THE UNIVERSITY OF TEXAS
MD Anderson Cancer Center
Making Cancer History®

CERTaIN

Methods for Comparative Effectiveness and Patient-Centered Outcomes Research

Maria E. Suarez-Almazor, MD, PhD
Barnts Family Distinguished Professor
Department of Health Services Research
University of Texas MD Anderson Cancer Center
msalmazor@mdanderson.org

Overview

- Basic methods of Comparative Effectiveness Research
- Evidence synthesis
- Observational studies
- Pragmatic clinical trials

CER: Comparative Effectiveness Research

PCOR: Patient-Centered Outcomes Research

RWE: Real-World Evidence

Key Elements of Comparative Effectiveness Research

1. Goal: to inform health care decisions for selected populations
2. Effectiveness (real-world) vs. efficacy (selected populations and controlled environment)
3. Comparisons of different alternatives

Major Methodologies in CER

- **Evidence synthesis**
  - Research on existing data

- **Generation of new data (real-world studies)**
  - Observational research (cohort and registry studies)
  - Effectiveness/pragmatic clinical trials

Evidence Synthesis

Evidence synthesis

- Research methodology that identifies, selects, appraises, combines and analyzes data from multiple sources of information

- The goal is to identify and synthesize all the available information on a topic of interest

- Systematic reviews
- Meta-analysis
- Decision analysis

Systematic Review

- A summary of research results (evidence) that uses explicit and reproducible methods to systematically search, critically appraise, and synthesize on a specific issue

- It synthesizes the results of multiple primary studies related to each other by using strategies that reduce biases and errors

https://guides.temple.edu/c.php?g=78618&p=9548660

A Systematic Review is Different from a Traditional Narrative Review

- Clearly stated objectives with well-defined question (PICOTS)
- *A priori* protocol (needs to be registered – PROSPERO, etc)
- Pre-defined eligibility criteria for studies
- Explicit, reproducible methodology
- Systematic search that attempts to identify all studies that would meet the eligibility criteria
- Assessment of the validity of the findings of the included studies (assessment of risk of bias)
- Systematic synthesis and reporting of the characteristics and findings of the included studies

Meta-analysis

- Meta-analysis is the use of statistical methods to summarize the results of independent studies - **quantitative**
- Many but not all systematic reviews contain meta-analyses
- By combining information from all relevant studies, meta-analyses can provide more precise estimates than those derived from individual studies
- Evaluate consistency of evidence vs. differences across studies

Steps in Systematic Reviews

1. **Formulate question**
2. **Develop protocol(e.g. PROSPERO)**
3. **Search strategy**
   - Librarian
   - Can use more than one database: Medline, EMBASE
   - Needs to be broad to increase sensitivity
4. **Study selection**
   - Explicit criteria (inclusion and exclusion)
   - Type of study, interventions, participants, outcomes, length of follow-up
   - 2 independent reviewers select studies. Disagreement resolved by consensus or by arbitrator
5. **Data extraction**
   - 2 reviewers or 1 reviewer cross-checked by second

Steps

6. Evaluate risk of bias of studies

- Different tools according to type of study: clinical trial, cohort study, etc

7. Synthesize results

- Qualitative
- Quantitative: meta-analysis if studies are similar enough

8. Assess quality of the evidence

- Robustness of findings

9. Present results

- PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses)

How to Formulate a Good Research Question

- Clear and understandable
- Focused
- Original and novel (new information)
- Relevant
    - change current ideas or practice
    - lead to new research
- Feasible (time and resources)

PICOTS Framework

- **P** Population (who)
- **I** Intervention or issue (what are we examining)
- **C** Comparator (compared to)
- **O** Outcome (what happens)
- **T** Timing (when)
- **S** Setting (where)

PRISMA flowchart

Ruiz JI, Lopez-Olivo MA, Geng Y, Suarez-Almazor ME. COVID-19 Outcomes in Patients with Cancer Receiving Immune Checkpoint Inhibitors: A Systematic Review. J Immunother Precis Oncol. 2023 Feb 22;6(2):103-110. doi: 10.36401/JIPO-22-24. PMID: 37214207; PMCID: PMC10195019.

25 studies
21 cohort studies
1 case series
3 case reports

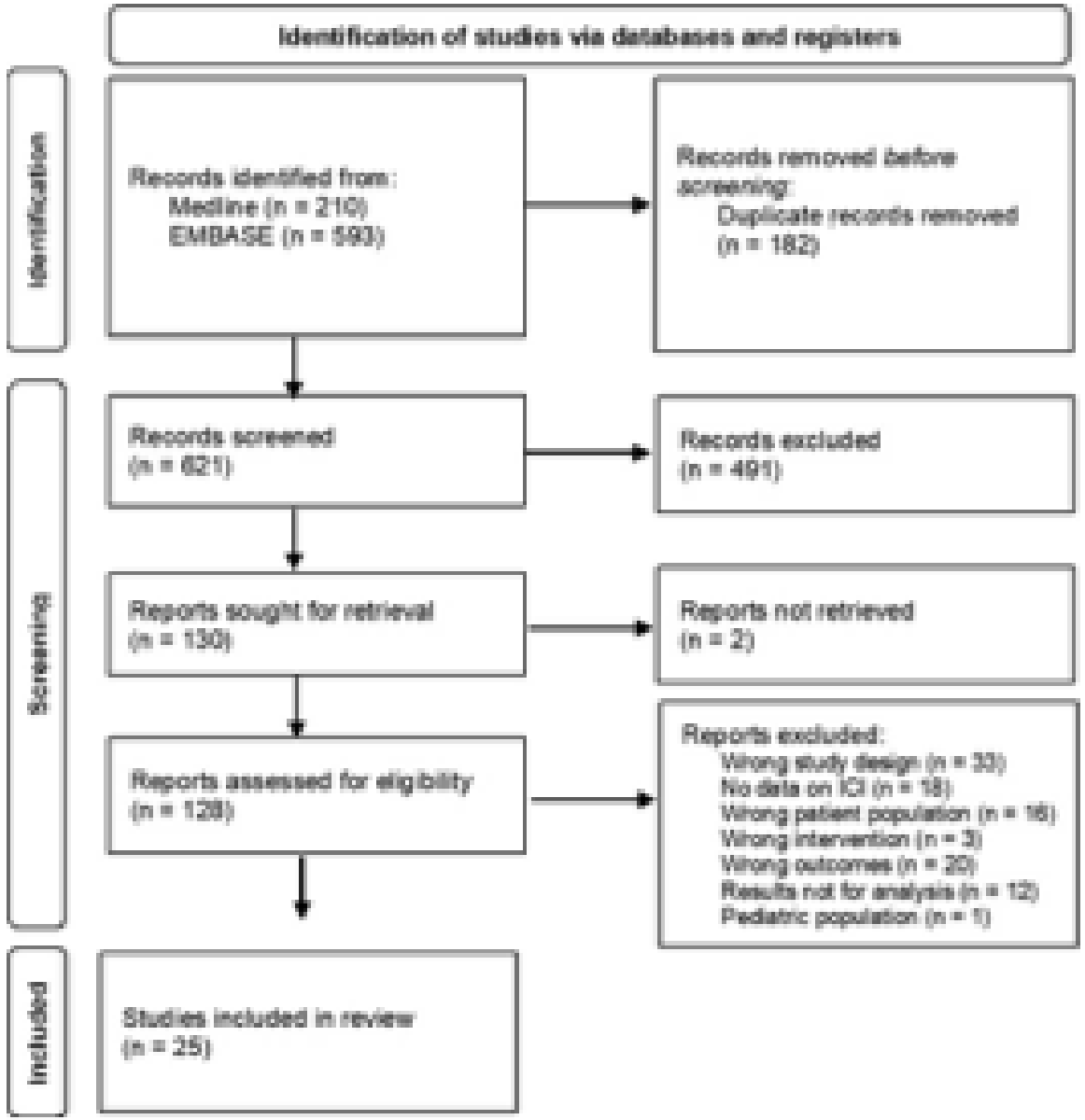


Figure 1. PRISMA flow diagram of study selection.

