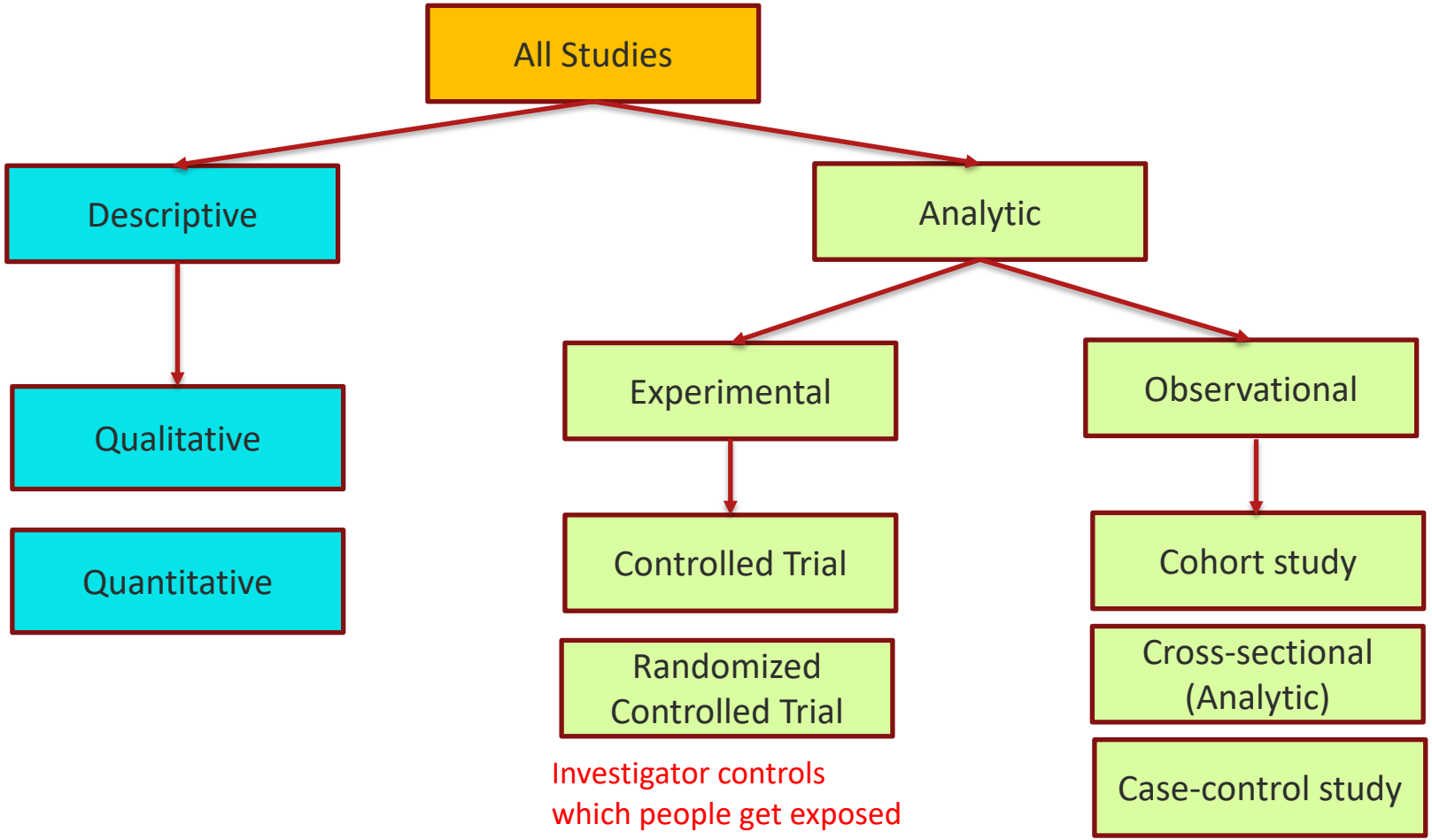
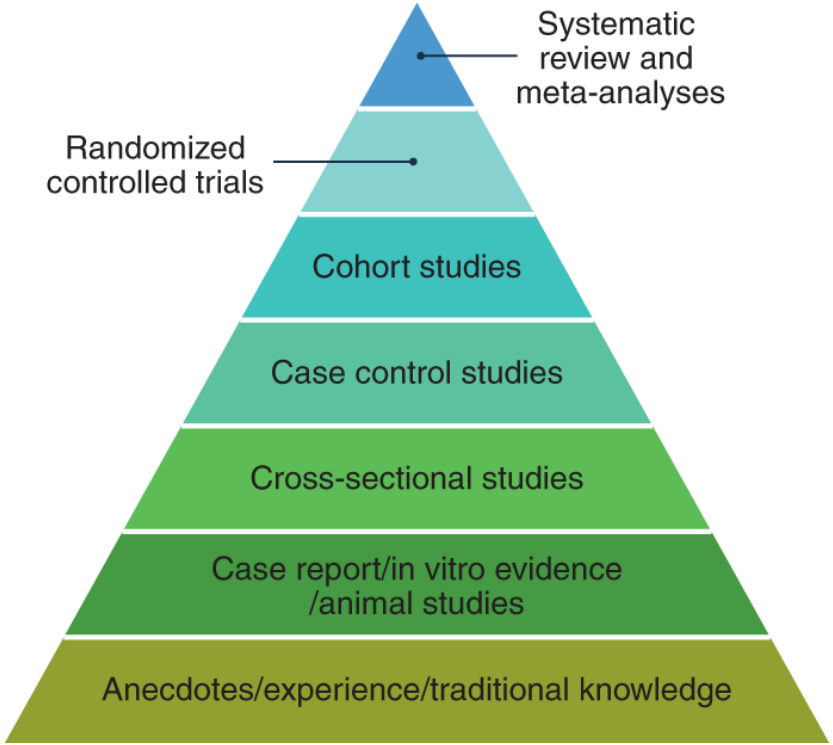
Evidence-Based Decision Making In Healthcare
Randomized Controlled Trials

