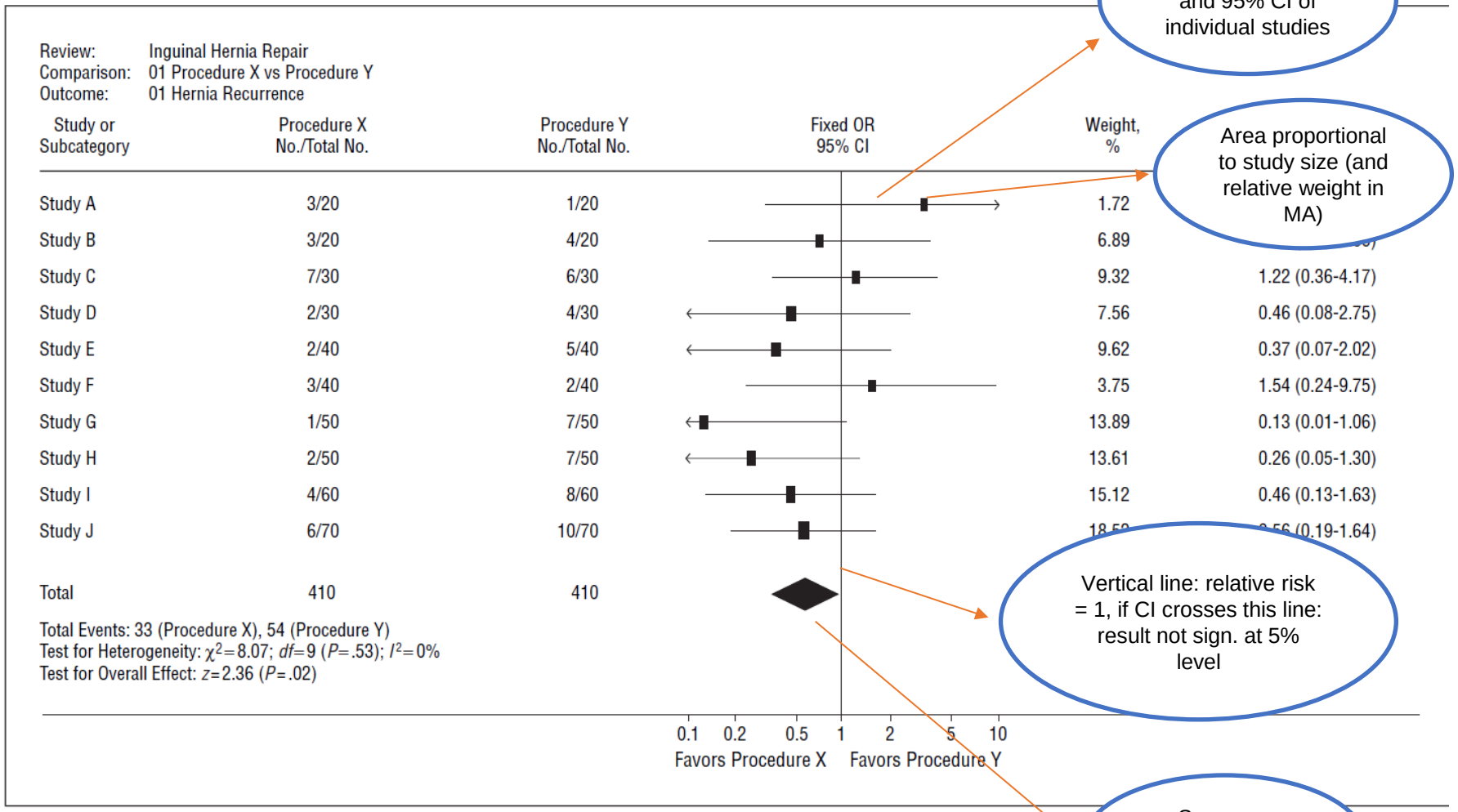
Meta-analysis

- Outcome measures
  - Binary: OR or RR
  - Continuous: weighted mean difference
- Calculation of overall effect
  - Fixed effects model
    - considers that variability is exclusively due to random variation, i.e. if all the studies were infinitely large they would give identical results and estimate the same treatment effect
    - More power to reject the null hypothesis
    - Justified when the test for heterogeneity is not significant
  - Random effects model
    - assumes a different underlying effect for each study and takes this into consideration as an additional source of variation