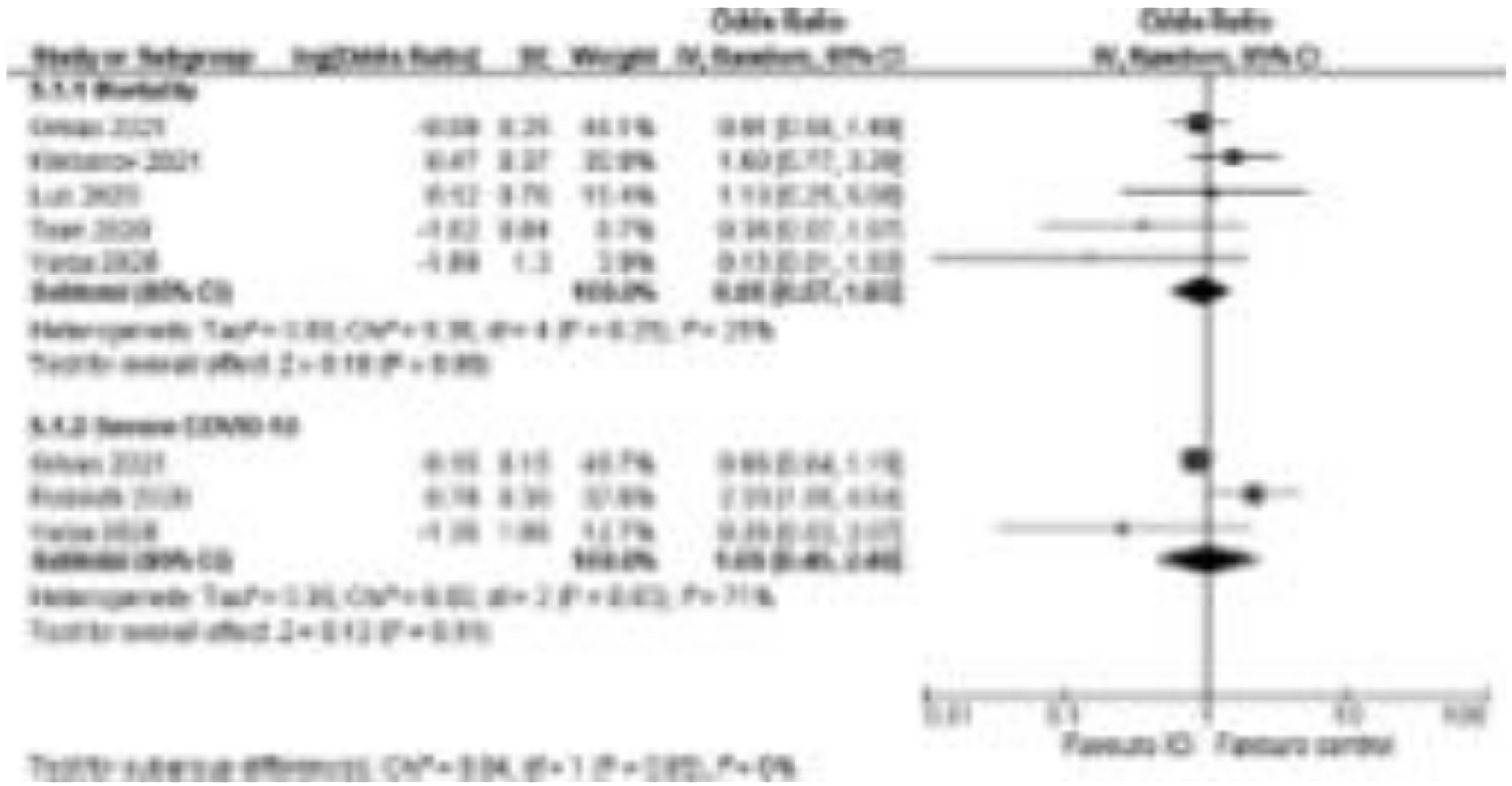


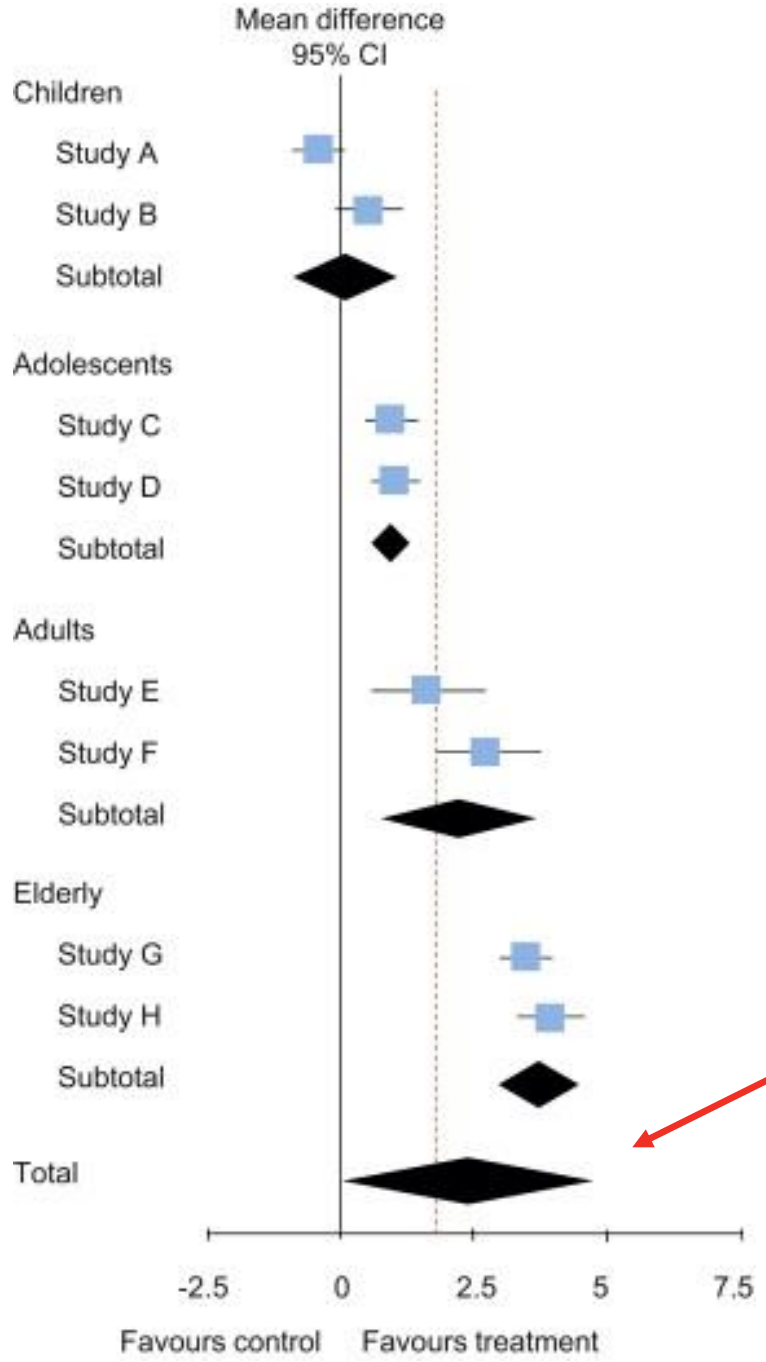
Figure 2. Forest plot comparing patients with cancer treated with ICIs versus those not treated with ICI. Adjusted ORs. ICI: immune checkpoint inhibitor; OR: odds ratio.

