Randomized Controlled Clinical Trials (RCTs) and Its Applications in Health Research

Chung-Yi Li PhD
Department of Public Health, College of Medicine
National Cheng Kung University
Tainan, Taiwan


Google 地圖
交通時間、路況和附近地點
寧德市
福州市
福建
泉州市
漳州市
廈門市
台北
台中市
台灣
台南市
高雄市
廣東
揭陽市
汕頭市
州市
東莞市
深圳市
澳門
香港
Hong Kong

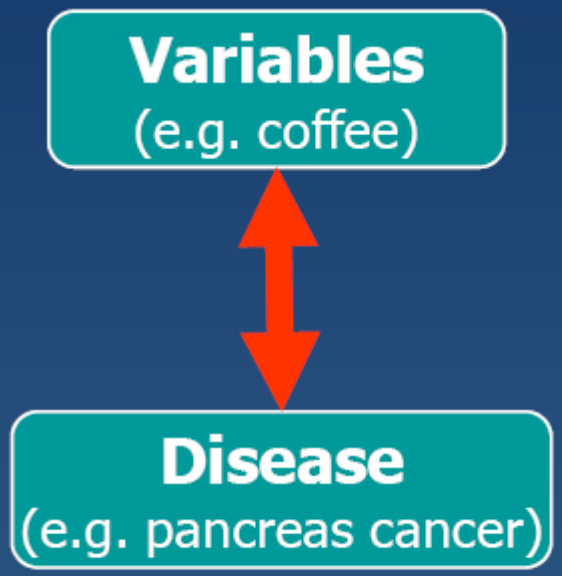
Outlines

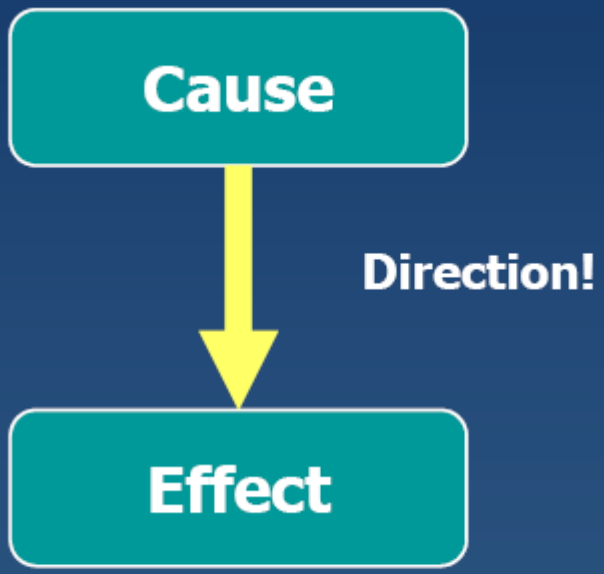
- Evidence for causal inference
- Why is randomization so important?
  - Confounding by indication
  - Assumption for difference-in-difference analysis
- Methodology of RCTs
  - Study design
  - Data analysis
- What can RCTs do, and what RCTs can't do?

Evidence for causal inference

Causal inference has never been easy in the real world

Association and Causal Relation



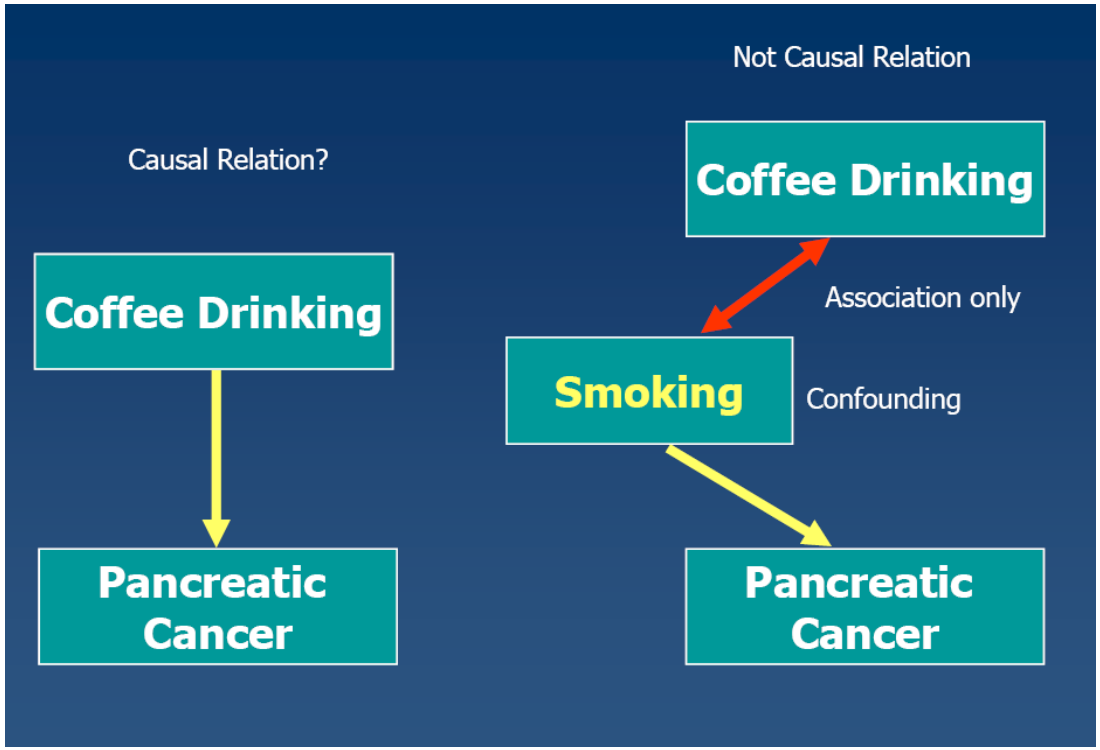


Statistic can tell association!

Statistic can NOT tell causality!

