Evolve or die: the urgent need to streamline randomized trials

Rory Collins

British Heart Foundation
Professor of Medicine & Epidemiology

Clinical Trial Service Unit &
Epidemiological Studies Unit (CTSU),
Nuffield Department of Population Health,
University of Oxford, UK

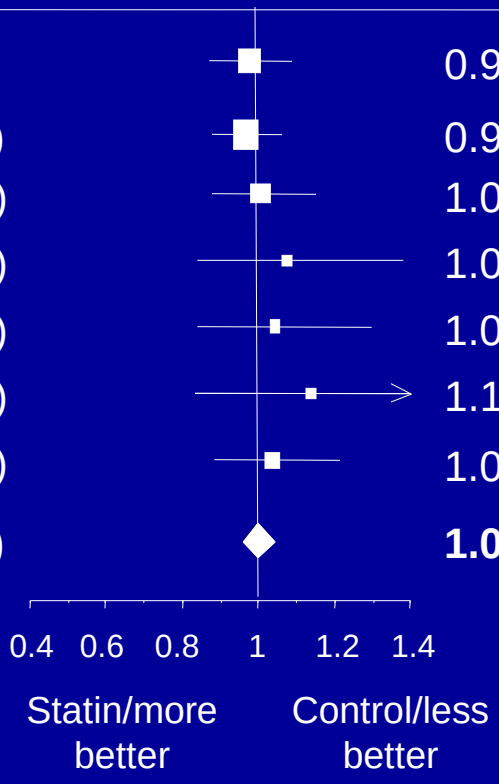
Limitations of observational studies for identifying causal effects of treatments on health outcomes

- Observational studies can help in identifying large effects of treatments on health outcomes that are otherwise rare (e.g. myopathy with statin therapy)
- However, observational studies of the associations of treatment with health outcomes may be prone to various sources of confounding and bias

So-called “real world” observational studies in large health-care databases may well yield associations of treatment with health outcomes that are precise (i.e. have small random errors) but are not causal

Meta-analysis of large RCTs of statin therapy: effects on site-specific CANCER INCIDENCE (Lancet 2010)

	No. of first cancers (% pa)		Relative risk (CI) per mmol/L LDL-C reduction
	Statin/more	Control/less	
Gastrointestinal	1166 (0.3%)	1194 (0.3%)	0.97 (0.87 - 1.09)
Genitourinary	1596 (0.5%)	1645 (0.5%)	0.97 (0.88 - 1.06)
Respiratory	813 (0.2%)	814 (0.2%)	1.00 (0.88 - 1.15)
Female breast	267 (0.3%)	241 (0.3%)	1.07 (0.84 - 1.38)
Haematological	305 (0.1%)	291 (0.1%)	1.04 (0.84 - 1.30)
Melanoma	159 (0.0%)	142 (0.0%)	1.14 (0.83 - 1.56)
Other/unknown	754 (0.2%)	737 (0.2%)	1.04 (0.89 - 1.21)
Any	**5060 (1.4%)**	**5064 (1.4%)**	**1.00 (0.96 - 1.04)**

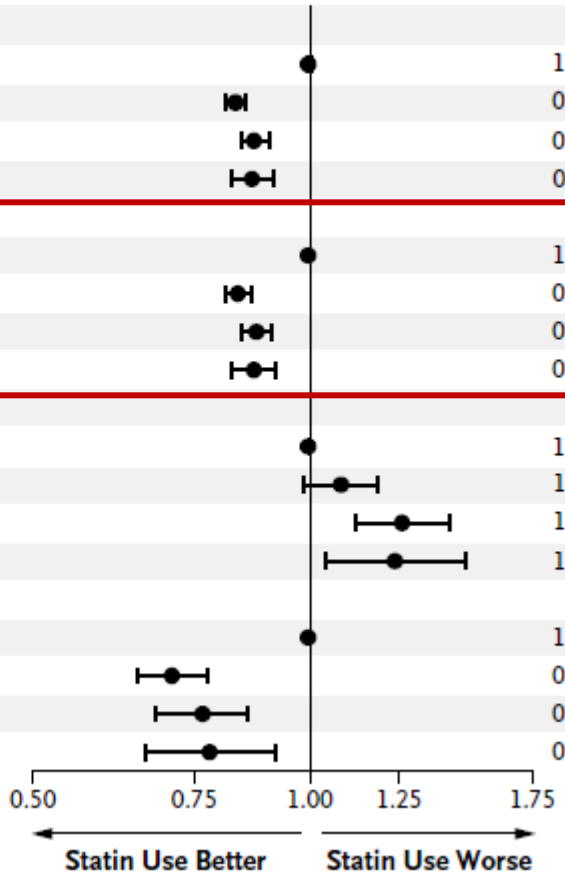


No emerging effect on cancer with prolonged follow-up by linkage to health-record systems