Heterogeneity

Hypothetical meta-analysis of trials, using a random-effects model, with subgrouping by age categories.

Dotted vertical line represents the threshold for clinical relevance.

Test for heterogeneity:
$I^2 = 67\%$, $p = 0.04$.

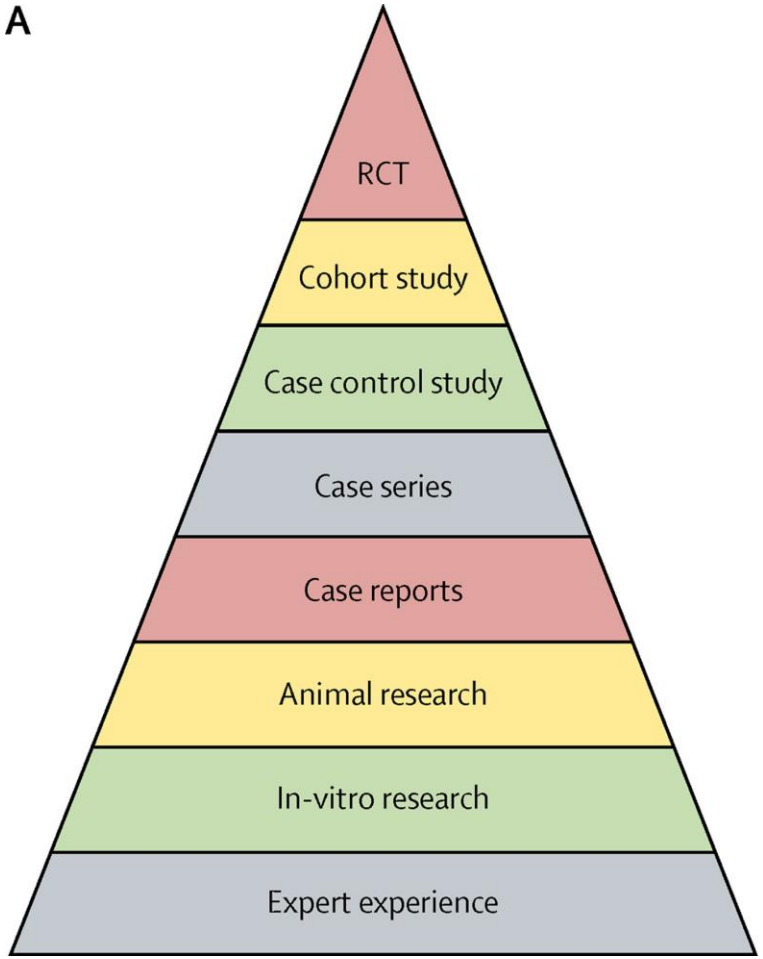


Not appropriate

Verhagen AP , Ferreira ML. Forest Plots.
J Physiotherapy. 2016
https://doi.org/10.1016/j.jphys.2014.06.021

Evidence-Based Medicine (EBM)
Hierarchy of Evidence from Studies

A



B

Quality of evidence	Study design	Lower quality if*	Higher quality if†
High	Randomised trial	Study limitations - 1 serious - 2 very serious Inconsistency - 1 serious - 2 very serious Indirectness - 1 serious - 2 very serious Imprecision - 1 serious - 2 very serious Publication bias - 1 likely - 2 very likely	Large effect + 1 large + 2 very large Dose response + 1 evidence of a gradient All plausible confounders + Would reduce a demonstrated effect or + Would suggest a spurious effect when results show no effect
Moderate			
Low	Observational study		
Very low			

Initial Pyramid

GRADE

Djulbegovic B, Guyatt GH. Progress in evidence-based medicine: a quarter century on. Lancet. 2017 PMID: 28215660.

HOW DO WE EXPRESS HOW CERTAIN WE ARE IN THE RESULTS?
GRADE: Four levels of certainty

Certainty level		Grade
High	We are very confident that the true effect lies close to that of the estimate of the effect	⨁⨁⨁⨁
Moderate	We are moderately confident in the effect estimate: The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different	⨁⨁⨁⊖
Low	Our confidence in the effect estimate is limited: The true effect may be substantially different from the estimate of the effect	⨁⨁⊖⊖
Very low	We have very little confidence in the effect estimate: The true effect is likely to be substantially different from the estimate of effect	⨁⊖⊖⊖

Summary of Results

Table 2. COVID-19 outcomes in patients with cancer receiving ICIs compared to those not receiving ICI Therapy

Outcomes	Anticipated Absolute Effects: Risk With No ICI	Anticipated Absolute Effects: Risk With ICI (95% CI)	Relative Effect, OR (95% CI)	No. of Studies	Certainty of the Evidence (GRADE)	Comments
Mortality	100 per 1000	95 per 1000 (60–151)	0.95 (0.57–1.60)[a]	5	Very low[b,d]	There is uncertainty based on the quality of evidence if the risk of COVID-19 mortality is higher in patients with cancer exposed to ICI compared with those not exposed to ICI.
Hospital admission	300 per 1000	464 per 1000 (291–647)	2.02 (0.96– 4.27)[a]	2	Very low[d]	There is uncertainty based on the quality of evidence if the risk of COVID-19 hospital admission is higher in patients with cancer exposed to ICI compared with those not exposed to ICI.
Severe COVID-19	120 per 1000	125 per 1000 (58–251)	1.05 (0.45–2.46)[a]	3	Very low[c,d]	There is uncertainty based on the quality of evidence if the risk of severe COVID-19 is higher in patients with cancer exposed to ICI compared with those not exposed to ICI.

[a]Adjusted OR.
[b]There is high risk of bias assessed with the Newcastle-Ottawa Scale specifically in the selection and outcome domains.
[c]There is heterogeneity not explained due to chance.
[d]The true effect can benefit either the experimental or the control group.
CI: confidence interval; GRADE: Grading of Recommendations, Assessment, Development and Evaluations; ICI: immune checkpoint inhibitor; OR: odds ratio.

