Critical appraisal of evidence

Mahasti Alizadeh MD,MPH

Associate professor of community medicine

Tabriz University of Medical Sciences

مراحل عملی تصمیم گیری مبتنی بر شواهد

۱. تهیه یک سوال قابل جستجو و قابل پاسخ
۲. جستجوی منابع بر اساس سوال طراحی شده
۳. نقد و بررسی مقالات بدست آمده از جستجو از نظر اعتبار و قابلیت بکار گیری نتایج آن
۴. بکار گیری نتایج در عملکرد حرفه ای

Your question (PICO)

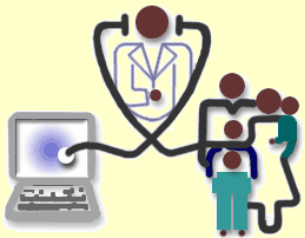
study

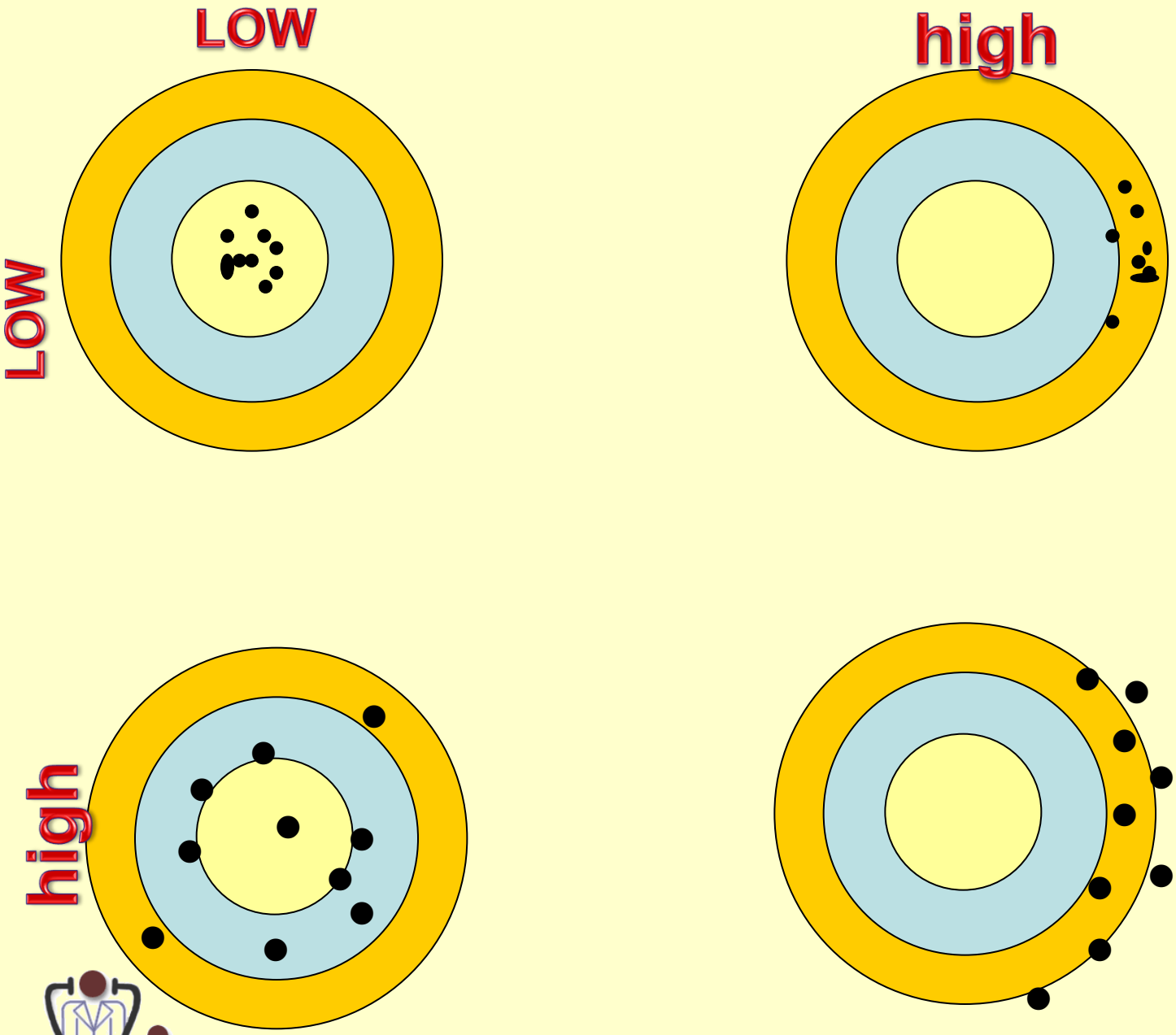
How well was the study done?