Jay K. Varma, MD
https://drjayvarma.com


All Studies
Descriptive
Analytic
Qualitative
Quantitative
Experimental
Observational
Controlled Trial
Randomized Controlled Trial
Cohort study
Cross-sectional (Analytic)
Case-control study
Investigator controls which people get exposed
Investigator has no control over which people get exposed

Hierarchy of Evidence



https://www.nature.com/articles/s43016-021-00388-5/figures/1

Randomized Clinical Trials

Clinically relevant question

Greatest impact if limited information or high variability in care or outcomes

Can be answered by RCT

Feasible to perform at your center

Systematic review

Identify available information

Justify importance of question

Help design study

Randomized Clinical Trials

- Define key elements of study
  - Population
  - Intervention
  - Comparator
  - Outcome
- State primary hypothesis
  - Expected result for primary outcome in population

STUDY DESIGN

Randomization

- Generate sequence of allocation
  - Computer generated random numbers table
  - Randomize in blocks
- Concealed allocation
  - When obtaining informed consent to enroll a patient, investigator does not know if patient will get new treatment or control
- Non-manipulable allocation schedule, e.g. off-site

Concealed Allocation

Concealed Allocation

Stratification

- To assure baseline factors that impact study outcome equally distributed in study groups
- Especially useful in smaller trials
- Choose factors that have greatest impact on primary outcome
  - Aspirin use in MI study
  - Separate randomization schedules for patients with and without high blood pressure

Randomization “Blocks”