    - 95% CI wider than that of a fixed effects analysis: both inter-patient variability and inter-study variability
    - Results in more weight given to smaller studies!



Fixed effects models assume that each trial represents a random sample of a single population with a single response to treatment. Random effects models assume that the different trial results may come from different populations with varying responses to treatment.

Forest Plot



Point estimate and 95% CI of individual studies

Area proportional to study size (and relative weight in MA)

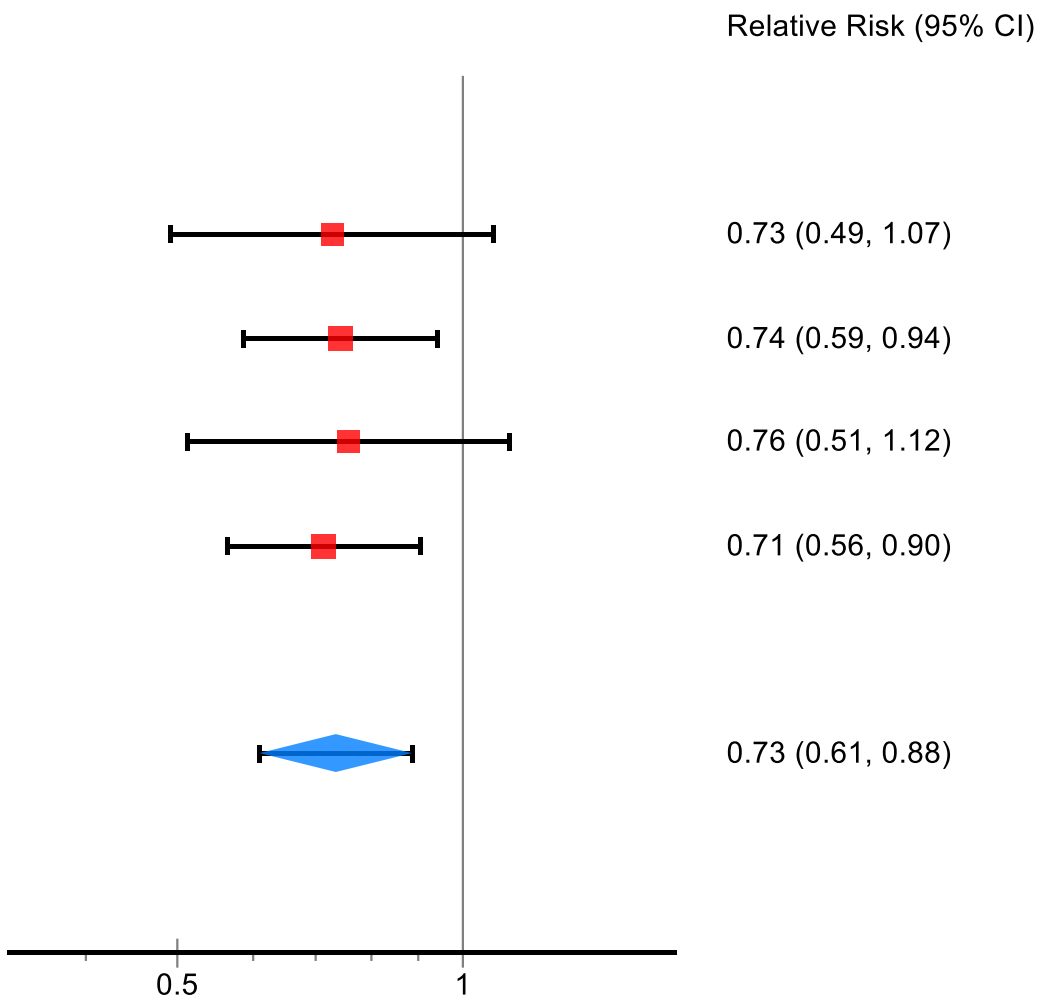
Vertical line: relative risk = 1, if CI crosses this line: result not sign. at 5% level