What do the results mean?

تا چه حد مطالعه درست انجام شده است؟

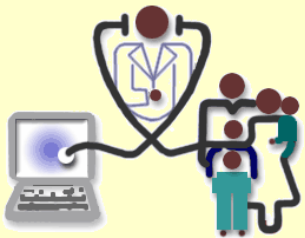
- تا چه اندازه نتایج مطالعه تحت تاثیر سوگرایی و عوامل مخدوش کننده قرار می گیرند؟
- سوگرایی چیست؟
- تفاوت سوگرایی با خطای تصادفی چیست؟



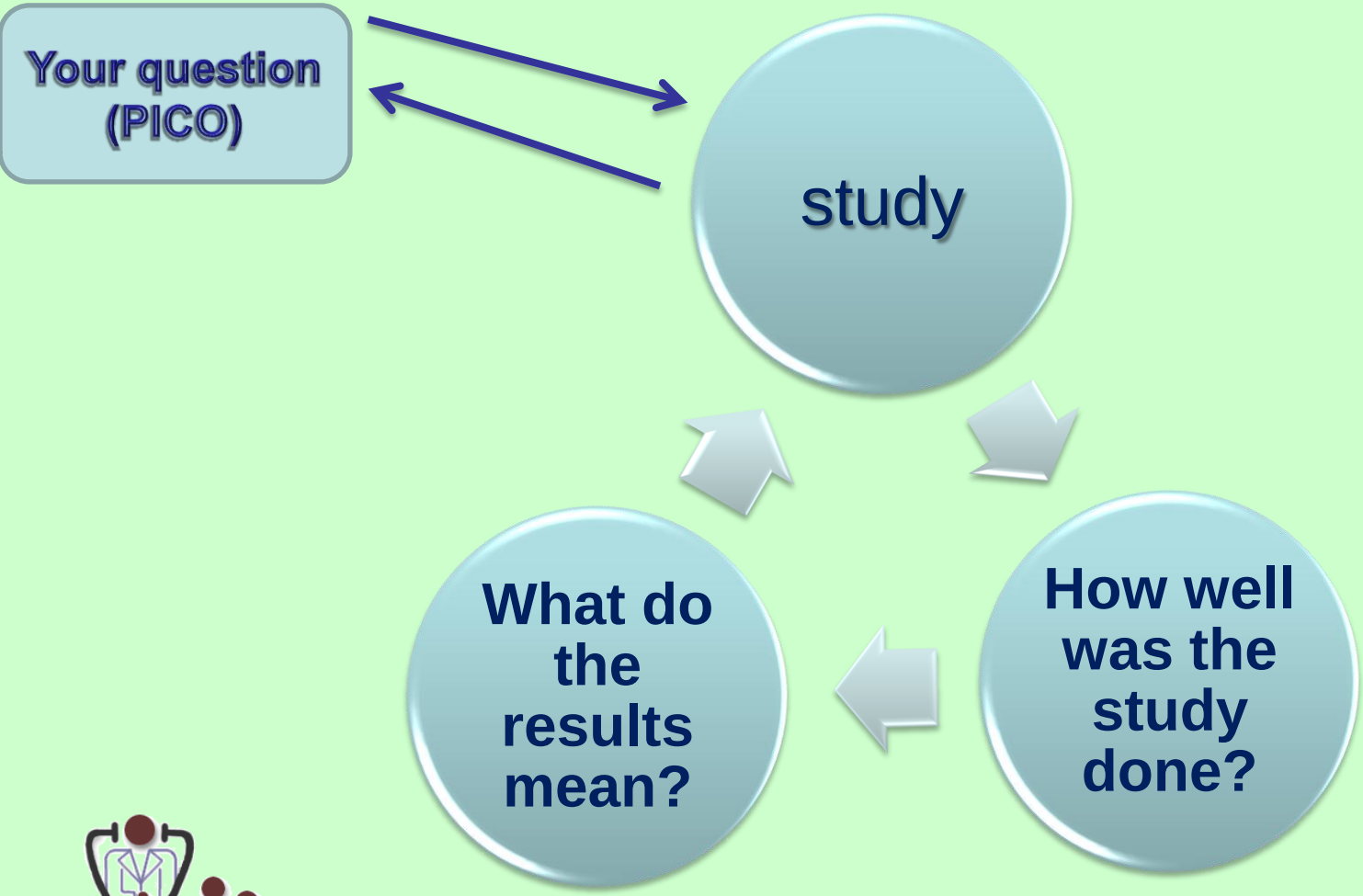

LOW
high