Cause

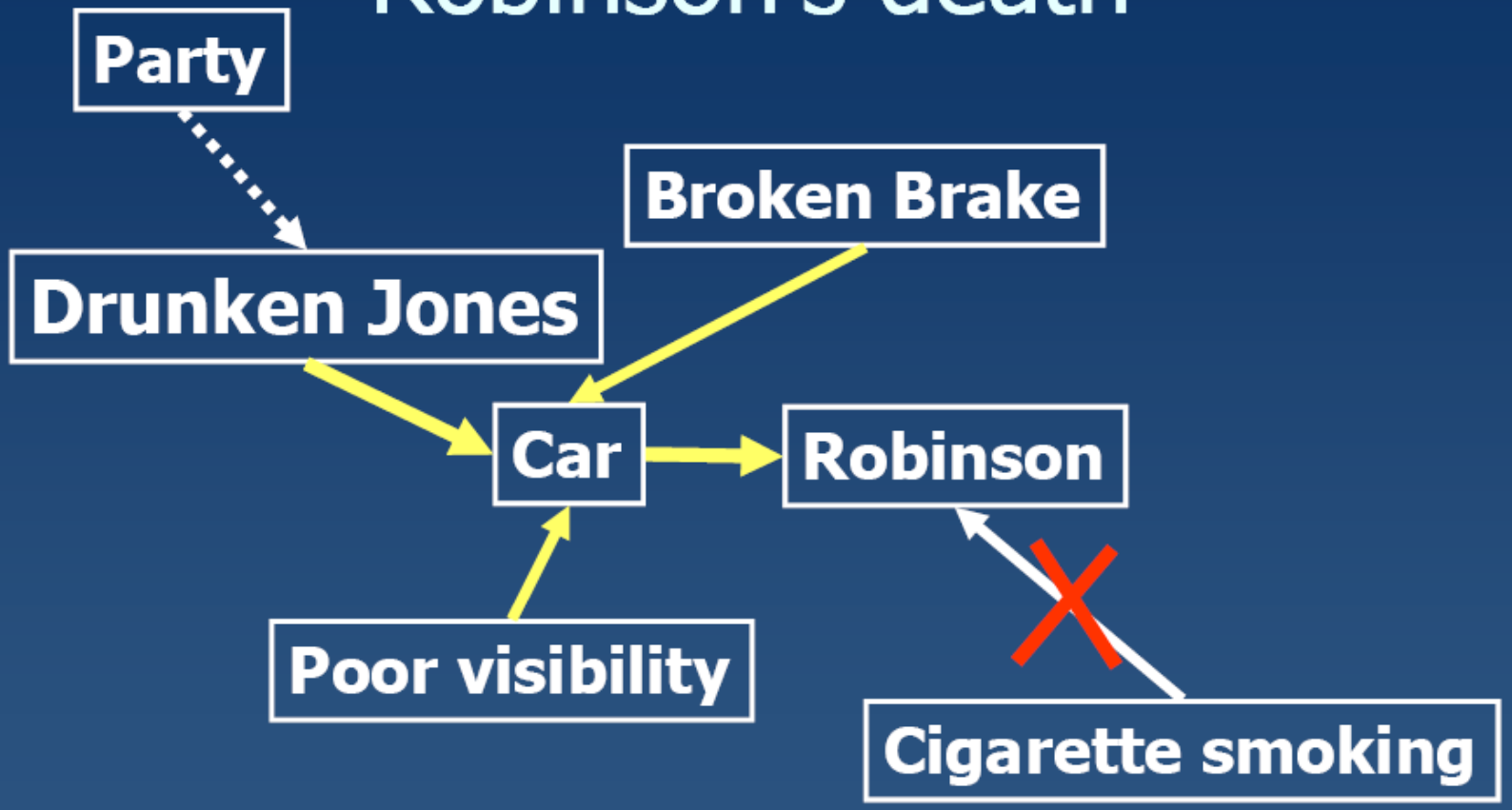
- Is that which produces "an effect, result, or consequence" or "the one, such as a person, event, or condition that is responsible for an action or result" (American Heritage Dictionary)



Who kill(s) Robinson

- Jones driving from a party where he has drunk too much, in a car whose brakes are defective, at an intersection with poor visibility runs down and kills Robinson, who was crossing the read to buy cigarettes.

Robinson's death



Types of Epidemiological Studies

EPIDEMIOLOGY SERIES

Epidemiology series

An overview of clin

David A Grimes, Kenneth F Schulz

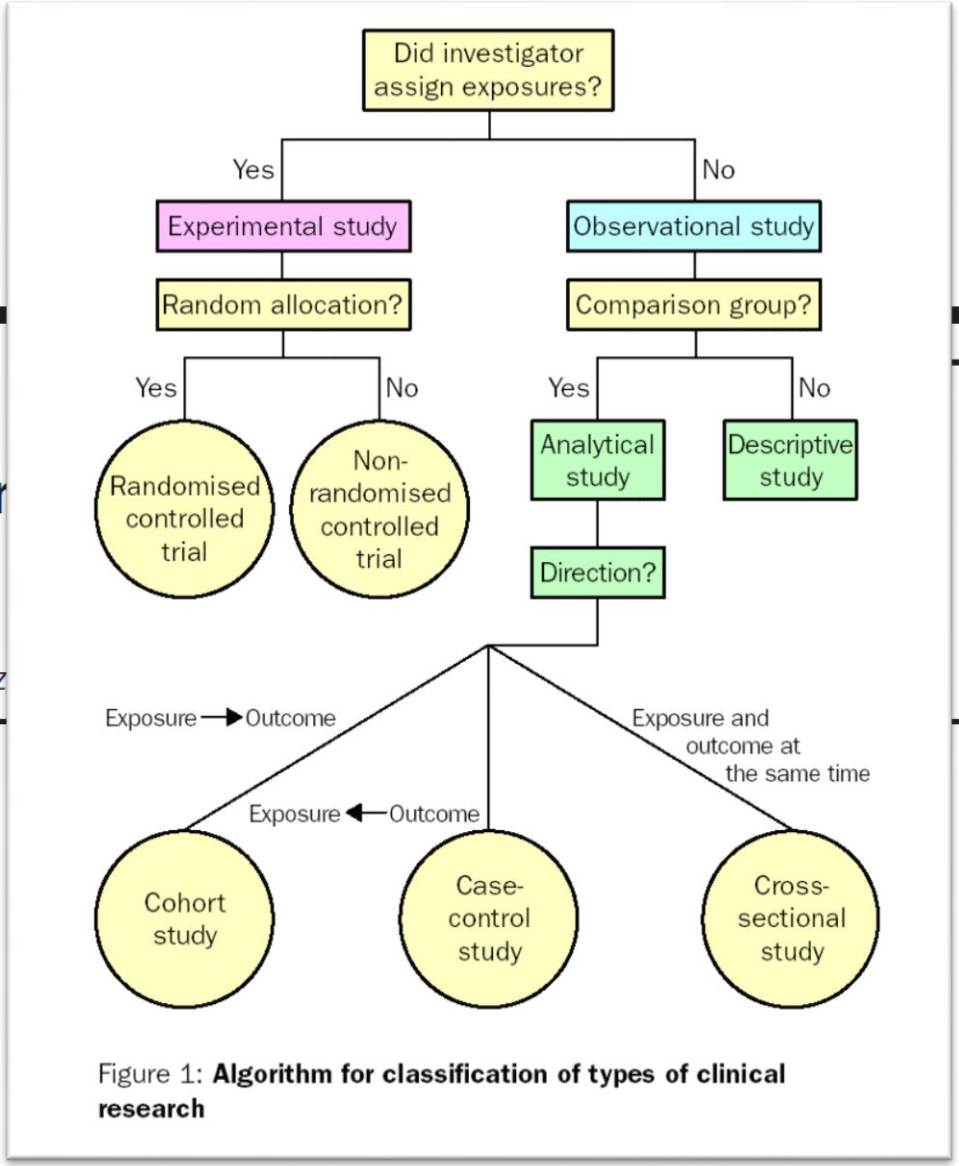
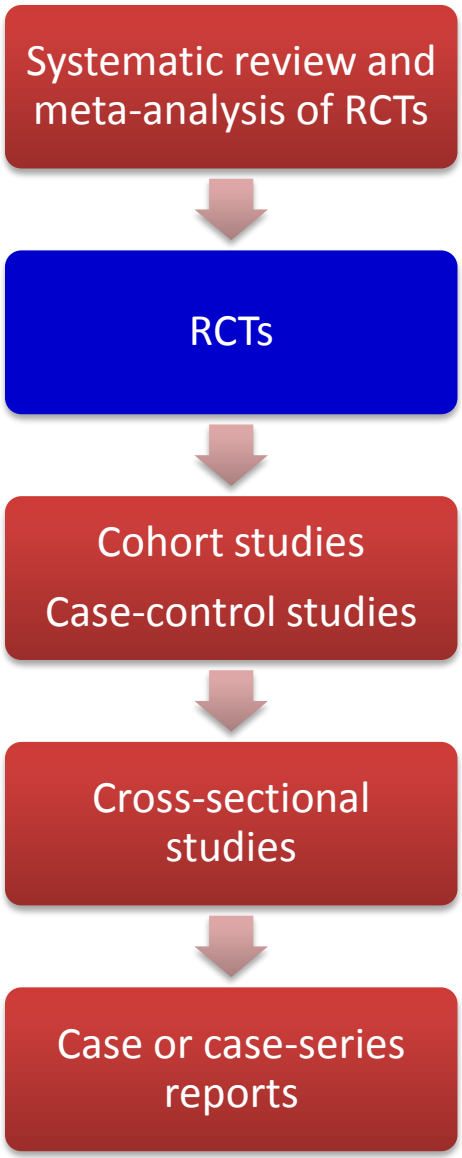