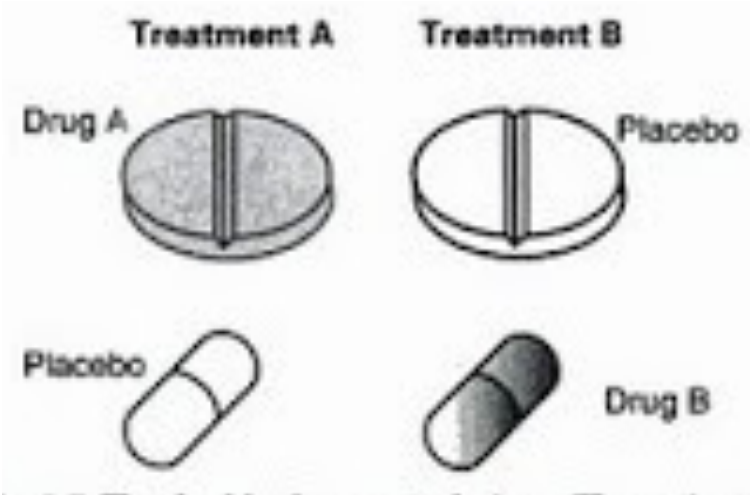
- Assures equal number in each study arm for every successive block of patients enrolled
- Prevents unequal numbers in study arms
- Prevents differences in distribution over time
  - eg, study intervention mostly early, comparator later
- Disadvantage
  - If block size figured out, next allocation may be predictable (unconcealed)

Blinding

- Knowing patient received placebo or drug may alter:
  - Whether patient gets another therapy that my impact outcome
  - Assessment of symptoms, signs, outcomes
- Blinding means patients, providers, & investigators do not know if patient got placebo or drug
- Reduces bias in management decisions and in assessment of outcomes by subject or investigator

Blinding

- Identical appearing therapies, eg, sham surgery
- Double-dummy, e.g., receive identical active and control together
- Side effect of therapy may unblind



Patient Population

- Will results be generalizable to population that will be treated in real world?
- Inclusion and exclusion criteria
  - Broad: exclude few, more generalizable
  - Restricted: exclude many, less generalizable
- Prospectively screen patients with condition of interest
  - Skipping patients may introduce bias
  - Keep log of subjects screened, but not enrolled

Choosing Control Intervention

- Placebo control vs. active control
- Placebo best to define efficacy of study therapy
  - May not be ethical or practical,
  - For example, because should not withhold standard care if documented effective
- Active control if hypothesis is that new therapy is superior or non-inferior (equivalent)

Primary vs. Additional Endpoints

- What is the most important impact you want to achieve with the new intervention?
- Trials should be designed to measure a small number of primary endpoints
- All others should be considered ”additional”

Surrogate vs. Clinically Meaningful

- Which study endpoint would alter practice?
  - Lab test (CRP) or clinical outcome (death)
- Studies of intermediate/surrogate endpoints may indicate areas for further research, but generally don't alter patient management
- Some surrogate endpoints are accepted as "true" indicators of clinical outcomes
  - e.g., blood pressure, cholesterol, colon polyps

Sample Size Based on Endpoints

- Primary endpoint result for intervention and comparator
- Assumptions based on available data, clinical judgment
- Difference should be clinically meaningful, realistic
- Power calculation based on:
  - Type 1 error: Probability of finding difference when does not exist
  - Type 2 error: Probability of not finding a difference when does exist

Feasible Sample Size

- Review medical records at study center(s) to
  - Determine number who meet enrollment criteria
  - Confirm assumptions about outcomes
- “Preparatory to research” review does not require IRB approval of protocol

Non-Inferiority Study

- Determine non-inferiority margin
- Maximal difference that would be considered clinically non-inferior, i.e., no unacceptably worse than the control
- Set a margin of difference less than the upper bound of the confidence interval observed in study
  - If you observe a difference more than the confidence intervals, then not ”non-inferior”

Recruitment and Retention

- Engagement before and throughout trial
  - Brochures, ads, social media, phone, text, email
  - Reminders for study personnel and participants
- Benefits of participation for subjects
  - Reimbursement for time and effort
- Identify and minimize barriers to participation
  - Easy access to study personnel and activities
  - Participation as non-onerous as possible

Follow Up of Patients

- If numerous patients are lost to follow-up, results of the trial may not be accurate
- Predefine method to deal with such patients
  - Last observation carried forward
  - Imputation methods
  - Re-calculate results assuming that patients lost to follow-up in treatment group had bad outcome and patients in control group had good outcome

STUDY ANALYSIS

Which Patients to Study?