Danish population study: Highly significant non-causal association of statin use with lower cancer mortality (Nielsen et al, NEJM 2012)

A Nationwide Study

Cause of Death and Statin Dose	No. of Patients	No. of Deaths	Hazard Ratio (95% CI)	P Value
Any cause				
0.00	277,204	184,895	1.00	
0.01–0.75	9,780	5,730	0.82 (0.81–0.85)	<0.001
0.76–1.50	6,181	3,438	0.87 (0.83–0.89)	<0.001
>1.50	2,760	1,531	0.87 (0.81–0.91)	<0.001
Cancer				
0.00	277,204	153,327	1.00	
0.01–0.75	9,780	4,680	0.83 (0.81–0.86)	<0.001
0.76–1.50	6,181	2,810	0.87 (0.83–0.91)	<0.001
>1.50	2,760	1,250	0.87 (0.81–0.92)	<0.001
Cardiovascular cause				
0.00	277,204	13,512	1.00	
0.01–0.75	9,780	529	1.08 (0.99–1.19)	0.08
0.76–1.50	6,181	314	1.25 (1.21–1.41)	<0.001
>1.50	2,760	134	1.24 (1.03–1.48)	0.01
Other cause				
0.00	277,204	18,056	1.00	
0.01–0.75	9,780	521	0.70 (0.64–0.77)	<0.001
0.76–1.50	6,181	314	0.76 (0.68–0.86)	<0.001
>1.50	2,760	147	0.77 (0.66–0.92)	0.003



"Real-World Evidence – What is it and what can it tell us." Sherman and others from the Office of the US FDA Commissioner (NEJM December 2016)

"… the confluence of large data sets of uncertain quality and provenance, the facile analytic tools that can be used by non-experts, and a shortage of researchers with adequate methodologic savvy could result in poorly conceived study and analytic designs that generate incorrect or unreliable conclusions"

Limitations of observational studies for identifying causal effects of treatments on health outcomes

- Observational studies can be useful for identifying large effects of treatments on health outcomes that are rare (e.g. myopathy with statin therapy)
- However, observational studies of the associations of treatment with health outcomes may be prone to various sources of confounding and bias
- Consequently, randomized trials are needed to detect plausibly moderate beneficial or adverse effects of treatments on common health outcomes (with wide eligibility criteria yielding generalizability)

Adverse impact of increased regulation and related bureaucracy on randomised trials

- Increased obstacles, delays and costs for trials (encouraging mis-use of “real world” evidence)
- Distorts research agenda and reduces creative collaborations between academia and industry
- Fewer new treatments developed by industry, and fewer academic trials of current therapies
- Undue focus on complying with rules rather than on innovation in the design and conduct of trials

International Conference on Harmonisation (ICH) is the key obstacle to better randomized trials

- Lack of transparency
  - Who decides at ICH?
  - How does one influence them?
- Lack of representativeness
  - Regulators and Industry only
  - Why not patients or academics?
- Lack of evidence of competence
  - Proven failures of ICH-GCP guidelines
  - Contradictory text in proposed amendment

Key issues with ICH-GCP guidelines for RCTs

- Fundamental: Not based on the key scientific principles of RCTs that are critical for the generation of reliable results