LOW
high
Random error
Bias

Confounding factor
عامل مخدوش کننده

- هر عاملی بجز عامل مورد سنجش که بر نتایج و پیامد مورد نظر تاثیر گذارد

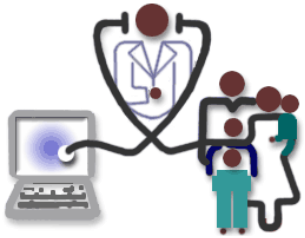


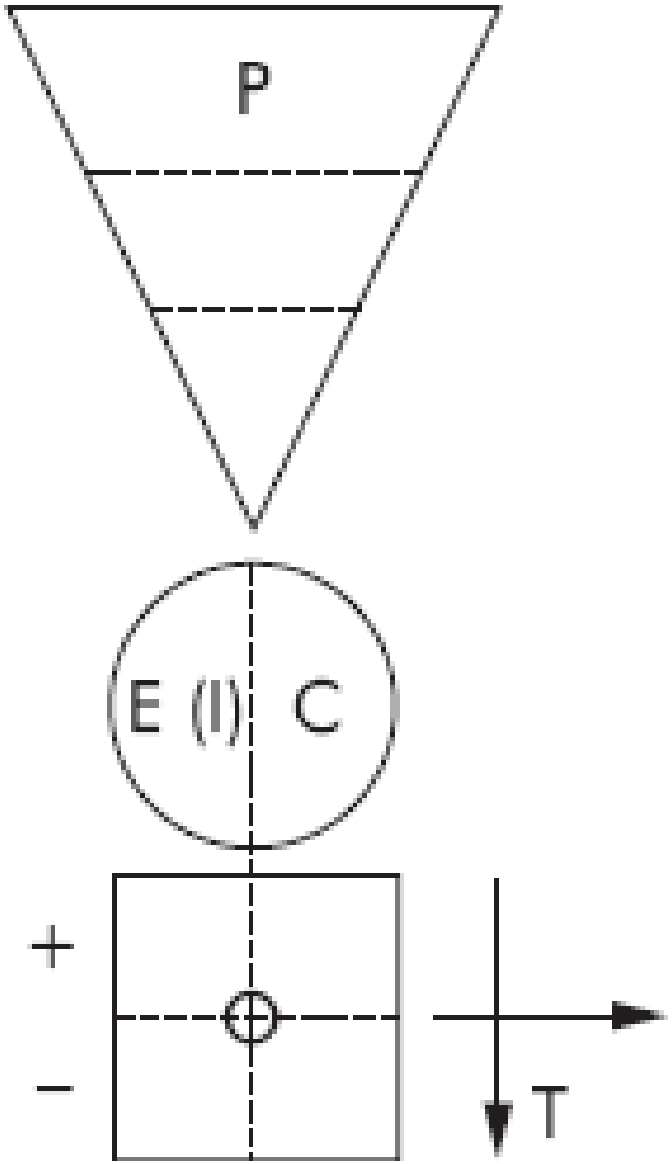
Critical appraisal



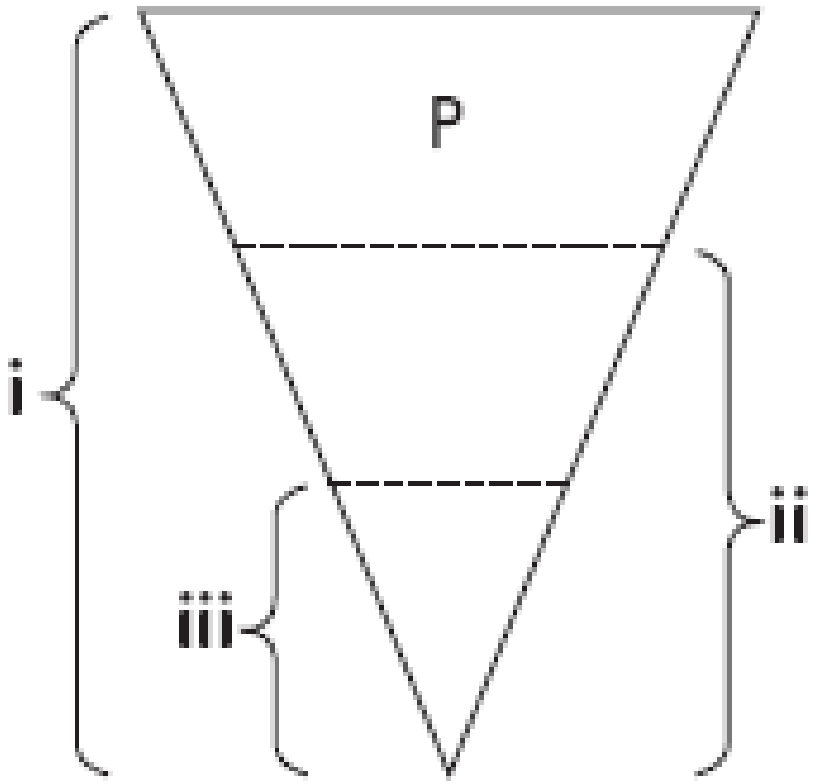
How well was the study done?

- انتخاب افراد و ورود آنها به مطالعه تا چه حد درست بوده است؟ (P)
- آیا افراد تحت مطالعه بدرستی در دو یا چند گروه قرار گرفته اند؟ (I and C)
- تا چه حد افراد گروههای تحت مطالعه بدرستی بررسی و پیگیری شده اند؟(IandC)
- پیامد یا Outcome تا چه حد درست اندازه گیری شده است؟(O)