Figure 1: **Algorithm for classification of types of clinical research**

Evidence for causal inference



Example for RCT-based Systematic Review & Meta-analysis

ACRM
AMERICAN CONGRESS OF REHABILITATION MEDICINE

Archives of Physical Medicine and Rehabilitation

journal homepage: www.archives-pmr.org

Archives of Physical Medicine and Rehabilitation 2017;98:1666-77

REVIEW ARTICLE (META-ANALYSIS)

Effects of Home-Based Supportive Care on Improvements in Physical Function and Depressive Symptoms in Patients With Stroke: A Meta-Analysis

Hui-Chuan Huang, PhD,[a,b] Yi-Chieh Huang, MSN,[b] Mei-Feng Lin, RN, PhD,[c] Wen-Hsuan Hou, PhD,[d] Meei-Ling Shyu, EdD,[b] Hsiao-Yean Chiu, PhD,[b] Hsiu-Ju Chang, PhD[b]

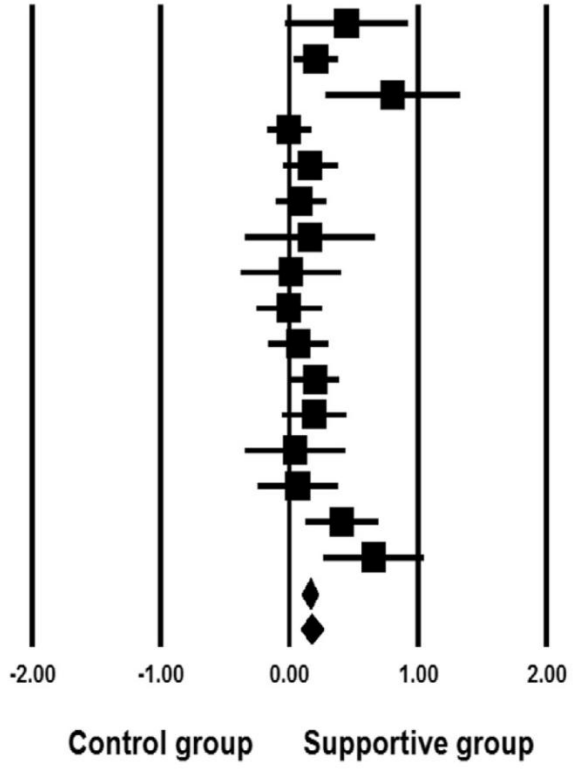
Effect on Physical Function

A

Model	Study name	Hedges's g	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
	Clark et al (2003)24	0.446	0.243	0.059	-0.030	0.922	1.838	0.066
	Mant et al (2000)25	0.207	0.088	0.008	0.035	0.379	2.356	0.018
	Zhang et al (2012)26	0.806	0.265	0.070	0.286	1.326	3.040	0.002
	Boter (2004)27	0.000	0.086	0.007	-0.169	0.169	0.000	1.000
	Tilling et al (2005)28	0.164	0.108	0.012	-0.049	0.376	1.509	0.131
	Dennis et al (1997)29	0.089	0.098	0.010	-0.102	0.281	0.914	0.360
	Chang and Li (2000)30	0.159	0.255	0.065	-0.341	0.660	0.625	0.532
	Anderson et al (2002)40	0.015	0.197	0.039	-0.371	0.401	0.077	0.939
	Forster and Young (1996)41	0.000	0.129	0.017	-0.252	0.252	0.000	1.000
	Glass et al (2004)42	0.073	0.117	0.014	-0.157	0.302	0.620	0.535
	Johsson et al (2014)43	0.203	0.093	0.009	0.020	0.386	2.170	0.030
	Lincoln et al(2003)44	0.196	0.126	0.016	-0.052	0.444	1.551	0.121
	Markle-Reid et al (2011)45	0.045	0.198	0.039	-0.342	0.433	0.230	0.818
	Ostwald et al(2014)46	0.068	0.158	0.025	-0.242	0.377	0.429	0.668
	Wang et al (2010)47	0.410	0.142	0.020	0.131	0.689	2.877	0.004
	Wang et al(2015)48	0.655	0.196	0.039	0.271	1.040	3.339	0.001
Fixed		0.157	0.032	0.001	0.095	0.220	4.921	0.000
Random		0.173	0.044	0.002	0.088	0.258	3.979	0.000

Statistics for each study

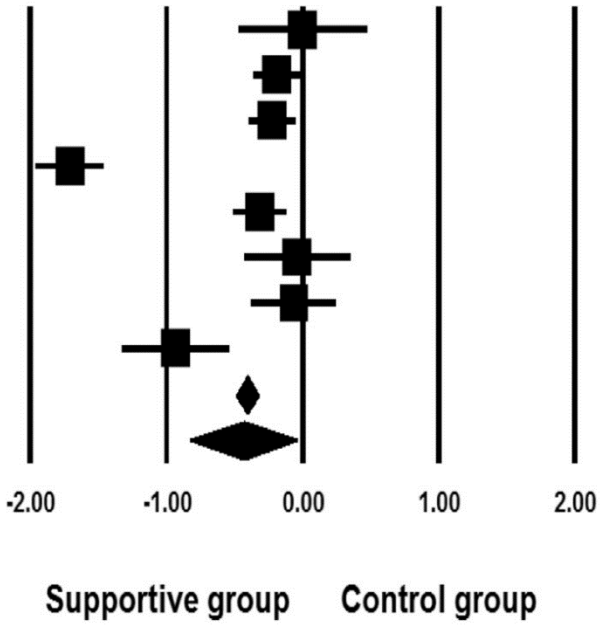
Hedges's g and 95% CI



Effect on Depression

B

Model	Study name	Hedges's g	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
	Clark et al (2003)24	0.000	0.240	0.058	-0.470	0.470	0.000	1.000
	Mant et al (2000)25	-0.189	0.088	0.008	-0.361	-0.017	-2.152	0.031
	Boter (2004)27	-0.225	0.087	0.007	-0.395	-0.055	-2.598	0.009
	Tilling et al(2005)28	-1.710	0.127	0.016	-1.958	-1.462	-13.516	0.000
	Dennis et al(1997)29	-0.316	0.098	0.010	-0.508	-0.123	-3.209	0.001
	Markle-Reid et al(2011)45	-0.041	0.198	0.039	-0.429	0.346	-0.209	0.834
	Ostwald et al (2014)46	-0.065	0.158	0.025	-0.374	0.245	-0.411	0.681
	Wang et al (2015)48	-0.934	0.201	0.041	-1.329	-0.540	-4.640	0.000
Fixed		-0.411	0.043	0.002	-0.496	-0.326	-9.524	0.000
Random		-0.440	0.199	0.040	-0.830	-0.050	-2.210	0.027



Assessment of Methodological Quality Using the Cochrane Risk of Bias Tool

Table 3 Risk ias assessment for the methodologic lity of the included

Author (Year)	Selection Bias		erformance	ion Bias	Attrition Bias	Reporting Bias
		ing			ete	
	Allo				me Data	Selective
					ressed	Reporting
Clark et al (2003)[24]						L
Mant et al						
Zhang et al						L
Boter (2004)[27]						L
Tilling et al (2005)[28]						L
Dennis et al (1997)[29]						L
Chang and Li (2000)[30]						
Anderson et al (200						
Forster and Yo						L
Glass et al (2004)	L					L
Jonsson et al (2014)[43]	L					L
Lincoln et al (2003)[44]	L					L
Markle-Reid et al (2011)[45]	L	L			H	L
Ostwald et al (2014)[46]	L	L			H	L
Wang et al (2010)[47]	?	?	H		L	L
Wang et al (2015)[48]	L	L	H		L	L

NOTE. Risk of bias assessment was according to the Cochrane aboration's tool for assessing risk of bias in randomized trials.
Abbreviations: ?, Unclear risk of bias; H, high risk of bias; L, risk of bias.