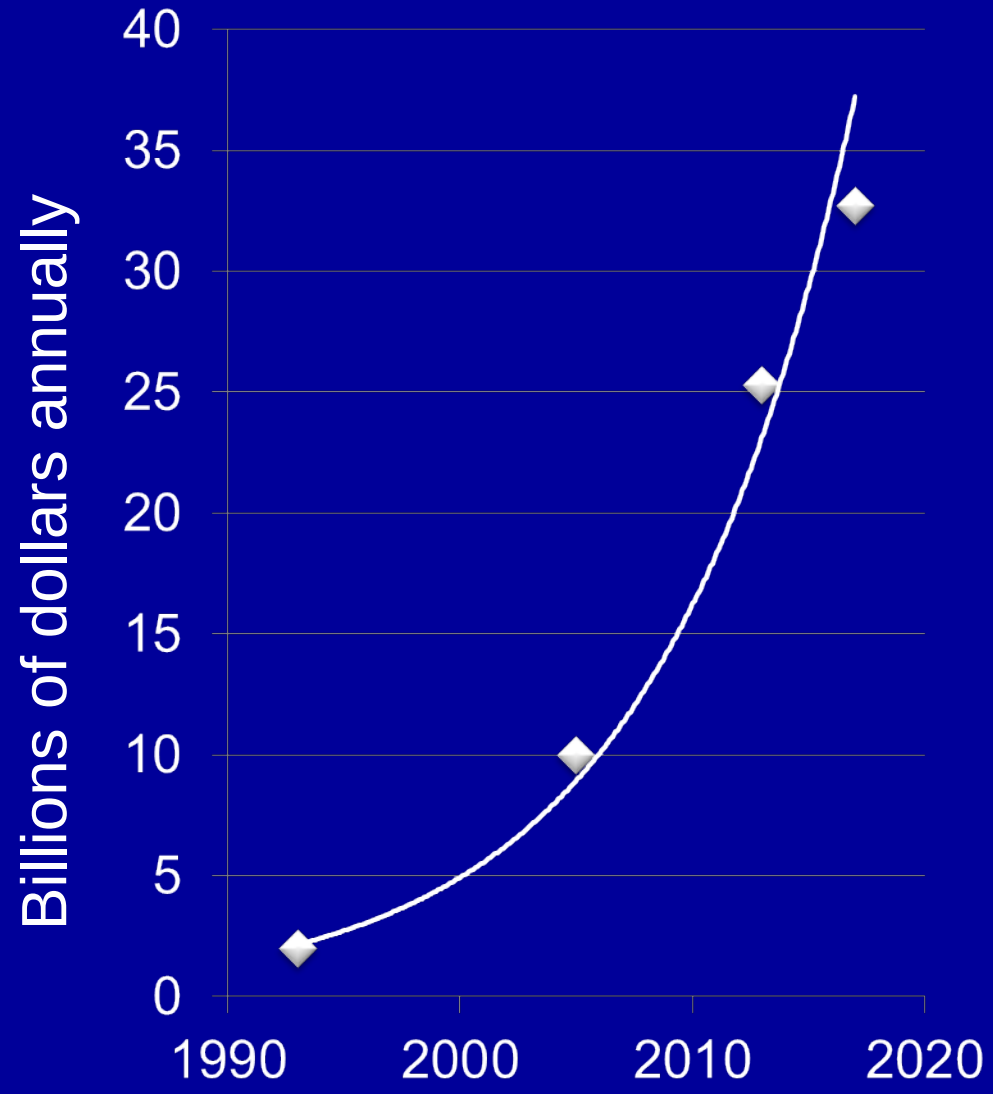
Key issues with ICH-GCP guidelines for RCTs

- Fundamental: Not based on the key scientific principles of RCTs that are critical for the generation of reliable results
- Not even working well for registration trials of new drugs: unsustainable costs; wasteful practices; poor quality

FOURIER trial of PCSK9 inhibitor therapy

- 1 Billion US dollars reported cost
- 28,000 patients for 2.2 y of follow-up
- Under-estimated LDL-lowering benefit
- Failed to demonstrate effect on mortality

Growth in the Contract Research Organization (CRO) market since the creation of ICH in 1990



2015 survey predicts 10% annual growth during the next 5 years (but plot is exponential)

Roberts et al Cancer 2016

Key issues with ICH-GCP guidelines for RCTs

- Fundamental: Not based on the key scientific principles of RCTs that are critical for the generation of reliable results
- Not even working well for registration trials of new drugs: unsustainable costs; wasteful practices; poor quality
- Applied more widely than intended (e.g. EU Regulation; Gates Foundation) to RCTs of all types of intervention

  EU Regulation: *"…ICH guidelines on good clinical practice should be taken appropriately into account for the application of the rules set out in this Regulation"*

  Gates Foundation: *"You will adhere to current Good Clinical Practice as defined by the International Council on Harmonisation (ICH) E-6 Standards (or local regulations if more stringent)"*

Key issues with ICH-GCP guidelines for RCTs

- Fundamental: Not based on the key scientific principles of RCTs that are critical for the generation of reliable results
- Not even working well for registration trials of new drugs: unsustainable costs; wasteful practices; poor quality
- Applied more widely than intended (e.g. EU Regulation; Gates Foundation) to RCTs of all types of intervention
- Even ICH has recognised that its GCP guidelines are not able to accommodate regulatory changes or innovations
- ICH has now initiated a detailed revision of its guidelines, but again is not involving patients or non-industry trialists

Why Not?