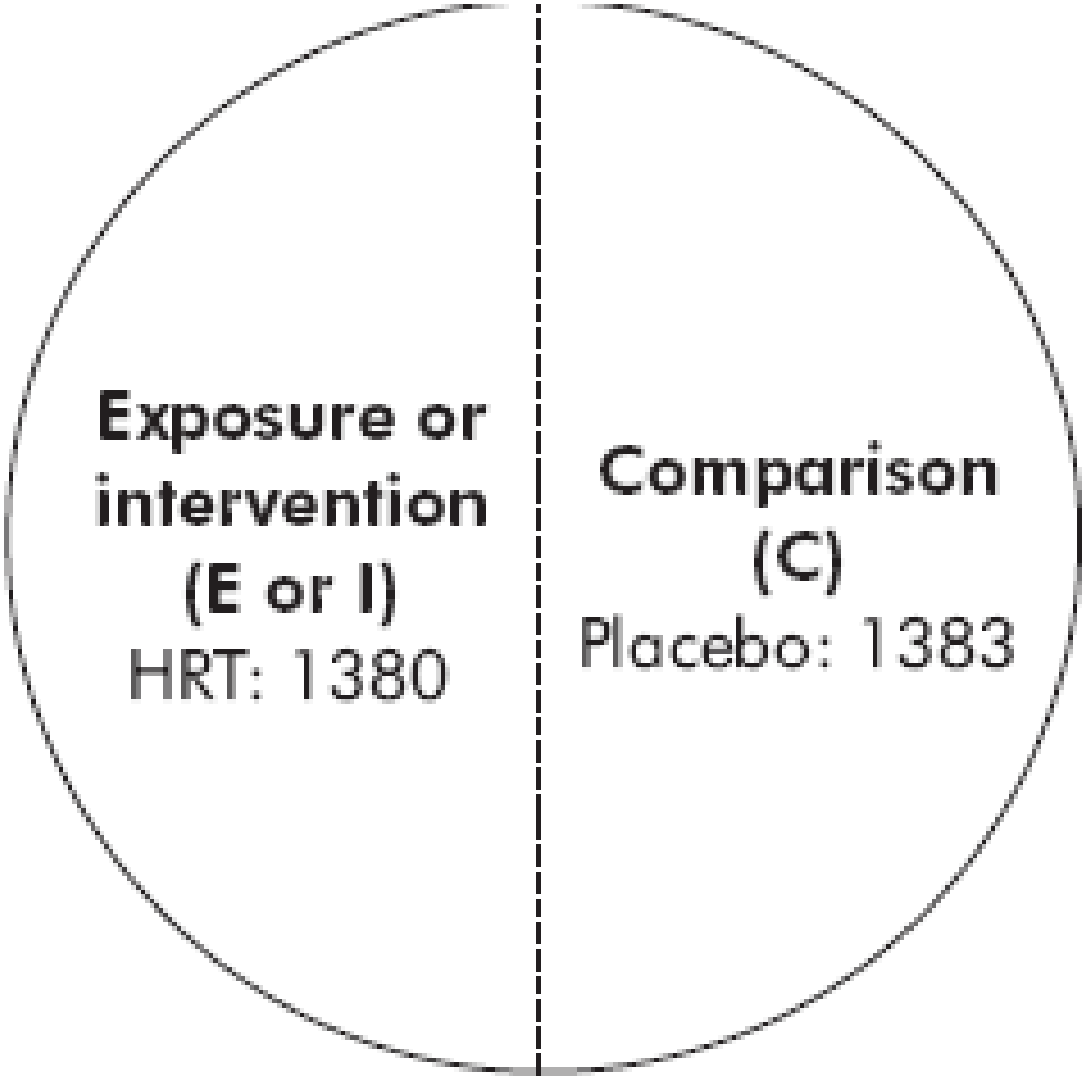
The GATE frame.