"*GIGO*" if Quality Assessment is not well performed.

Why is randomization so important?

Confounding by Indication

Curr Epidemiol Rep (2014) 1:1–8
DOI 10.1007/s40471-013-0004-y

REPRODUCTIVE AND PERINATAL EPIDEMIOLOGY (EF SCHISTERMAN, SECTION EDITOR)

Confounding by Indication and Related Concepts

K. S. Joseph · Azar Mehrabadi · Sarka Lisonkova

Table 1 Estimated numbers and rates of maternal deaths and maternities by type of delivery, United Kingdom 2000–2002 [6]

Type of delivery	Estimated number of maternities	Number of maternal deaths	Death rate per 100,000 maternities	Rate ratio 95 % confidence interval
Vaginal	1,571,000	75	4.8	1.0 (−)
Cesarean	425,000	73	17.2	3.7 (2.6–5.0)
Emergency and urgent	212,000	44	20.8	4.3 (3.0–6.3)
Scheduled and elective*	214,000	29	13.6	2.8(1.9–4.4)

A Community Trial w/o Randomization

RESEARCH AND PRACTICE

Increasing Use of Mammography Among Older, Rural African American Women: Results From a Community Trial

Jo Anne Earp, ScD, Eugenia Eng, DrPH, Michael S. O'Malley, PhD, Mary Altpeter, PhD, Garth Rauscher, MPH, Linda Mayne, PhD, RN, Holly F. Mathews, PhD, Kathy S. Lynch, MPH, and Bahjat Qaqish, MD, PhD

American Journal of Public Health | April 2002, Vol 92, No. 4

Methods

- **OBJECTIVES:**
  - A community trial was undertaken to evaluate the effectiveness of the North Carolina Breast Cancer Screening Program, a lay health advisor network intervention intended to increase screening among rural African American women 50 years and older.
- **METHODS:**
  - A stratified random sample of 801 African American women completed baseline (1993-1994) and follow-up (1996-1997) surveys. The primary outcome was self-reported mammography use in the previous 2 years

TABLE 1—Baseline Characteristics of Black Female Respondents: North Carolina Breast Cancer Screening Program, 1993–1994 and 1996–1997

Characteristic	Intervention (n = 390), %	Comparison (n = 411), %
Personal		
Age, y		
50–64	46	44
65–74	31	32
≥75	23	24
Married	39	35
Education		
Grades 1–8	37	32
Grades 9–11	31	35
High school or more	32	33
Annual family income below $12 000	81	63**
Health		
Personal history of breast problems	9	9
Family history of breast cancer	8	9
1 or more medications taken regularly	79	81
More than 3 chronic health problems	27	29

Access		
Regular physician		
Obstetrician/gynecologist	2	3
Other	87	86
No regular physician	11	10
Has health insurance coverage	83	84
No. of medical visits in past year		
4 or more	49	65
1–3	42	28
None	9	7**
Physician recommendation in past year	39	51**
Attitudes/knowledge barriers		
Perceived susceptibility to breast cancer	23	11**
Perceived severity of breast cancer	29	24
Breast cancer knowledge (7 items)		
High (6–7 correct)	23	28
Medium (4–5 correct)	34	41
Low (0–3 correct)	43	31**
Barriers to mammography (18 items)		
Low (0–4)	44	52
High (5–17)	56	48*
Social norms and support		
Support for breast cancer screening		
High	19	16
Medium	38	49
Low	43	35**
Spirituality (5 items)		
Low (0–2)	26	25
High (3–5)	74	75

Difference of the differences?

TABLE 4—Self-Reported Mammography Use in Past 2 Years From Baseline to Follow-Up: North Carolina Breast Cancer Screening Program, 1993–1994 and 1996–1997

					Difference of differences,[a] %				
					Unadjusted		Adjusted		
	No.	Baseline, %	Follow-Up, %	Increase, %	Estimate	95% CI	Estimate	95% CI	*P*
Overall									
Intervention	387	41	58	17	6	(-1, 14)	7[b]	(0, 14)	.05
Comparison	409	56	67	11					
Low income[c]									
Intervention	279	37	59	22	12	(2, 21)	11[b]	(2, 21)	.02
Comparison	235	49	60	11					
High income									
Intervention	66	56	59	3	-6	(-18, 7)	1[d]	(-10, 11)	.92
Comparison	138	73	82	9					