- Intention to treat
  - Analyze all patients randomized regardless of whether they completed study
- Per-protocol
  - Did not receive sufficient study intervention
  - Did not return for necessary follow-up visits
  - Major violations of inclusion or exclusion criteria
  - Major violations during study

New Surgery to Stop Gastric Bleeding?

- Compare surgery vs. medications and endoscopy
- Patient randomized to receive surgery
- Dies from bleeding before can be operated on
- Do you include them in the study analysis?

Whom to Include in Analysis?

- Choose most conservative analysis
- Less likely to favor intervention, be overly optimistic

Superiority Study

- Per-protocol assumes optimal circumstances
  - Study done perfectly, no one dropped out, everyone took medication as prescribed
- ***Intention-to-treat analysis*** avoids bias to treatment difference and superiority

Non-Inferiority Study

- ITT can bias to no treatment difference
  - Bias to “non-inferiority”
  - e.g., non-adherence, drop-outs, misclassified subjects/endpoints
- ***Per protocol analysis*** should be included

Define Analysis Before Enrollment

- Predefine presentation of data
  - Proportions vs. time-to-event curves
  - Mean vs. median
- Predefine statistical analyses
  - Comparisons for primary, additional outcomes
  - Subgroup analyses
  - Sensitivity analyses

Is Difference Due to Chance?

- Null hypothesis means “true” proportion of success w/treatment = “true” proportion of success w/ control
- If the null hypothesis correct and treatments are equally effective, p-value indicates
  - Probability of observing a difference between treatment and control at least this large
  - Probability that difference at least this large is due to chance

Is Difference Due to Chance?

- A small p-value (< 0.05) means finding a difference at least this large is unlikely if the null hypothesis (treatments equally effective) is true
  - Reject the null hypothesis
- The more comparisons you make, the more likely you will get a result with a p<0.05
  - Single primary outcome > multiple outcomes
  - Statistical methods to ‘correct’ for multiple comparisons

Confidence Intervals

	Success (Small Study)	Success (Large Study)
Treatment	13/26 (50%)	1000/2000 (50%)
Control	10/26 (38%)	973/2000 (48.7%)
p-value	0.40	0.40
Difference	12%	1.3%
95% CI	-15% to 38%	-1.8% to 4.4%
Interpretation	Large effect size; high uncertainty	Small effect size; low uncertainty

Connor. AJG 2004;99:1638

Subgroup Analysis

Some groups respond differently	• Does drug produce benefit in men but not in women?
Pre-define and justify	• Post hoc more risk of bias since results known
Less power than overall analysis	• More analyses increases chance of significance occurring by chance • Smaller number in each subgroup

DATA SAFETY MONITORING BOARD

Data Safety Monitoring Board

"Ensures that a clinical trial is stopped if the benefit-risk balance for participants or the expected value to society no longer justifies continuing"

Lewis et al. JAMA 2016;316:2359

Data Safety Monitoring Board

Independent, external experts

Periodic assessments of recruitment, retention, data quality, adverse events, endpoints

Make recommendations to continue, modify, or terminate study

Interim Analysis

- Modify study
- Stop study
  - Futility: no possibility can document benefit
  - Efficacy: unequivocal benefit in clinically important outcome
- Stopping should be based on pre-defined rules

Stopping Study

- Pre-defined rules
- Usually very conservative p-value (e.g., 0.001)
  - Prevents incorrect conclusion (vs. p <0.05)
- Usually focused on safety
  - Stop if unacceptable risk-benefit balance
  - Consider additional monitoring and/or therapies

TRIAL REGISTRATION

Trial Registration

- Prevent selective publication and selective reporting of research outcomes
- Prevent unnecessary duplication of research
- Help public know of planned or ongoing trials into which they might want to enroll
- Give ethics boards considering new studies a view of similar work and data relevant to the research

http://icmje.org/recommendations/browse/publishing-and-editorial-issues/clinical-trial-registration.html

What is a Clinical Trial?