Initial rating depends on study design for body of evidence

BODY OF EVIDENCE FROM RCT

4 – high

3 – moderate

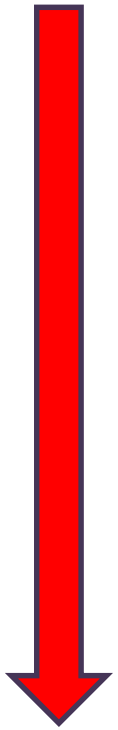
2 – low

1 – very low

BODY OF EVIDENCE FROM NON – RANDOMIZED STUDIES

Certainty of evidence: rating down

- risk of bias
- imprecision
- inconsistency
- indirectness
- publication bias



4 – high (RCT)

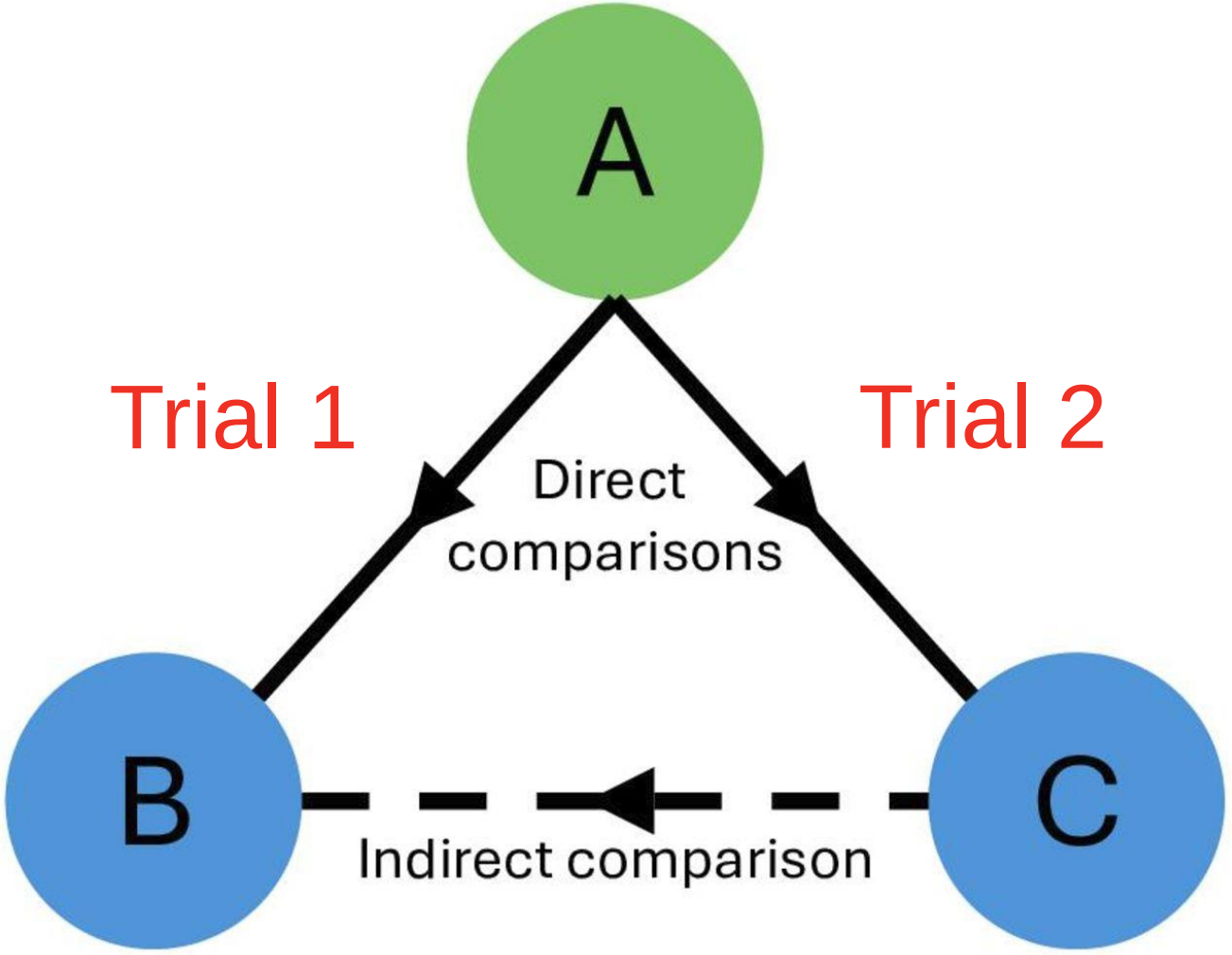
3 – moderate

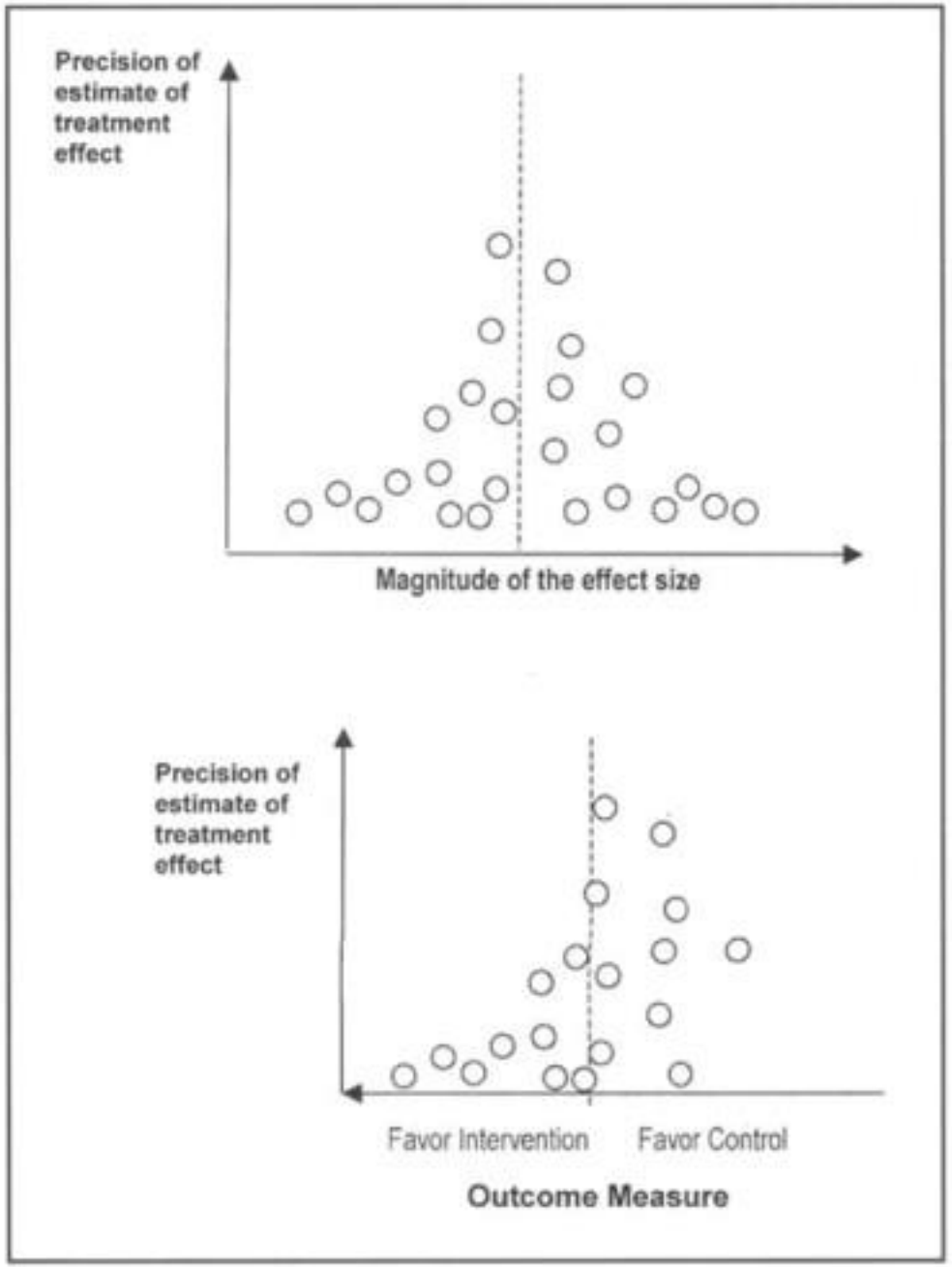
2 – low

1 – very low

- rating down 1 level for serious concerns
- rating down 2 levels for very serious concerns
- some (minor) concerns on more than one domain can amount to one full level rated down