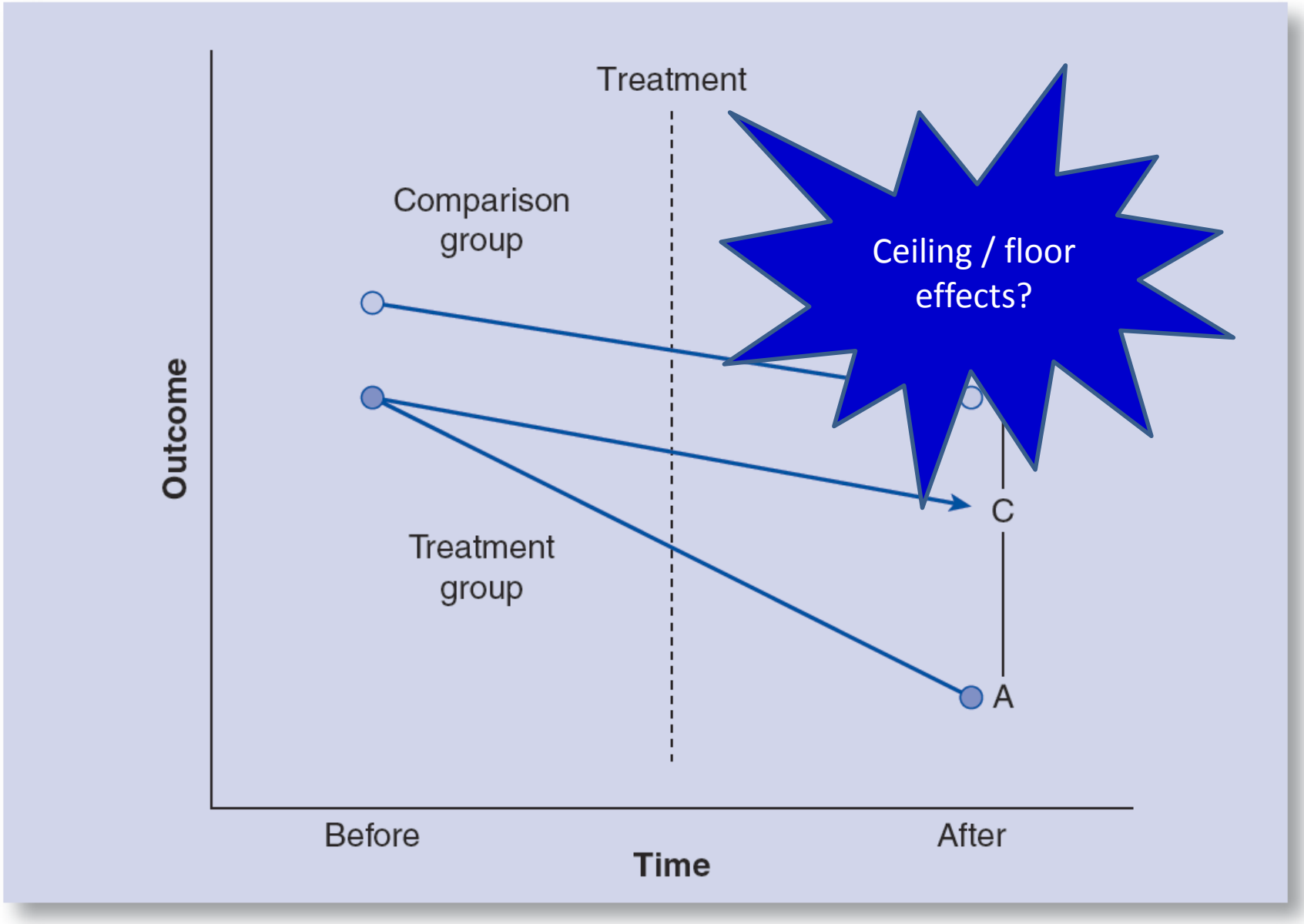
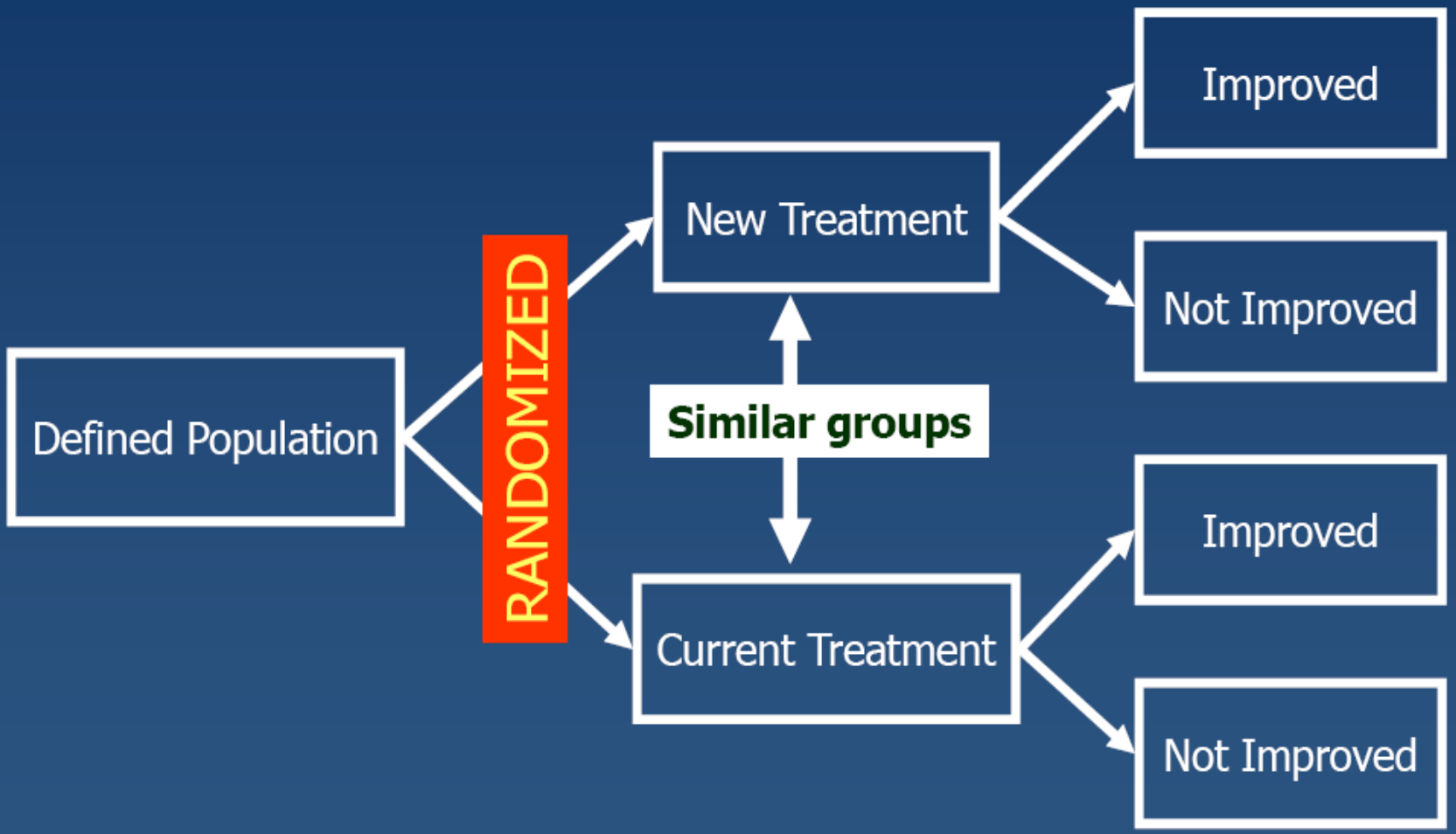


Figure 13.5 Difference-in-Differences Assumptions

Methodology of Randomized Controlled Clinical Trials

Design of Randomized Trial



Randomized Controlled Clinical Trials

- New drugs are only introduced into medical care when they have been shown to be effective.
- This involves a formal assessment using a rigorous study design: the randomized controlled clinical trials
- This method provides the best way of determining whether a proposed new treatment represents an advance on the current best treatment.