ClinicalTrials.gov	• Participants receive intervention per protocol • Outcomes measured for safety, efficacy
Medical Journal Editors	• Prospectively assigns people to an intervention…to study cause-and-effect relationship between a health intervention and a health outcome
NIH	• Human subject(s) prospectively assigned to intervention to evaluate the effects on health-related biomedical or behavioral outcomes

Obligation to Register

- Required by law (U.S. FDA)
  - Controlled clinical investigations of FDA-regulated drug, biologic, or device other than Phase 1 or small feasibility studies
- Enable publication (ICMJE)
  - All interventional studies, including Phase 1
- Ethical
  - “scientific, ethical and moral responsibility” *(WHO)*

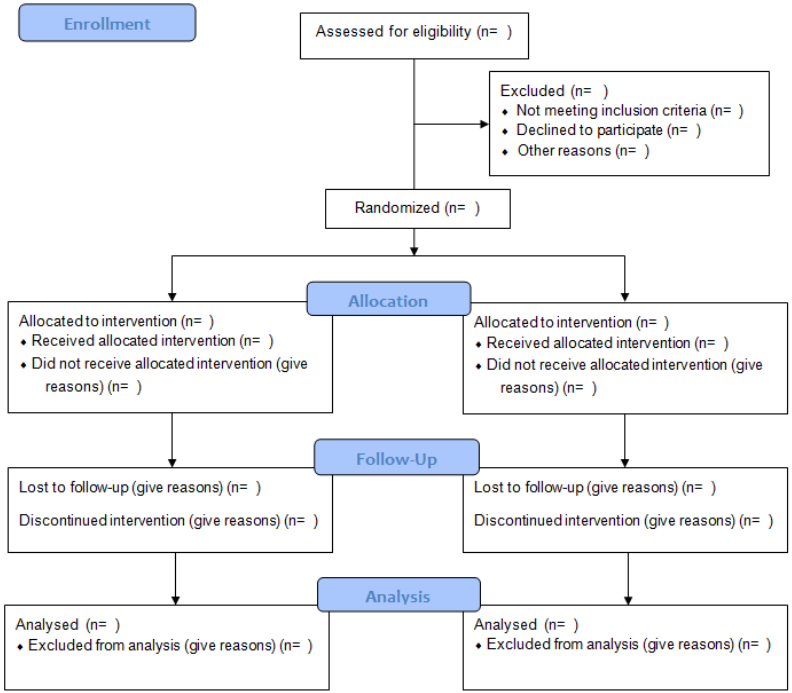
TRIAL REPORTING

CONSORT Guidelines on Reporting

- Framework for reporting RCTs
- Required by many journals
- Checklist for the content of the title, abstract, introduction, methods, results, and discussion
- Pre-trial registration ensures no change from original design without explanation
- Reporting guidelines ensure all important elements provided in publication

CONSORT
TRANSPARENT REPORTING of TRIALS

CONSORT 2010 Flow Diagram



TRIAL ASSESSMENT

General

- Background
  - Relevance, importance, novelty of topic
  - Rationale clear
- Hypothesis
  - Primary hypothesis stated
- Trial Registered, e.g., Clinicaltrials.gov
- Trial reported per guidelines, CONSORT

Design

- Randomization
  - Computer generated by uninvolved individual
  - Allocation concealed (opaque covering)
  - Block size stated
- Blinding
  - Patients, caregivers, investigators blinded
  - Method of blinding

Design

- Population
  - Inclusion, exclusion criteria clearly stated
  - Consecutive patients enrolled
- Intervention and Control
  - Characterized fully
  - Clinically appropriate
    - Simulate standard practice
    - Control is an acceptable standard of care

Design

- Primary and additional outcomes defined
  - Clinically relevant
  - Appropriately measured
- Sample size assumptions, calculation
  - Superiority vs. non-inferiority study
  - Assumptions reasonable

Analysis

- Population
  - ITT vs. per-protocol
  - Full accounting of subjects
- Primary and additional analyses
  - Predefined analyses presented
  - Post hoc analyses

Conclusions

- Do results support conclusions
- Limitations discussed
  - Potential sources of bias
  - Magnitude and precision of results
  - Generalizability
- Results placed in context
  - Outcomes not studied, e.g., cost, availability,
  - Consider relevant evidence outside study

WHY NOT DO RCT?