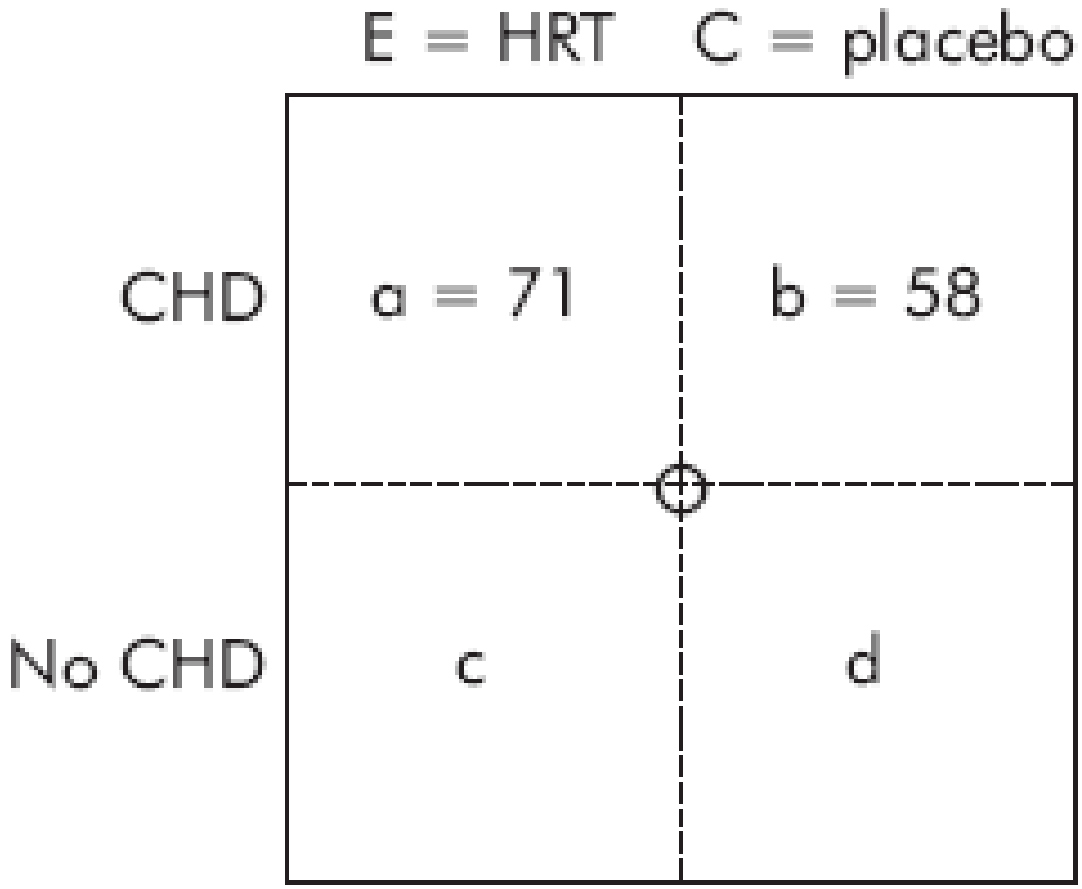
i. Source Population:
from multiple sources.
68 561 screened by phone.

ii. Eligible Population:
post-menopausal, <80 yrs,
with CHD. No MI < 6 months
or HRT < 3 months
TG > 3.39 mmol/l.
Number not given.