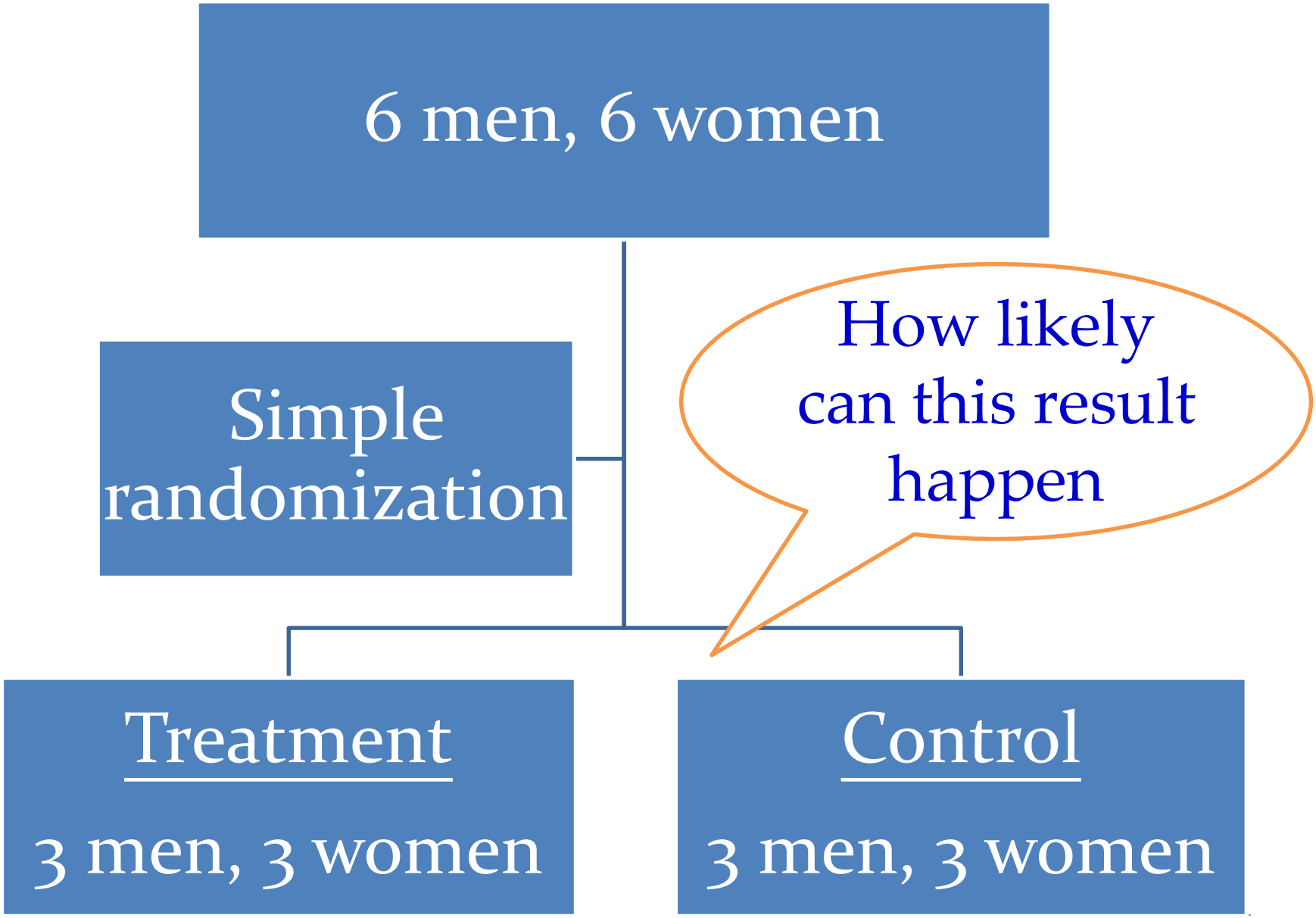
Design of a Randomized Clinical Trial

- Definition of study groups
- Allocating treatments
- Ethics
- Outcome assessment

Allocating Treatments (Randomization)

- A key feature of the trial is that once a set of patients has been recruited, they are allocated randomly to one of the two treatments being compared.
- This helps guard against systematic differences occurring between the two groups.

- If the two groups of patients are not similar at entry to the study, a fair comparison between the treatments cannot be made


6 men, 6 women
Simple randomization
How likely can this result happen
Treatment
3 men, 3 women
Control
3 men, 3 women

Stratified Randomization

- Stratification might be useful in small trials in which it can avert severe imbalances on prognostic factors. It will confer adequate balance (on the stratified factors) and probably slightly more statistical power and precision.
- The gain from stratification becomes minimal, however, once the number of participants per group is more than 50.

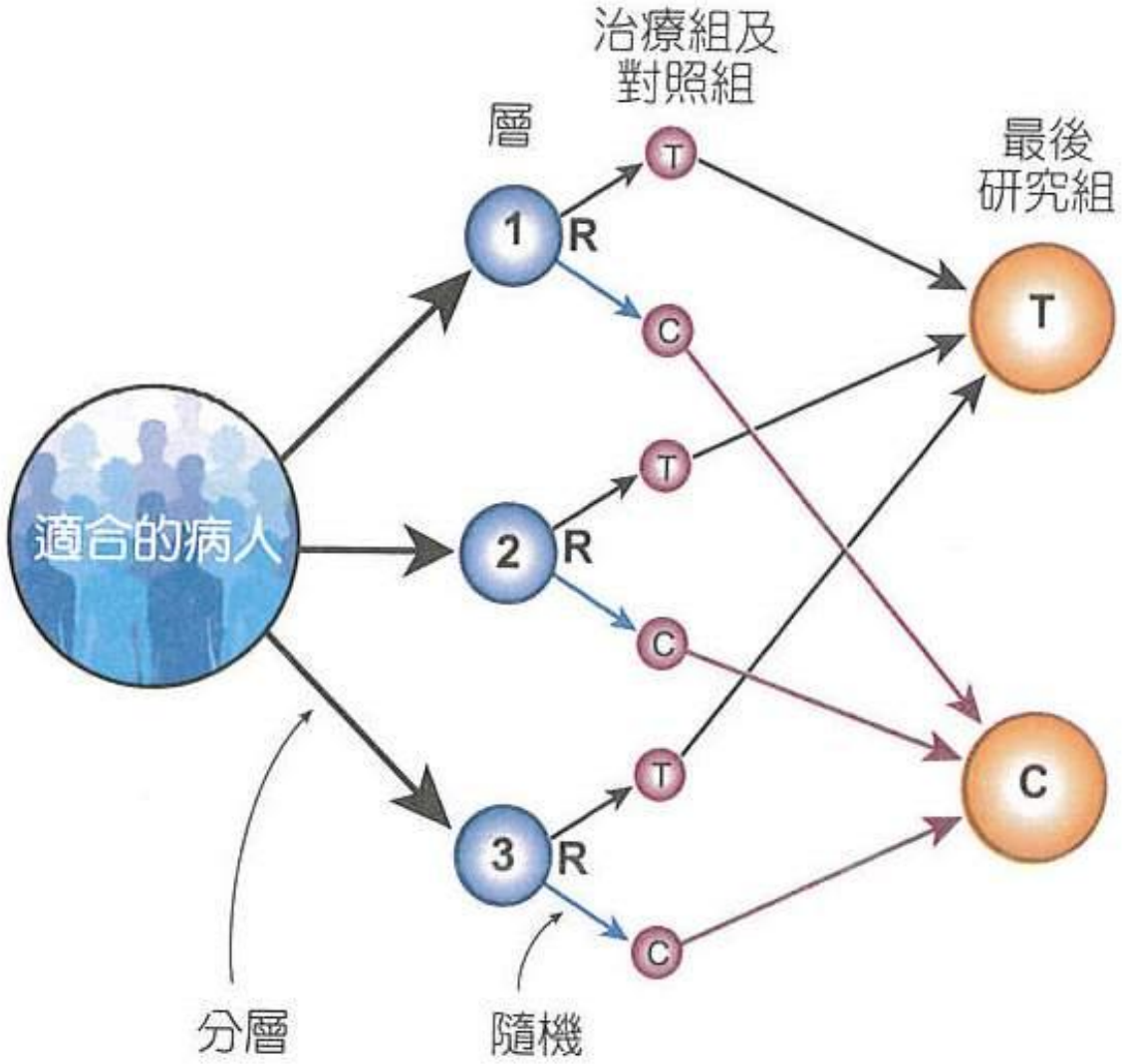
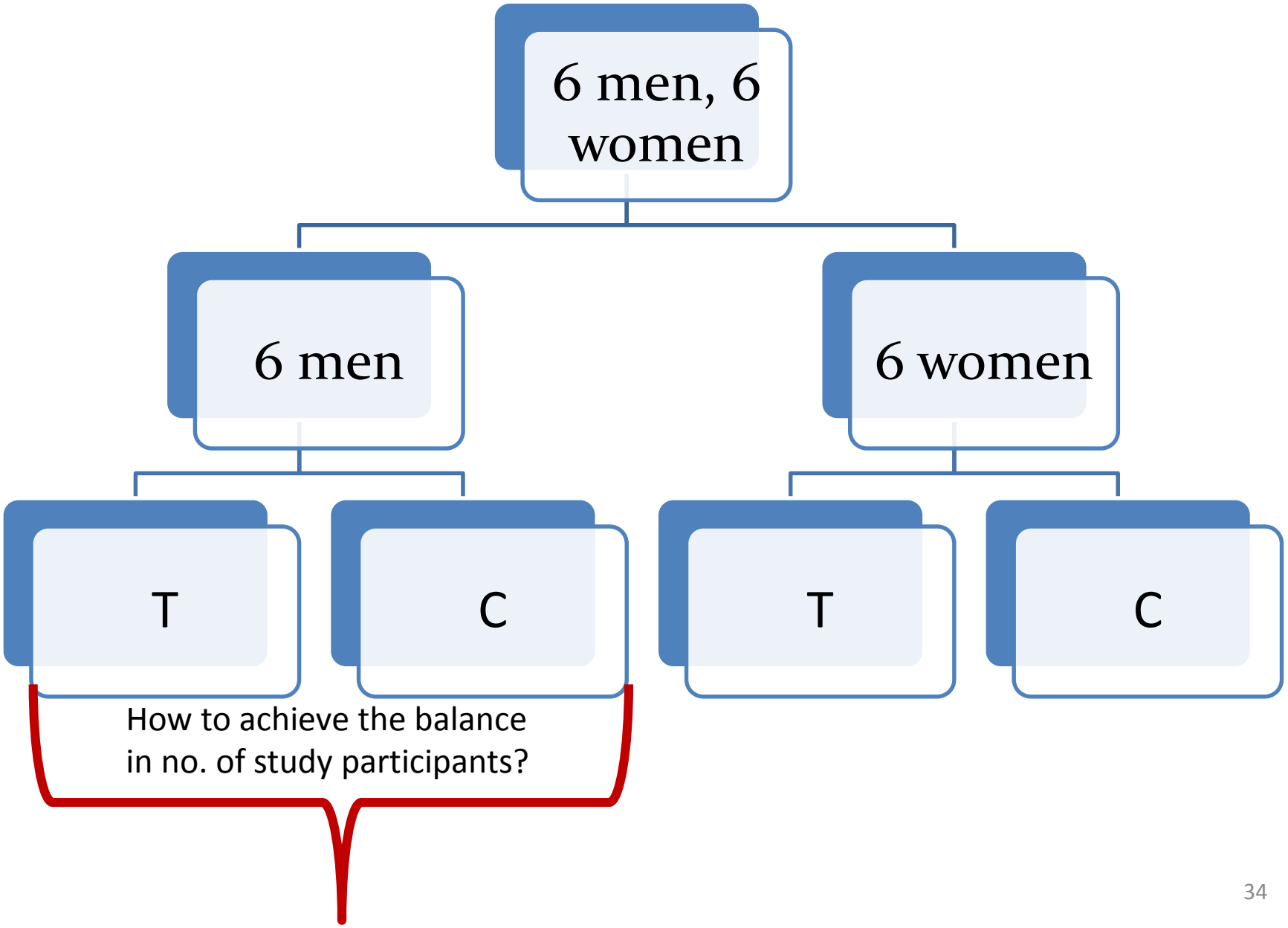