Examples of inappropriate “safety” monitoring in randomised trials driven by regulatory pressures (often due to over-interpretation of unclear rules)

- Requirement to record all AEs not just serious AEs (despite good evidence of lack of effects)
- Requirement to record narratives for all serious AEs in case there is an excess of a particular AE
- Demands for unblinded results for AEs (including even primary outcomes) during ongoing trials
- Demands for blinded line-lists quarterly despite it being pointed out that these are uninterpretable
- Annual reports required by regulatory authorities that are so long that safety signals risk being lost

Lack of logical basis for Regulatory rules for pharmacovigilance in randomized trials

- Reporting of all Suspected Unexpected Serious Adverse Reactions (SUSARS), as is required by Regulations, may be able to detect LARGE effects on RARE outcomes
  - but this does not require randomized control (and has only rarely detected major signals)
- Reliable assessment of MODERATE effects on COMMON outcomes does, however, need LARGE-SCALE RANDOMIZED evidence
  - and such effects are likely to have a far bigger impact on public health than rare side-effects

THRIVE trial: Unexpected adverse effects of ER niacin only emerged from a large randomised comparison

SAE outcomes	Niacin	Placebo	Odds ratio (& 95% CI)
New onset diabetes	494 (5.7%)	376 (4.3%)	1.32 (1.16–1.51)
Any infection	1031 (8.0%)	853 (6.6%)	1.22 (1.12-1.34)
Any bleeding	326 (2.5%)	238 (1.9%)	1.38 (1.17-1.62)

None of these excesses was identified by regulatory SUSAR or SAE reporting (despite 50 years of pharmacovigilance)

ICH-GCP: Guidance on monitoring of collaborating sites in clinical trials

"The purposes of trial monitoring are to verify that… reported trial data are accurate, complete, and verifiable from source documents."

ICH-GCP 5.18.1

"In general there is a need for on-site monitoring before, during and after the trial"

ICH-GCP 5.18.3

Undue emphasis in ICH-GCP on the quality of outcome assessment data (e.g. source data verification)

High quality DATA ≠ Reliable RESULT

Reliable RESULT ≠ High quality DATA

Examples of inefficient/ineffective site monitoring driven by compliance with ICH-GCP guidance

- Investigators' qualifications
  - Curriculum vitae
  - GCP training
- Consent
  - Review of consent forms but not process
- Source data verification
  - Non-critical blood results & physical measures
  - Use of routine concomitant medications
  - Unimportant adverse events
- Regulatory documentation
  - Approval letters, etc in established centres
  - Individual SAR (15-day) reports
- Drug accountability
  - Pill counts

Need more efficient monitoring of RCTs (since site visits are a large part of costs)

- Implement central statistical monitoring (rather than routine site visits) to improve study quality (rather than merely detecting past problems)
- Use study data to detect potential deviations from the study protocol and other local issues
- Direct visits to sites that have been identified with potential problems by central monitoring
- Focus study visits on mentoring local site staff (not on source data verification or “box ticking”)

Encourage (not inhibit) innovation so that the conduct of RCTs is smarter

Limited value (in most cases) of data checks on major health outcomes in randomized trials

- Put greater reliance on comparison with the randomly-allocated control group
- Missing data have little impact if this is unbiased with respect to allocated group

Minimal impact on findings of adding false events or of missing real events

	Active (10,000)	Control (10,000)	OR (& 95%CI)	P-value
True events	800	1000	0.78 (0.71-0.86)	<0.00001

Minimal impact on findings of adding false events or of missing real events

	Active (10,000)	Control (10,000)	OR (& 95%CI)	P-value
True events	800	1000	0.78 (0.71-0.86)	<0.00001
Missing real events (unevenly distributed)				
- 10%	720	900	0.78 (0.71-0.87)	<0.00001
- 20%	640	800	0.79 (0.71-0.88)	=0.00001

Limited value (in most cases) of data checks on major health outcomes in randomized trials