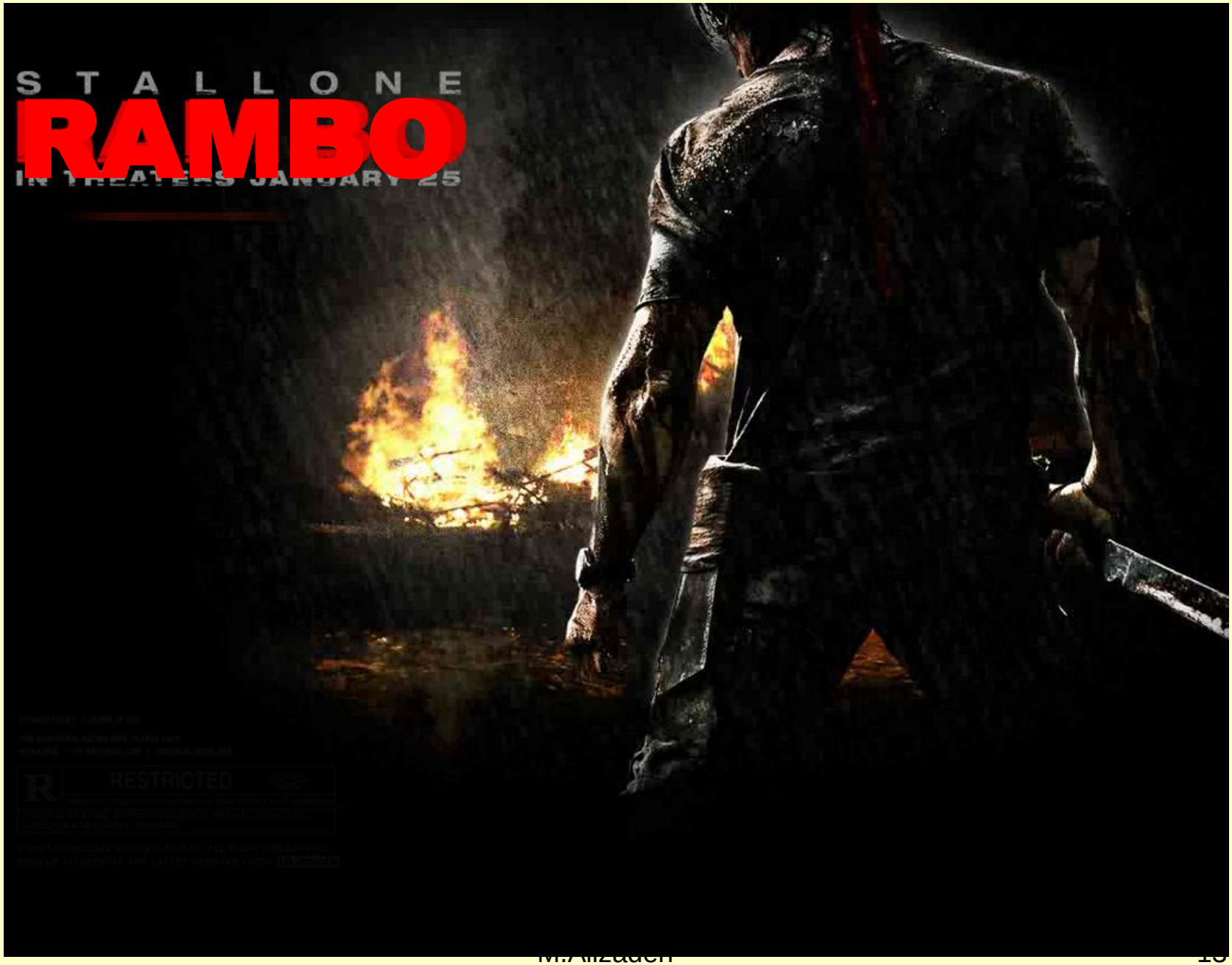
iii. Participant Population:
All eligibles invited. Of 3463
attending 2nd screen.
2763 randomised.