圖 9.5■ 分層隨機分派圖。T＝治療組，C＝對照組，R＝隨機


6 men, 6 women
6 men
6 women
T
C
T
C
How to achieve the balance in no. of study participants?

Randomized Block Design, RBD

- Used to ensure close balance of the numbers in each group at any time during the trial.
- After a block of every 10 participants, for example, five would be allocated to each arm of the trial.
- For example: a block size of 4
  - Sequence is determined randomly
  - TTCC, TCTC, CCTT.....

Stratified / Blocked Randomization

- Define strata
- Randomization is performed within each stratum and is usually blocked
- Example: Age, < 40, 41-60, ≥60; Sex, M, F
  Total number of strata = 3 x 2 = 6

Age	Male	Female
40	ABBA, BAAB, ...	BABA, BAAB, ...
41-60	BBAA, ABAB, ...	ABAB, BBAA, ...
≥60	AABB, ABBA, ...	BAAB, ABAB, ..

An Example of Clinical Trial

Articles

Placebo-controlled multicentre randomised trial of interferon β-1b in treatment of secondary progressive multiple sclerosis

*European Study Group on Interferon β-1b in Secondary Progressive MS**

THE LANCET • Vol 352 • November 7, 1998

Methods

- In this multicentre, double-masked, randomised, placebo-controlled trial, outpatients with SP-MS having scores of 3.0-6.5 on the Expanded Disability Status Scale (EDSS) received either 8 million IU interferon β -1b every other day subcutaneously, or placebo, for up to 3 years.

- The primary outcome was the time to confirmed progression in disability as measured by a 1.0 point increase on the EDSS, sustained for at least 3 months, or a 0.5 point increase if the baseline EDSS was 6.0 or 6.5.
- A prospectively planned interim analysis of safety and efficacy of the intention-to-treat population was done after all patients had been in the study for at least 2 years.