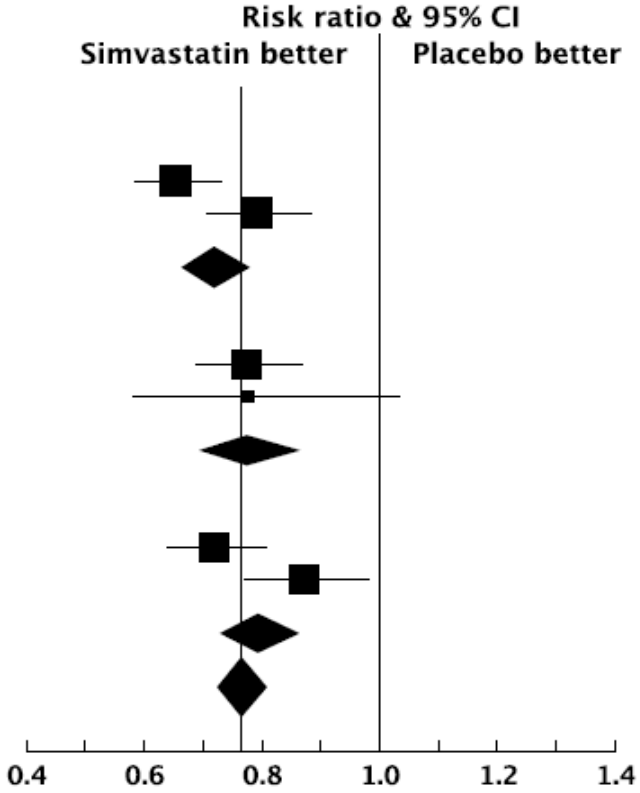
- Put greater reliance on comparison with the randomly-allocated control group
- Missing data have little impact if this is unbiased with respect to allocated group
- Adjudication of study outcomes adds substantial cost, but typically little gain

Minimal impact on findings of adding false events or of missing real events

	Active (10,000)	Control (10,000)	OR (& 95%CI)	P-value
True events	800	1000	0.78 (0.71-0.86)	<0.00001
Missing real events (unevenly distributed)				
- 10%	720	900	0.78 (0.71-0.87)	<0.00001
- 20%	640	800	0.79 (0.71-0.88)	=0.00001
Extra false events (evenly distributed)				
+ 10%	890	1090	0.80 (0.73-0.88)	<0.00001
+ 20%	980	1180	0.81 (0.74-0.89)	<0.00001

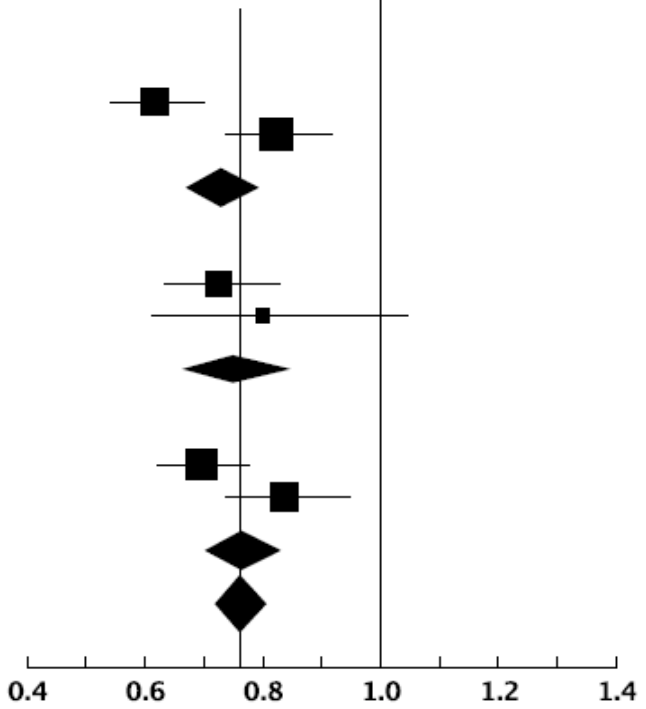
HPS: Effects of simvastatin-allocation on UN-ADJUDICATED major vascular events