Indirect comparison





Publication bias: funnel plots

Montori, Victor M. et al. Publication Bias: A Brief Review for Clinicians. Mayo Clinic Proceedings 2000: 75, 1284 - 1288

Decision Analysis Model

- **Quantitative approach that uses mathematical modeling to aid in decision-making by providing probabilities of events under different scenarios**
- **Often used in cost-effectiveness evaluations to establish value**
  - How much it costs to attain a certain outcome
- **Can use data from different sources**
- **Different methodologies**
  - Simple model is a decision tree

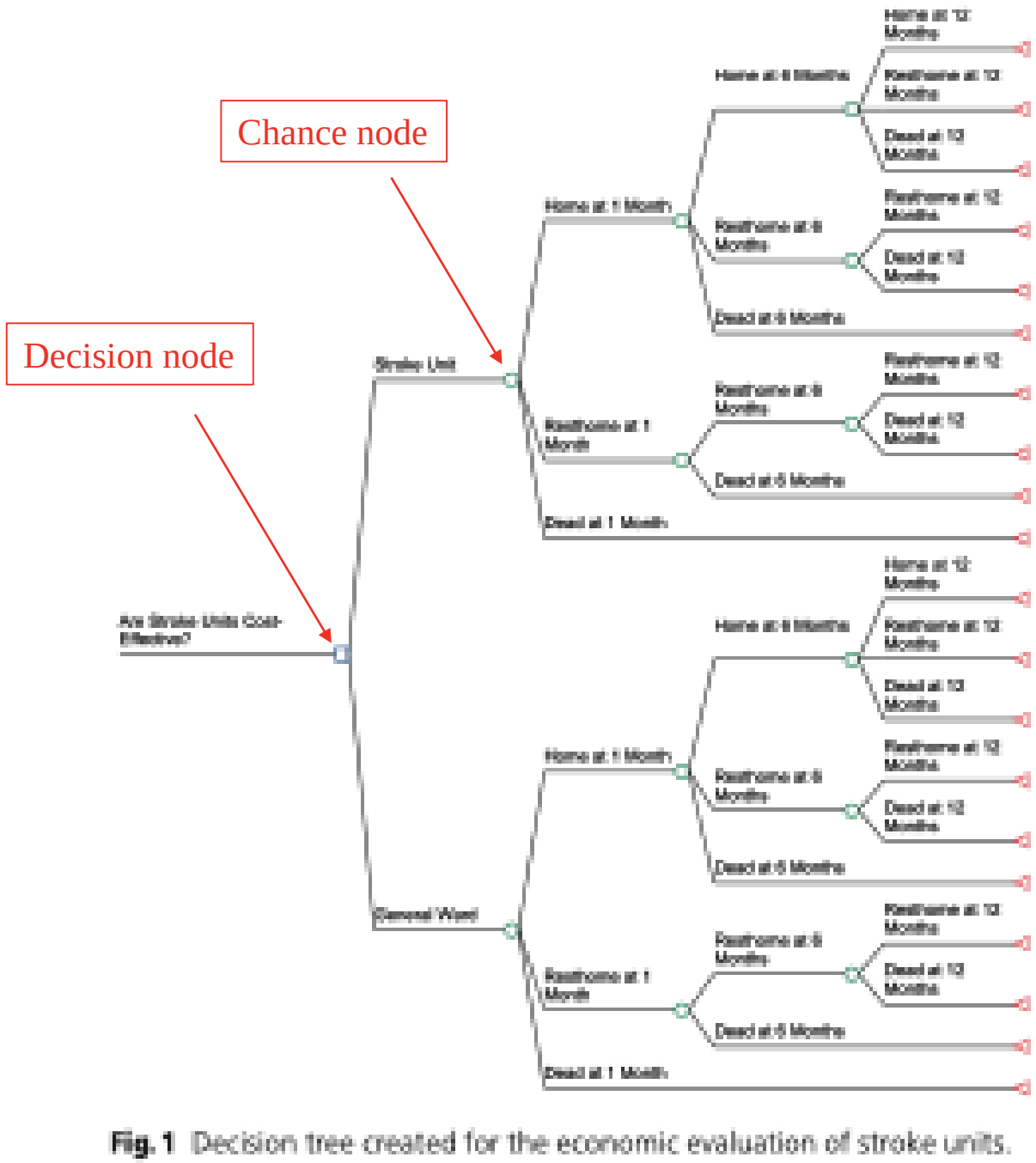


Fig. 1 Decision tree created for the economic evaluation of stroke units.

Te Ao BJ, Brown PM, Feigin VL, Anderson CS. **Are stroke units cost effective?** Evidence from a New Zealand stroke incidence and population-based study. Int J Stroke. 2012 Dec;7(8):623-30. doi: 10.1111/j.1747-4949.2011.00632.x. Epub 2011 Oct 20. PMID: 22010968.

Observational Studies

Observational Studies vs Clinical Trials

Observational Studies

Research studies in which researchers collect information from participants or look at data that was already collected

- Cross-sectional
- Case-control (always retrospective)
- Cohort studies (prospective or retrospective)

Clinical Trials (always experimental)

Research studies in which researchers test health-related interventions by assigning participants to receive intervention(s)

- Explanatory (efficacy) → ideal conditions
- Pragmatic (effectiveness) → real-world

Sources of Data for Observational Studies in CER

- **Large prospective population-based cohorts**
- **Retrospective**
  - Existing cohort studies
  - Registries (specific populations, health interventions)
  - Claims data (Medicare, MarketScan, IQVIA, etc)
  - Electronic Health Records (EHR)

RWE: Real-World Evidence

- **Real-world data (RWD)** are data relating to patient health status and/or the delivery of health care routinely collected from a variety of sources.
- **Examples of RWD** include data derived from electronic health records, medical claims data, data from product or disease registries, and data gathered from other sources (such as digital health technologies).
- **Real-world evidence (RWE)** is the clinical derived from analysis of RWD.

https://www.fda.gov/science-research/science-and-research-special-topics/real-world-evidence

Uses and Advantages of Observational Studies

- Real-world
- Large sample sizes
- Multiple interventions can be compared
- Longer follow-ups
- Often needed for safety studies/rare outcomes (e.g. mortality)
- Readily available data to answer research question (retrospective)
- Less expensive