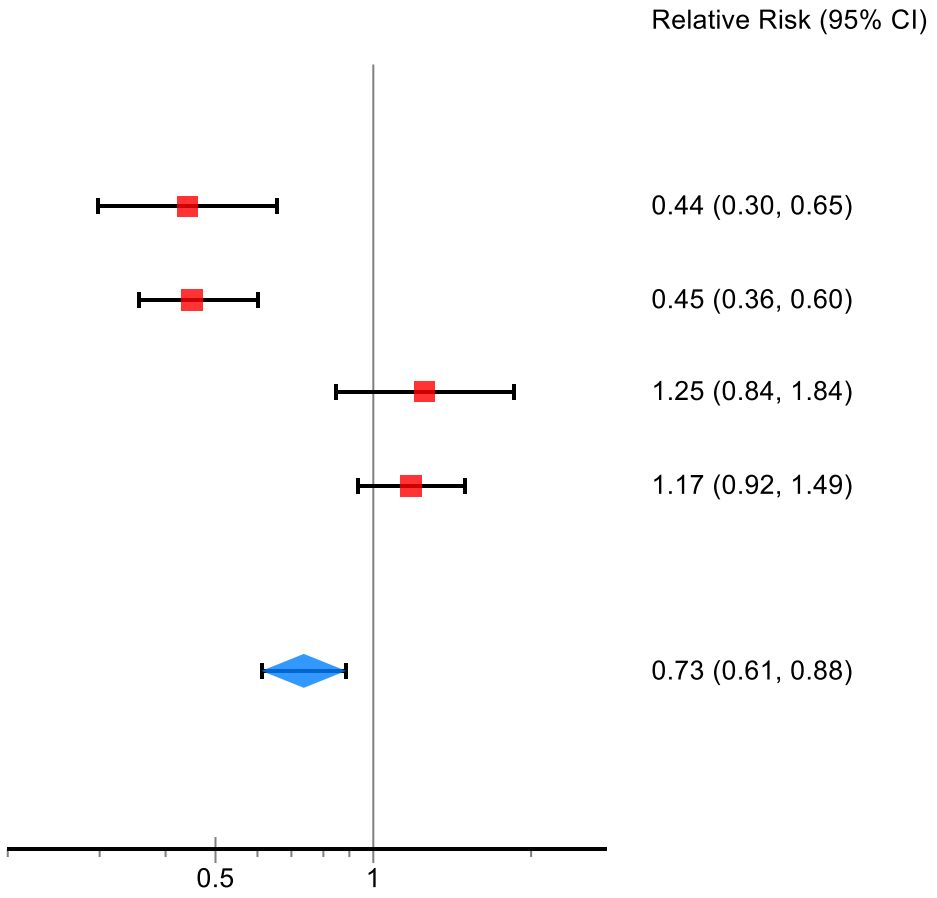
Summary statistic (pooled); width=95%CI

Meta-analysis

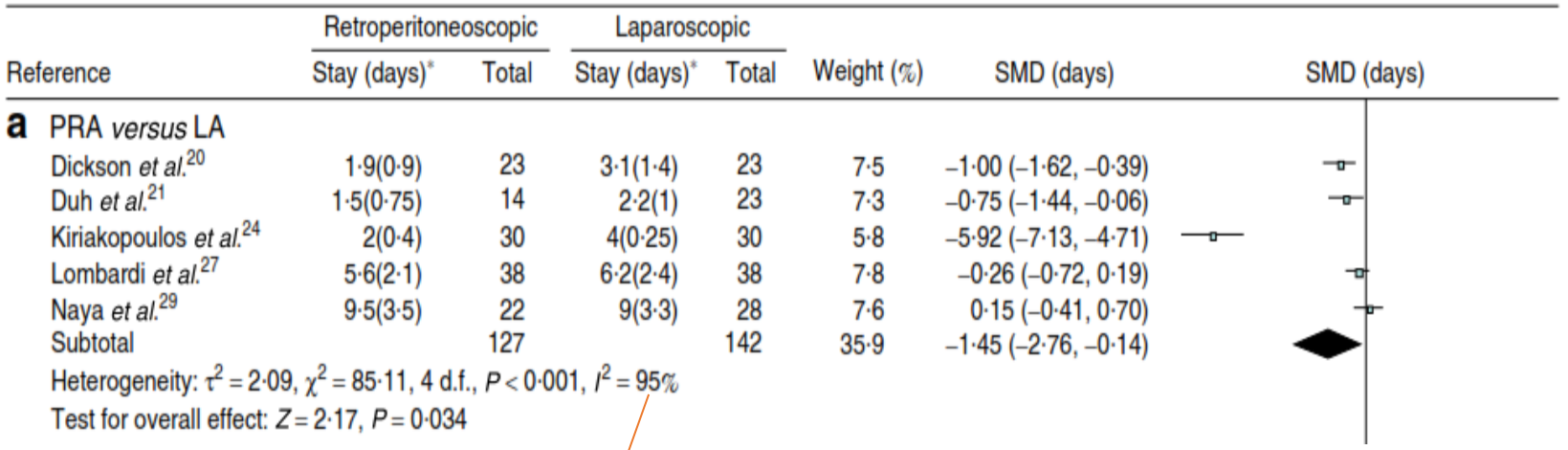
- Heterogeneity
  - Comonly used: I^2 test: [(Q-df/Q)]/ 100, where Q is the chi-square, 0 - 100%.
  - Defines percentage of variability in treatment effect estimates due to between study heterogeneity rather than chance