Parachute use to prevent death and major trauma when jumping from aircraft: randomized controlled trial

Robert W Yeh,[1] Linda R Valsdottir,[1] Michael W Yeh,[2] Changyu Shen,[1] Daniel B Kramer,[1] Jordan B Strom,[1] Eric A Secemsky,[1] Joanne L Healy,[1] Robert M Domeier,[3] Dhruv S Kazi,[1] Brahmajee K Nallamothu[4] On behalf of the PARACHUTE Investigators

[1]Richard A and Susan F Smith Center for Outcomes Research in Cardiology, Beth Israel Deaconess Medical Center, Harvard Medical School, 375 Longwood Avenue, Boston, MA 02215, USA
[2]David Geffen School of Medicine, University of California, Los Angeles, CA, USA
[3]Department of Emergency Medicine, University of Michigan and Saint Joseph Hospital, Ann Arbor, MI, USA
[4]Michigan Integrated Center for Health Analytics and Medical Prediction, Department of Internal Medicine and Institute for Healthcare Policy and Innovation, University of Michigan, Ann Arbor, MI, USA
Correspondence to: R W Yeh ryeh@bidmc.harvard.edu (or @rwyeh on Twitter)
Additional material is published online only. To view please visit the journal online.
Cite this as: *BMJ* 2018;363:k5094 http://dx.doi.org/10.1136/bmj.k5094
Accepted: 22 November 2018

ABSTRACT

OBJECTIVE
To determine if using a parachute prevents death or major traumatic injury when jumping from an aircraft.

DESIGN
Randomized controlled trial.

SETTING
Private or commercial aircraft between September 2017 and August 2018.

PARTICIPANTS
92 aircraft passengers aged 18 and over were screened for participation. 23 agreed to be enrolled and were randomized.

INTERVENTION
Jumping from an aircraft (airplane or helicopter) with a parachute versus an empty backpack (unblinded).

MAIN OUTCOME MEASURES
Composite of death or major traumatic injury (defined by an Injury Severity Score over 15) upon impact with the ground measured immediately after landing.

RESULTS
Parachute use did not significantly reduce death or major injury (0% for parachute *v* 0% for control; P>0.9). This finding was consistent across multiple subgroups. Compared with individuals screened but not enrolled, participants included in the study were on aircraft at significantly lower altitude (mean of 0.6 m for participants *v* mean of 914.6 m for non-participants; P<0.001) and lower velocity (mean of 0 km/h *v* mean of 800 km/h; P<0.001).

CONCLUSIONS
Parachute use did not reduce death or major traumatic injury when jumping from aircraft in the first randomized evaluation of this intervention. However, the trial was only able to enroll participants on small stationary aircraft on the ground, suggesting cautious extrapolation to high altitude jumps. When beliefs regarding the effectiveness of an intervention exist in the community, randomized trials might selectively enroll individuals with a lower perceived likelihood of benefit, thus diminishing the applicability of the results to clinical practice.

WHAT IS ALREADY KNOWN ON THIS TOPIC

Parachutes are routinely used to prevent death or major traumatic injury among individuals jumping from aircraft, but their efficacy is based primarily on biological plausibility and expert opinion

No randomized controlled trials of parachute use have yet been attempted, presumably owing to a lack of equipoise

WHAT THIS STUDY ADDS

This randomized trial of parachute use found no reduction in death or major injury compared with individuals jumping from aircraft with an empty backpack

Lack of enrolment of individuals at high risk could have influenced the results of the trial

Introduction

Parachutes are routinely used to prevent death or major traumatic injury among individuals jumping from aircraft. However, evidence supporting the efficacy of parachutes is weak and guideline recommendations for their use are principally based on biological plausibility and expert opinion.[1 2] Despite this widely held yet unsubstantiated belief of efficacy, many studies of parachutes have suggested injuries related to their use in both military and recreational settings,[3 4] and parachutist injuries are formally recognized in the World Health Organization's ICD-10 (international classification of diseases, 10th revision).[5] This could raise concerns for supporters of evidence-based medicine, because numerous medical interventions believed to be useful have ultimately failed to show efficacy when subjected to properly executed randomized clinical trials.[6 7]

Previous attempts to evaluate parachute use in a randomized setting have not been undertaken owing to both ethical and practical concerns. Lack of equipoise could inhibit recruitment of participants in such a trial. However, whether pre-existing beliefs about the efficacy of parachutes would, in fact, impair the enrolment of participants in a clinical trial has not been formally evaluated. To address these important gaps in evidence, we conducted the first randomized clinical trial of the efficacy of parachutes in reducing death and major injury when jumping from an aircraft.