Major disadvantage: subject to bias and confounding

Observational Studies in CER

- **Hypothesis-testing vs. descriptive**
- **Broad population**
- **Explicit protocol**
  - Easy to go on fishing expedition
- **Use methodologies that control confounding and bias**
  - Try to approach what a RCT would be

JAMA Network Open

Original Investigation | Neurology

Telestroke and Timely Treatment and Outcomes in Patients With Acute Ischemic Stroke

Brian Stamm, MD, MSc; Rachael T. Whitney, PhD; Regina Royan, MD, MPH; Ghada Ibrahim, MS; Adrienne V. Nickles, MPH; Rebecca A. Ferber, MPH; Wen Ye, PhD; Wan-Ling Hsu, PhD; Nikita Chhabra, DO; Rodney A. Hayward, MD; Mollie McDermott, MD, MS; Phillip A. Scott, MD; Kevin N. Sheth, MD; Mathew J. Reeves, BVSc, PhD; Deborah A. Levine, MD, MPH

Abstract

IMPORTANCE Telestroke has the potential to revolutionize acute stroke treatment by improving access to optimal stroke care, including time-sensitive care such as thrombolysis. However, it is unclear how treatment times and stroke outcomes compare between patients evaluated and not evaluated by telestroke.

OBJECTIVE To evaluate the association between telestroke use and acute stroke treatment times and outcomes.

DESIGN, SETTING, AND PARTICIPANTS This retrospective cohort study included patients with acute ischemic stroke aged 18 years or older presenting to 42 Paul Coverdell Michigan Stroke Registry hospitals from January 1, 2022, to December 31, 2023. All patients were potentially eligible for thrombolysis (ie, presented ≤4 hours of last known well, no contraindications to thrombolysis documented).

EXPOSURE Telestroke (vs nontelestroke) encounter.

MAIN OUTCOMES AND MEASURES The primary outcomes were receipt of thrombolysis and door-to-needle (DTN) time as a continuous variable and a categorical variable (≤60 vs >60 minutes). Secondary outcomes included discharge ambulatory status, discharge destination, and door-in-door-out (DIDO) time in transferred patients. Multivariable hierarchical models evaluated associations between telestroke (vs nontelestroke) activation and outcomes, sequentially adjusting for demographics, medical history, presentation or arrival, and hospital characteristics.

RESULTS Among the 3036 patients with acute ischemic stroke potentially eligible for thrombolysis (mean [SD] age, 69.7 [14.5] years; 1563 male [51.5%]), 785 (25.9%) were evaluated using telestroke and 2251 (74.1%) without telestroke. A total of 1673 patients (55.1%) were treated with thrombolysis. In the fully adjusted models, patients evaluated by telestroke had a significantly higher odds of receiving thrombolysis (adjusted odds ratio, 1.61; 95% CI, 1.17-2.23) but longer DTN times (6.55 minutes longer; 95% CI, 2.12-10.97 minutes longer) and lower odds of meeting a guideline-concordant DTN time within 60 minutes (adjusted odds ratio, 0.56; 95% CI, 0.39-0.81) compared with those not evaluated by telestroke. Among 255 patients who underwent interhospital transfer, 207 (81.2%) received thrombolysis, and patients with telestroke had significantly longer DIDO times (46.90 minutes longer; 95% CI, 1.08-92.72 minutes longer).

CONCLUSIONS AND RELEVANCE In this cohort study of patients with acute ischemic stroke potentially eligible for thrombolysis, those evaluated by telestroke had a 61% higher odds of receiving thrombolysis but a 44% lower odds of meeting guideline-concordant DTN times within 60

Key Points

Question How do treatment times and stroke outcomes in patients with acute ischemic stroke evaluated by telestroke compare with those not evaluated by telestroke?

Findings In this cohort study of 3036 patients with acute ischemic stroke potentially eligible for thrombolysis, telestroke was associated with a higher odds of receiving thrombolysis, but significantly prolonged door-to-needle and door-in-door-out times and a lower odds of meeting guideline-concordant door-to-needle times within 60 minutes, compared with nontelestroke.

Meaning These findings suggest that there is room to improve timely stroke treatment for patients evaluated by telestroke to ensure that all patients with ischemic stroke receive guideline-concordant, time-sensitive care.

+ Invited Commentary

+ Supplemental content

Author affiliations and article information are listed at the end of this article.

PCORI Methodology Standards for Use of Registries

Registries must have the following characteristics

1. **Documentation registry purpose and protocol**
2. **Data safety and security**
3. **Data elements and quality**
   - Standardized data elements
   - Quality assurance plan
4. **Availability of data to control for confounding**
5. **Systematic participant enrollment**
6. **Explicit participant follow-up**
7. **Clear documentation of registry materials and protocols**

JAMA Network Open

Original Investigation | Public Health

Comparative Risks of Potential Adverse Events Following COVID-19 mRNA Vaccination Among Older US Adults

Daniel A. Harris, PhD; Kaleen N. Hayes, PhD; Andrew R. Zullo, PhD; Vincent Mor, PhD; Preeti Chachlani, MA; Yalin Deng, PharmD; Ellen P. McCarthy, PhD; Djeneba Audrey Djibo, PhD; Cheryl N. McMahill-Walraven, PhD; Stefan Gravenstein, MD

Abstract

IMPORTANCE Head-to-head safety comparisons of the mRNA vaccines for SARS-CoV-2 are needed for decision making; however, current evidence generalizes poorly to older adults, lacks sufficient adjustment, and inadequately captures events shortly after vaccination. Additionally, no studies to date have explored potential variation in comparative vaccine safety across subgroups with frailty or an increased risk of adverse events, information that would be useful for tailoring clinical decisions.

OBJECTIVE To compare the risk of adverse events between mRNA vaccines for COVID-19 (mRNA-1273 and BNT162b2) overall, by frailty level, and by prior history of the adverse events of interest.

DESIGN, SETTING, AND PARTICIPANTS This retrospective cohort study was conducted between December 11, 2020, and July 11, 2021, with 28 days of follow-up following the week of vaccination. A novel linked database of community pharmacy and Medicare claims data was used, representing more than 50% of the US Medicare population. Community-dwelling, fee-for-service beneficiaries aged 66 years or older who received mRNA-1273 vs BNT162b2 as their first COVID-19 vaccine were identified. Data analysis began on October 18, 2022.

EXPOSURE Dose 1 of mRNA-1273 vs BNT162b2 vaccine.

MAIN OUTCOMES AND MEASURES Twelve potential adverse events (eg, pulmonary embolism, thrombocytopenia purpura, and myocarditis) were assessed individually. Frailty was measured using a claims-based frailty index, with beneficiaries being categorized as nonfrail, prefrail, and frail. The risk of diagnosed COVID-19 was assessed as a secondary outcome. Generalized linear models estimated covariate-adjusted risk ratios (RRs) and risk differences (RDs) with 95% CIs.

RESULTS This study included 6 388 196 eligible individuals who received the mRNA-1273 or BNT162b2 vaccine. Their mean (SD) age was 76.3 (7.5) years, 59.4% were women, and 86.5% were White. A total of 38.1% of individuals were categorized as prefrail and 6.0% as frail. The risk of all outcomes was low in both vaccine groups. In adjusted models, the mRNA-1273 vaccine was associated with a lower risk of pulmonary embolism (RR, 0.96 [95% CI, 0.93-1.00]; RD, 9 [95% CI, 1-16] events per 100 000 persons) and other adverse events in subgroup analyses (eg, 11.0% lower risk of thrombocytopenia purpura among individuals categorized as nonfrail). The mRNA-1273 vaccine was also associated with a lower risk of diagnosed COVID-19 (RR, 0.86 [95% CI, 0.83-0.87]), a benefit that was attenuated by frailty level (frail: RR, 0.94 [95% CI, 0.89-0.99]).