  - More than 40%: important
- Funnel plots: detect publication bias
  - Large studies→ precise estimates
  - Symmetrical distribution

GHENT UNIVERSITY


Relative Risk (95% CI)
0.73 (0.49, 1.07)
0.74 (0.59, 0.94)
0.76 (0.51, 1.12)
0.71 (0.56, 0.90)
0.73 (0.61, 0.88)
0.5
1



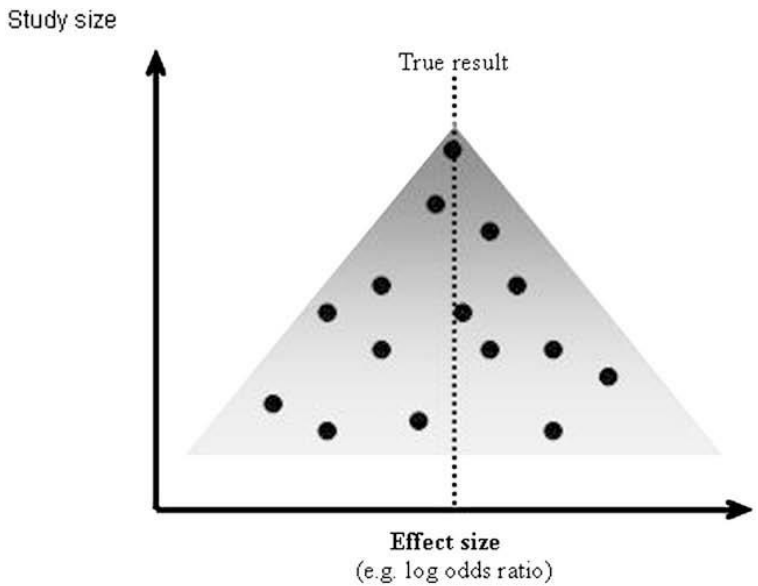
p-value for heterogeneity < 0.001
I^2=89%