Methods

Study protocol

Between September 2017 and August 2018, individuals were screened for inclusion in the PArticipation in RAndomized trials Compromised by widely Held beliefs aboUt lack of Treatment Equipoise (PARACHUTE) trial. Prospective participants were approached and screened by study investigators on commercial or private aircraft.

For the commercial aircraft, travel was related to trips the investigators were scheduled to take for business or personal reasons unrelated to the present study. Typically, passengers seated close to the study investigator (typically not known acquaintances) would be approached mid-flight, between the time of initial seating and time of exiting the aircraft. The

ABSTRACT

OBJECTIVE

To determine if using a parachute prevents death or major traumatic injury when jumping from an aircraft.

DESIGN

Randomized controlled trial.

SETTING

Private or commercial aircraft between September 2017 and August 2018.

PARTICIPANTS

92 aircraft passengers aged 18 and over were screened for participation. 23 agreed to be enrolled and were randomized.

INTERVENTION

Jumping from an aircraft (airplane or helicopter) with a parachute versus an empty backpack (unblinded).

MAIN OUTCOME MEASURES

Composite of death or major traumatic injury (defined by an Injury Severity Score over 15) upon impact with the ground measured immediately after landing.

RESULTS

Parachute use did not significantly reduce death or major injury (0% for parachute *v* 0% for control; $P>0.9$). This finding was consistent across multiple subgroups. Compared with individuals screened but not enrolled, participants included in the study were on aircraft at significantly lower altitude (mean of 0.6 m for participants *v* mean of 9146 m for non-participants; $P<0.001$) and lower velocity (mean of 0 km/h *v* mean of 800 km/h; $P<0.001$).

CONCLUSIONS

Parachute use did not reduce death or major traumatic injury when jumping from aircraft in the first randomized evaluation of this intervention. However, the trial was only able to enroll participants on small stationary aircraft on the ground, suggesting cautious extrapolation to high altitude jumps. When beliefs regarding the effectiveness of an intervention exist in the community, randomized trials might selectively enroll individuals with a lower perceived likelihood of benefit, thus diminishing the applicability of the results to clinical practice.

WHAT IS ALREADY KNOWN ON THIS TOPIC

Parachutes are routinely used to prevent death or major traumatic injury among individuals jumping from aircraft, but their efficacy is based primarily on biological plausibility and expert opinion

No randomized controlled trials of parachute use have yet been attempted, presumably owing to a lack of equipoise

WHAT THIS STUDY ADDS

This randomized trial of parachute use found no reduction in death or major injury compared with individuals jumping from aircraft with an empty backpack

Lack of enrolment of individuals at high risk could have influenced the results of the trial

Not Ethical or Practical

- Not ethical/possible to assign intervention
  - Cigarette smoking and lung cancer
  - H. pylori infection and ulcers
- Impractically large sample size
  - Very low-incidence event, eg, rare side effect
- Impractically long duration
  - Outcome requires years to develop, eg, cancer

More than One Way to Find “Truth”

- Medicine is different than other natural sciences in elevating the RCT to gold standard
- If you can directly measure the outcomes and all the relevant variables, then you do not need an RCT to control for noise (“unmeasured confounding”)
- This is true for engineering
  - You do not need an RCT to prove a new Iphone performs faster and better than an old Iphone

Parachute Example Again

- Design involves
  - High quality numeric models
  - Scale testing
  - Full-scale testing
  - Certification testing with live parachute trials, involving real people
- Initial simulation, then physical design, with parameters measured with high accuracy