CONCLUSIONS AND RELEVANCE In this cohort study of older US adults, the mRNA-1273 vaccine was associated with a slightly lower risk of several adverse events compared with BNT162b2, possibly due to greater protection against COVID-19. Future research should seek to formally disentangle

Key Points

Question Are there safety differences between mRNA vaccines for COVID-19, and do those differences vary by frailty level?

Findings In this cohort study of 6 388 196 older US adults, a 4% lower risk of pulmonary embolism, a 2% lower risk of thromboembolic events, and a 14% lower risk of diagnosed COVID-19 were observed among those who received the mRNA-1273 vaccine compared with the BNT162b2 vaccine. Although both vaccines were safe across frailty subgroups, differences were generally greater in individuals without frailty.

Meaning These findings suggest that compared with BNT162b2, mRNA-1273 was associated with a lower risk of adverse events, possibly due to improved protection against COVID-19.

+ Supplemental content

Author affiliations and article information are listed at the end of this article.

Pragmatic Clinical Trials

Clinical Trials (always experimental)

Research studies in which researchers test health-related interventions by assigning participants to receive intervention(s)

- Explanatory (efficacy) — ideal conditions
- Pragmatic (effectiveness) — real-world

Pragmatic Clinical Trials

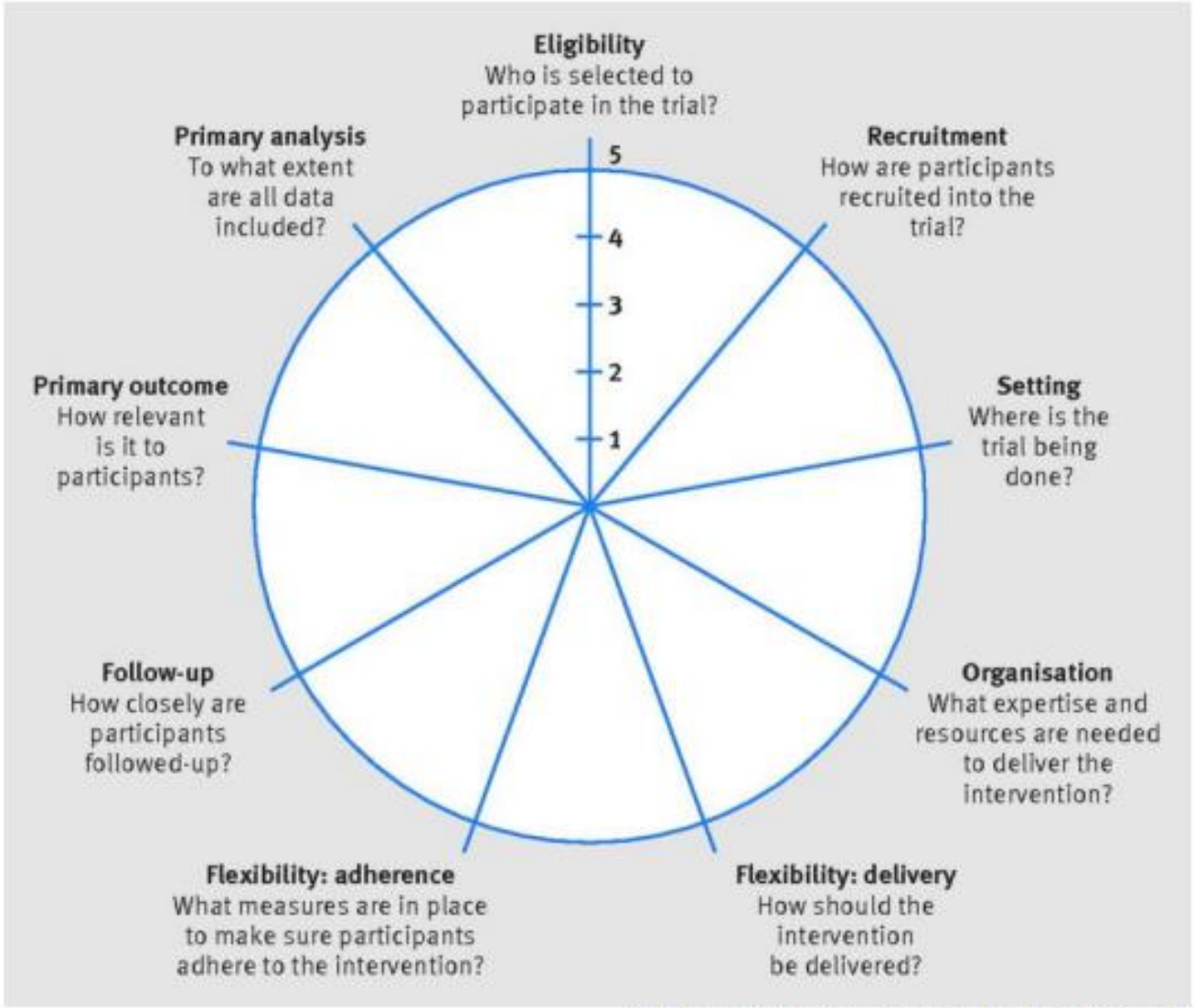
- Designed to study a health intervention in a real-world setting that is similar or identical to the one in which the intervention will be implemented
- In contrast to explanatory or traditional trials, which are designed to demonstrate the safety and efficacy of an intervention
  - under highly controlled conditions
  - in carefully selected groups of participants
- Most clinical trials are situated somewhere along the spectrum between pragmatic and explanatory

https://rethinkingclinicaltrials.org/archived-what-is-a-pragmatic-clinical-trial/pragmatic-elements-an-introduction-to-precis-2/

Explanatory vs Pragmatic Trials

	Explanatory	Pragmatic
Population	Highly selected	Little selection - Broad
Interventions	Strictly applied	Flexibility in co-interventions, dosage, etc
Comparator	Placebo or another intervention	Usual care
Outcome measures	Sometimes surrogate or clinical only	Patient-centered
Data collection	Additional procedures outside clinical care	In clinical settings