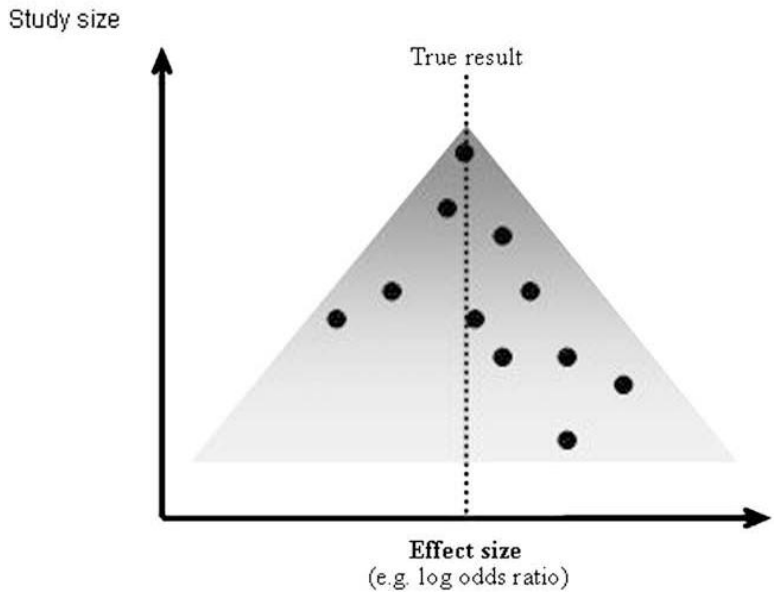


Reference	Retroperitoneoscopic Stay (days)*	Retroperitoneoscopic Total	Laparoscopic Stay (days)*	Laparoscopic Total	Weight (%)	SMD (days)
a PRA *versus* LA						
Dickson *et al.*[20]	1·9(0·9)	23	3·1(1·4)	23	7·5	−1·00 (−1·62, −0·39)
Duh *et al.*[21]	1·5(0·75)	14	2·2(1)	23	7·3	−0·75 (−1·44, −0·06)
Kiriakopoulos *et al.*[24]	2(0·4)	30	4(0·25)	30	5·8	−5·92 (−7·13, −4·71)
Lombardi *et al.*[27]	5·6(2·1)	38	6·2(2·4)	38	7·8	−0·26 (−0·72, 0·19)
Naya *et al.*[29]	9·5(3·5)	22	9(3·3)	28	7·6	0·15 (−0·41, 0·70)
Subtotal		127		142	35·9	−1·45 (−2·76, −0·14)
Heterogeneity: $\tau^2 = 2{\cdot}09$, $\chi^2 = 85{\cdot}11$, 4 d.f., $P < 0{\cdot}001$, $I^2 = 95\%$						
Test for overall effect: $Z = 2{\cdot}17$, $P = 0{\cdot}034$						

Fixed effects MA used even if I^2 = 95%!

Constantinides Br J Surg 2012


Study size
True result
Effect size
(e.g. log odds ratio)


Study size
True result
Effect size
(e.g. log odds ratio)

English Deutsch Español فارسی Français Hrvatski 日本語 한국어 Bahasa Malaysia ▼ | Media | Contact us | Community | My Account

Cochrane

Trusted evidence.
Informed decisions.
Better health.

Search...

Our evidence | About us | Join Cochrane | News and jobs | Cochrane Library ▶

Free Cochrane Lecture 2022
30 November

Latest News and Events

Cochrane Library Editorial: Cochrane reviews role in directing researchers

24 November 2022

Which medicines, taken by mouth or injected, work best to treat a skin condition called

22 November 2022

Latest Cochrane evidence | Top 10

How accurate are rapid antigen tests for diagnosing COVID-19?