Type of event	Simvastatin allocation: Simvastatin (n=10269)		Simvastatin allocation: Placebo (n=10267)		Risk ratio & 95% CI
Coranary events					
Non-fatal MI	492	(4.8%)	743	(7.2%)	0.65 (0.58-0.73)
Coronary death	547	(5.3%)	687	(6.7%)	0.79 (0.71-0.88)
Any coronary event	**988**	**(9.6%)**	**1350**	**(13.1%)**	**0.72 (0.66-0.78)**
Strokes					
Non-fatal stroke	487	(4.7%)	621	(6.0%)	0.77 (0.69-0.87)
Fatal stroke	82	(0.8%)	105	(1.0%)	0.78 (0.58-1.03)
Any stroke	**550**	**(5.4%)**	**700**	**(6.8%)**	**0.77 (0.69-0.87)**
Revascularisations					
Coronary	487	(4.7%)	667	(6.5%)	0.72 (0.64-0.81)
Non-coronary	492	(4.8%)	559	(5.4%)	0.87 (0.77-0.98)
Any revascularisation	**943**	**(9.2%)**	**1166**	**(11.4%)**	**0.79 (0.73-0.86)**
ANY MAJOR VASCULAR EVENT	**2187**	**(21.3%)**	**2765**	**(26.9%)**	**0.77 (0.72-0.81)**



HPS: Effects of simvastatin-allocation on ADJUDICATED major vascular events (Lancet 2002)

Type of event	Simvastatin allocation: Simvastatin (n=10269)		Simvastatin allocation: Placebo (n=10267)		Risk ratio & 95% CI
Coranary events					
Non-fatal MI	357	(3.5%)	574	(5.6%)	0.62 (0.54-0.70)
Coronary death	587	(5.7%)	707	(6.9%)	0.82 (0.74-0.92)
Any coronary event	**898**	**(8.7%)**	**1212**	**(11.8%)**	**0.73 (0.67-0.79)**
Strokes					
Non-fatal stroke	366	(3.6%)	499	(4.9%)	0.72 (0.63-0.83)
Fatal stroke	96	(0.9%)	119	(1.2%)	0.80 (0.61-1.05)
Any stroke	**444**	**(4.3%)**	**585**	**(5.7%)**	**0.75 (0.66-0.85)**
Revascularisations					
Coronary	513	(5.0%)	725	(7.1%)	0.70 (0.62-0.78)
Non-coronary	450	(4.4%)	532	(5.2%)	0.84 (0.74-0.95)
Any revascularisation	**939**	**(9.1%)**	**1205**	**(11.7%)**	**0.76 (0.70-0.83)**
ANY MAJOR VASCULAR EVENT	**2033**	**(19.8%)**	**2585**	**(25.2%)**	**0.76 (0.72-0.81)**



Adjudication represents a very large effort to obtain and review medical records for all relevant outcomes

Strengths of randomised placebo-controlled trials for providing reliable evidence about efficacy and safety

Randomisation: Provides groups of patients that differ only randomly from each other in terms of risk of events (so differences in outcome can be inferred to be causal)

Blinded-control: Allows unbiased comparison of events ascertained similarly in the randomized treatment groups (so differences in event identification are applied equally)

Consequently, randomised blinded-controlled trials can provide unbiased assessments of both the efficacy and safety of treatments (including symptomatic side-effects)

By contrast, in routine care, patients know they are taking a treatment, so may falsely attribute health outcomes to it (and other biases render observational studies unreliable)

Adverse impact of increased regulation and related bureaucracy on randomised trials

- Increased obstacles, delays and costs for trials (encouraging the use of “real world” evidence)
- Distorts research agenda and reduces creative collaborations between academia and industry
- Fewer new treatments developed by industry, and fewer academic trials of current therapies
- Undue focus on complying with rules rather than on innovation in the design and conduct of trials

Solution: Develop evidence-based strategies for randomized trials (do not increase inappropriate reliance on non-randomized observational data)

Back to the Future?

WHY DO WE NEED SOME LARGE, SIMPLE RANDOMIZED TRIALS?

SALIM YUSUF* RORY COLLINS AND RICHARD PETO

Clinical Trial Service Unit, Radcliffe Infirmary, Oxford, UK

The criteria for a good trial are similar in many serious diseases: first and foremost, ask an 'important' question and, secondly, answer it 'reliably'. These two very general criteria obviously require further elaboration, but even as they stand they can suggest some surprisingly specific consequences for clinical trial design. Particularly, they can be used to suggest both the possibility and the desirability of *large*, *simple randomized* trials of the effects on *mortality* of various *widely practicable* treatments for *common* conditions.

Statistics in Medicine; 1984