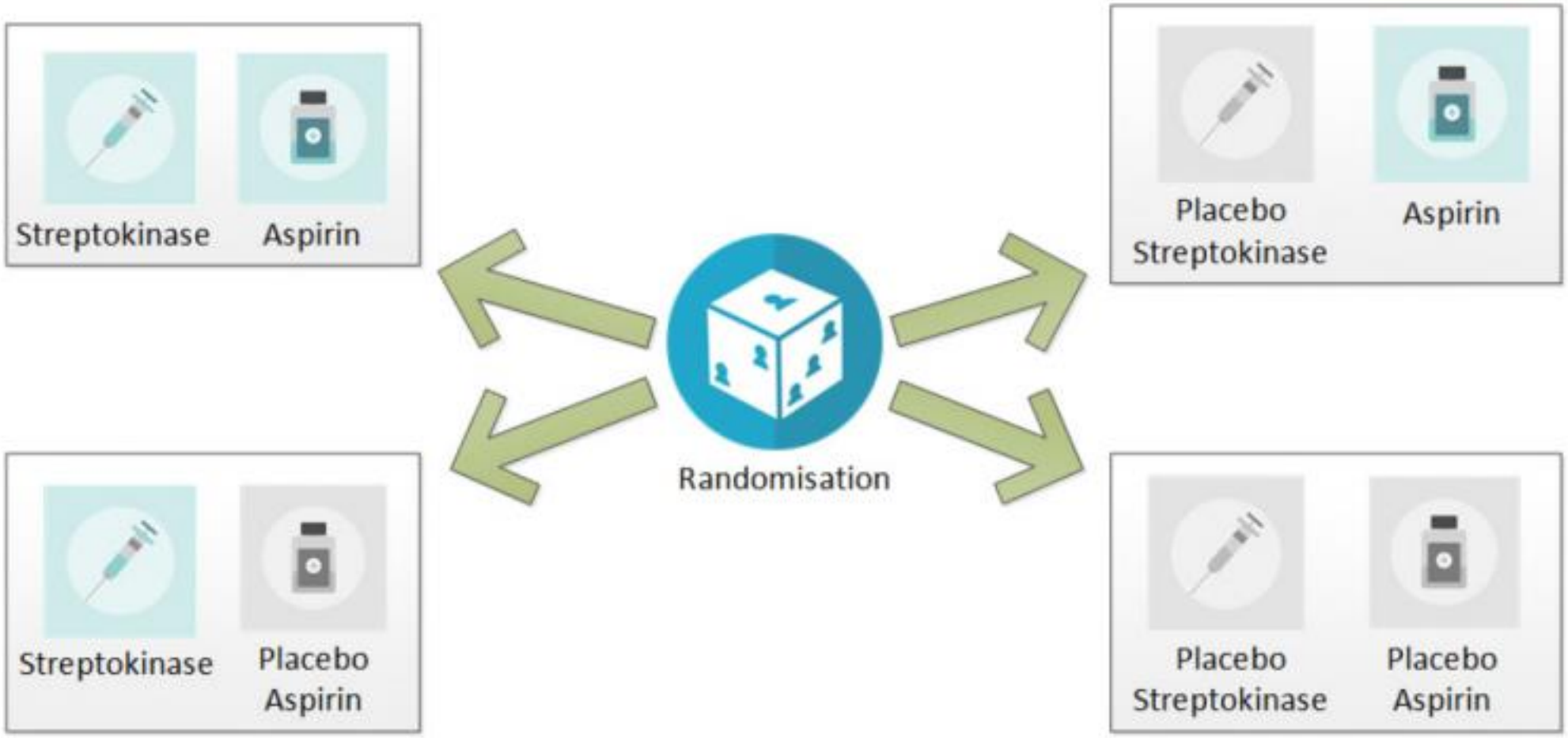
The Pragmatic-Explanatory Continuum Indicator Summary 2 (PRECIS-2)



Kirsty Loudon et al. BMJ 2015;350:bmj.h2147

1. Very explanatory
2. Rather explanatory
3. Equally pragmatic and explanatory
4. Rather pragmatic
5. Very pragmatic

ISIS-2 trial – 17,187 cases of suspected acute myocardial infarction (Lancet 1988)



ACP- PEACE. Promoting Effective and Aligned Communication in the Elderly

JAMA Network | Open

Original Investigation | Geriatrics

An Intervention to Increase Advance Care Planning Among Older Adults With Advanced Cancer

A Randomized Clinical Trial

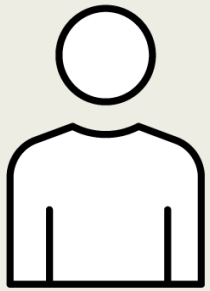
Angelo E. Volandes, MD, MPH; Yuchiao Chang, PhD; Joshua R. Lakin, MD; Michael K. Paasche-Orlow, MD; Charlotta Lindvall, MD, PhD; Seth N. Zupanc, BA; Diana Martins-Welch, MD; Maria T. Carney, MD; Edith A. Burns, MD; Jennifer Itty, MPH; Kaitlin Emmert-Tangredi, MSW; Narda J. Martin, MSN-ED, RN; Shreya Sanghani, MS; Jon Tilburt, MD; Kathryn I. Pollak, PhD; Aretha Delight Davis, MD, JD; Cynthia Garde, MBA; Michael J. Barry, MD; Areej El-Jawahri, MD; Lisa Quintiliani, PhD; Kate Sciacca, NP; Julie Goldman, MPH; James A. Tulsky, MD

JAMA Network Open

RCT: Advance Care Planning Among Older Adults With Advanced Cancer

POPULATION

15 344 Men, 14 013 Women



Adults with advanced cancer

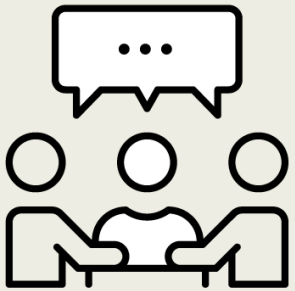
Mean (SD) age, 74.5 (6.6) y

SETTINGS / LOCATIONS

29 Cancer clinics in the US

INTERVENTION

29 537 Patients randomized



15 754 ACP video and communication skills training
Advanced care planning (ACP) video decision aids and clinician communication skills training

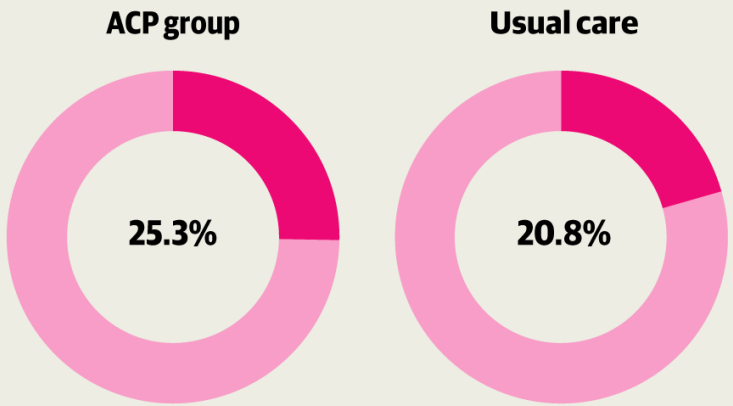
13 603 Control
Usual care

PRIMARY OUTCOME

ACP documentation frequency was a composite outcome of documented goals-of-care discussion, palliative care, hospice, and limitation of life-sustaining treatments in the electronic health record

FINDINGS

ACP documentation increased more in the ACP group than the usual care group



Documentation frequency:
ACP: 3980 of 15 754 (25.3%)
Usual care: 2834 of 13 603 (20.8%)
Adjusted difference: 6.8% (95% CI, 2.8%-10.8%; *P* < .001)

Volandes AE, Chang Y, Lakin JR, et al. An intervention to increase advance care planning among older adults with advanced cancer: a randomized clinical trial. *JAMA Netw Open*. 2025;8(5):e259150. doi:10.1001/jamanetworkopen.2025.9150

“Far better an approximate answer to the right question, which is often vague, than an exact answer to the wrong question, which can always be made precise.”

John Tukey