Figure 1. Trial profile

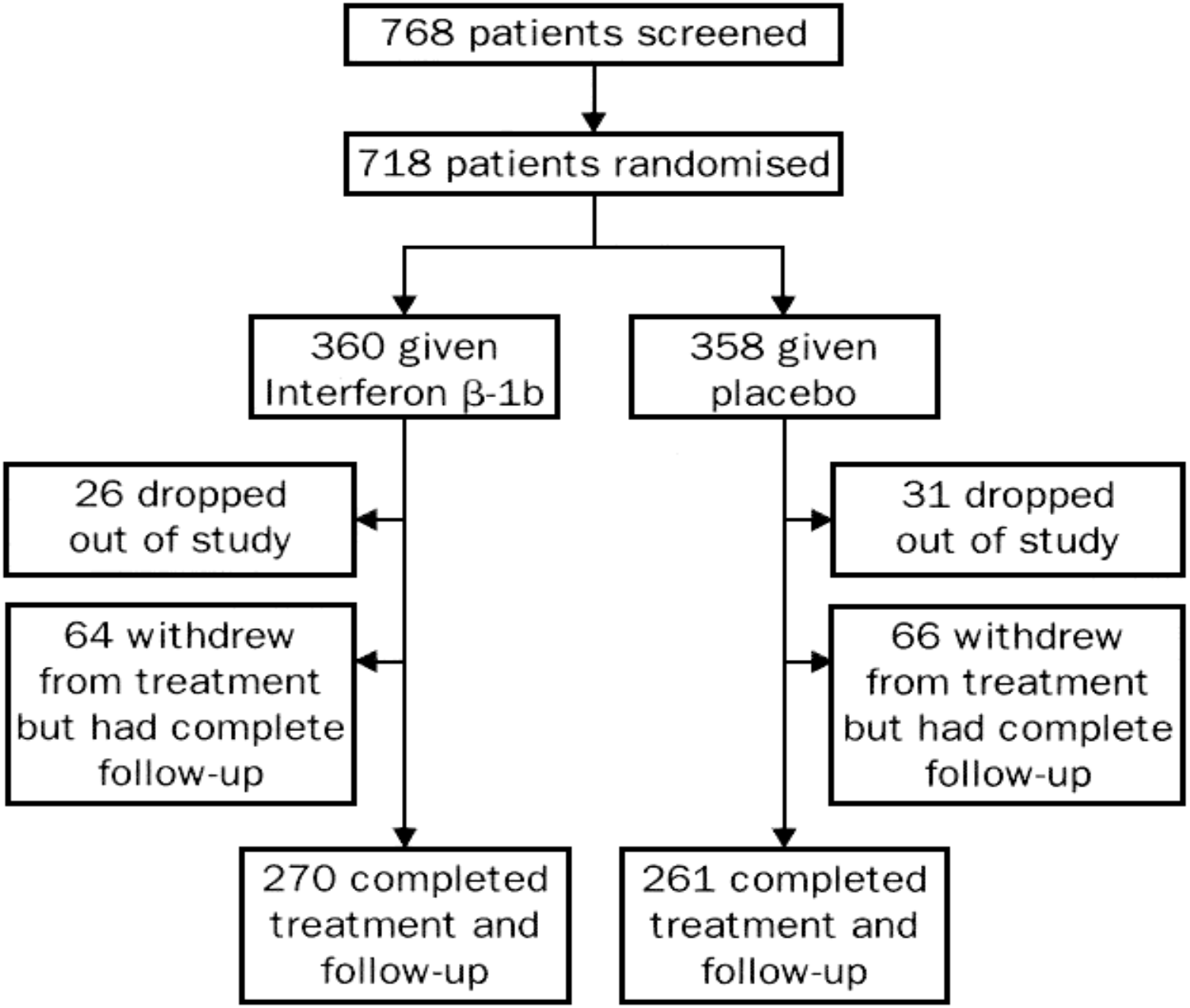


Table 2. Patient population (baseline characteristics)

	Placebo (n=358)*	Interferon β-1 (n=360)*
Mean age (SD, years)	40·9 (7·2)	41·1 (7·2)
Women	64·2%	58·1%
Mean disease duration (SD, years)	13·4 (7·5)	12·8 (6·6)
Mean time since diagnosis of relapsing risk MS (SD, years)	8·2 (6·1)	8·1 (5·6)
Mean time since evidence of progressive deterioration (SD, years)	3·8 (3·4)	3·8 (2·7)
Mean time since diagnosis of SP-MS (SD, years)	2·1 (2·2)	2·2 (2·4)
Mean EDSS at baseline (SD)	5·2 (1·1)	5·1 (1·1)
EDSS by category		
⩽3·5	47 (13·1%)	67 (18·6%)
4·0–5·5	142 (39·7%)	140 (38·9%)
⩾6·0	169 (47·2%)	153 (42·5%)
Patients without relapse in 2 years before study†	101 (28·2%)	115 (31·9%)

*No significant differences between treatment groups (p>0·05).

†Data missing for seven patients (four placebo, three interferon β–1b) who were included in the subgroup of patients without relapse.

Table 3. Reasons for dropping out of study and stopping treatment

	Placebo (n=358)	Interferon β-1b (n=360)
Reason for drop out		
Adverse event, laboratory deviation	4 (1·1%)	5 (1·4%)
Progression of disease	10 (2·8%)	5 (1·4%)
Death*	1 (0·3%)	3 (0·8%)
Lost to follow-up	4 (1·1%)	8 (2·2%)
Other	12 (3·4%)	5 (1·4%)
Total	31 (8·7%)	26 (7·2%)
Reason for stopping treatment (including drop outs)		
Adverse events†	15(4·2%)	45 (12·5%)
Illness, independent from trial medication	3 (0·8%)	0
Patient uncooperative/rejects treatment*	19 (5·3%)	8 (2·2%)
Deviation from trial protocol	0	3 (0·8%)
Inefficacy of trial medication†‡	44 (12·3%)	23 (6·4%)
Death	0	2 (0·6%)
Pregnancy	0	1 (0·3%)
Other	16 (4·5%)	8 (2·2%)
Total	97 (27·1%)	90 (25·0%)

*One suicide in each group. †$p<0·05$ two-sided Fisher's Exact Test.
‡Includes two patients (one placebo, one interferon β-1b) who died after premature discontinuation of treatment.

Table 5. Results of secondary and tertiary efficacy variables

Efficacy variable	Placebo (n=358)	Interferon β-1b (n=360)	p
Proporportion of patients with confirmed EDSS progression*	49·7%	38·9%	0·0048
Loss of mobility			
Time to becoming wheelchair-bound			0·0133
Estimated probability of not becoming wheelchair-bound			
Year 1	0·90	0·96	0·0129
Year 2	0·81	0·89	0·0094
Year 3	0·66	0·77	0·0133
Mean EDSS			
At endoint	5·84	5·57	0·0750
Change at endpoint†	0·60	0·47	0·0299
Mean annual relapse rate			
Overall	0·64	0·44	0·0002
Year 1	0·82	0·57	0·0095
Year 2	0·47	0·35	0·0201
Year 3	0·35	0·24	0·1624
Median time to first relapse (days)	403	644	0·0030
Proportion of patients with moderate or severe	53·1%	43·6%	0·0083

*Patients lost to follow-up counted as not progressed.
†Endpoint minus baseline.

What can RCTs do, and what RCTs can't do?

Research Paper

W.-H. Hou et al. • 157 (2016) 1954–1959

PAIN®

Dipeptidyl peptidase-4 inhibitor use is not associated with elevated risk of severe joint pain in patients with type 2 diabetes: a population-based cohort study

Wen-Hsuan Hou[a,b,c,d], Kai-Cheng Chang[e], Chung-Yi Li[f,g], Huang-Tz Ou[e,*]

Table 1

Comparison of baseline characteristics between DPP4i and non-DPP4i users before and after propensity score matching.

	Before propensity score matching			After propensity score matching		
	Non-DPP4i users (n = 48,190)	DPP4i users (n = 4743)	*P*	Non-DPP4i users (n = 4743)	DPP4i users (n = 4743)	*P*
Male, %	60.37	55.42	<0.0001	54.07	55.42	0.1869
Age, y	54.4 ± 12.6	55.6 ± 12.9	<0.0001	55.4 ± 13.1	55.6 ± 12.9	0.5596
CCI	2.9 ± 1.4	2.9 ± 1.4	0.0003	3.0 ± 1.5	2.9 ± 1.4	0.1846
DCSI	0.5 ± 1.1	0.8 ± 1.3	<0.0001	0.8 ± 1.4	0.8 ± 1.3	0.1140
Duration of diabetes	1.7 ± 2.7	4.1 ± 3.5	<0.0001	4.1 ± 3.6	4.1 ± 3.5	0.6323
Depression, %	2.25	2.51	0.2651	2.55	2.51	0.8960
Anxiety, %	7.48	7.52	0.9251	7.69	7.52	0.7568
Obesity, %	1.63	3.46	<0.0001	3.65	3.46	0.6176
Connective tissue disease, %	0.54	0.67	0.2210	0.67	0.67	1.0000
Infectious or crystal arthropathy, %	0.13	0.11	0.6436	0.11	0.11	1.0000
Rheumatism, %	22.39	22.78	0.5431	22.36	22.78	0.6234
Dorsopathy, %	18.82	18.61	0.7165	18.98	18.61	0.6363
Osteopathy, chondropathy, and acquired musculoskeletal deformities, %	2.07	2.40	0.1332	2.63	2.40	0.4711
Fracture, %	3.25	2.70	0.0385	2.55	2.70	0.6530
Dislocation, %	0.27	0.19	0.2947	0.17	0.19	0.8082
Sprain and strain, %	13.24	11.95	0.0119	12.28	11.95	0.6148

CCI, charlson comorbidity index; DCSI, diabetes complications severity index; DPP4i, dipeptidyl peptidase-4 inhibitor.