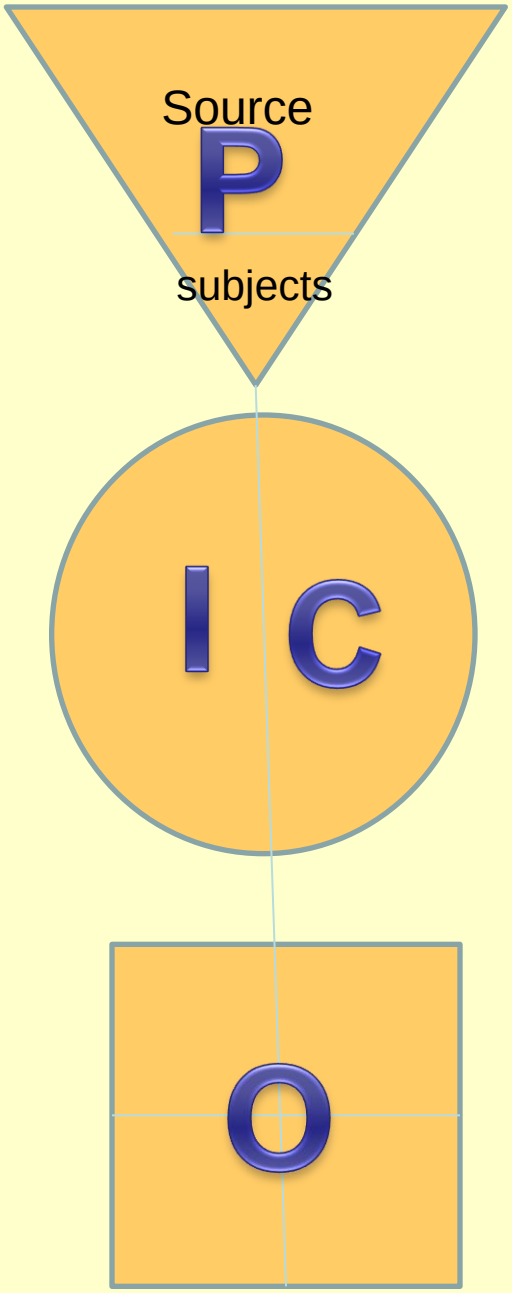


Exposure (E) & comparison (C)—the HERS example.



Outcomes (O)—the HERS example (fatal CHD).


STALLONE
RAMBO
IN THEATERS JANUARY 25
R
RESTRICTED



آیا افراد انتخاب شده نماینده گروه هدف هستند؟

Recruitment
Representative

تخصیص به گروهها تصادفی و مخفی بوده است؟
(Random and cocealed)

Allocation
Adjustment

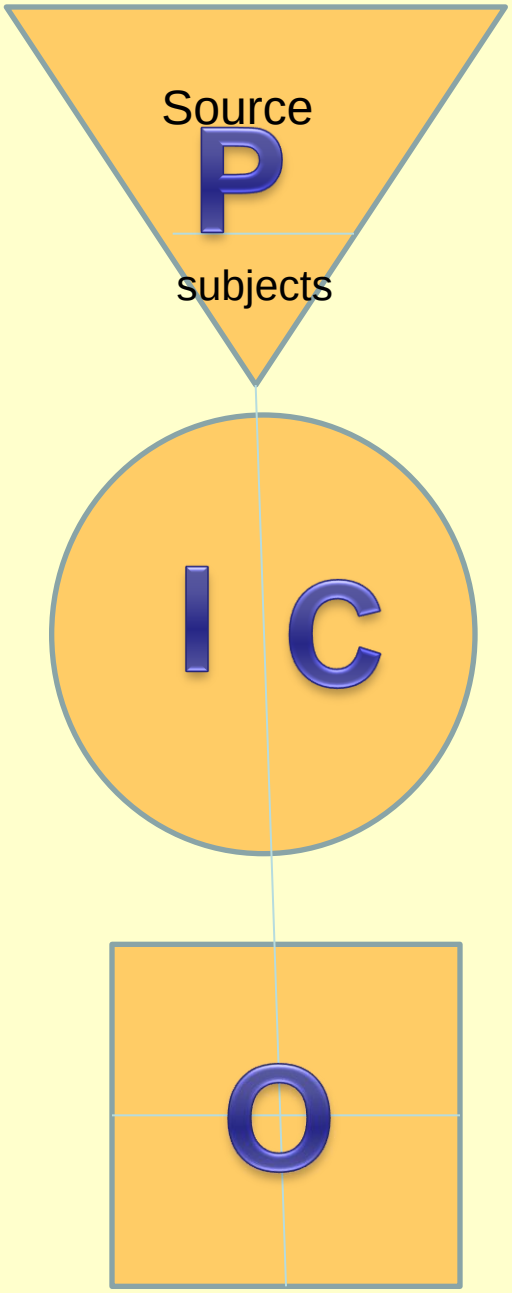
بجز مداخله مورد نظر ایاا در تمام موارد گروهها یکسان بوده اند؟

Maintenance

Measurement
Blinding
Objective

آیا سنجش پیامد مورد انتظار بصورت عینی و blindبوده؟

R A M M b o



آیا افراد انتخاب شده نماینده گروه هدف هستند؟

Recruitment
Representative

تخصیص به گروهها تصادفی و مخفی بوده است؟
(Random and cocealed)

Allocation

بجز مداخله مورد نظر ایاا در تمام موارد گروهها یکسان بوده اند؟

Maintenance

Measurement
Blinding
Objective

آیا سنجش پیامد مورد انتظار بصورت عینی و blindبوده؟

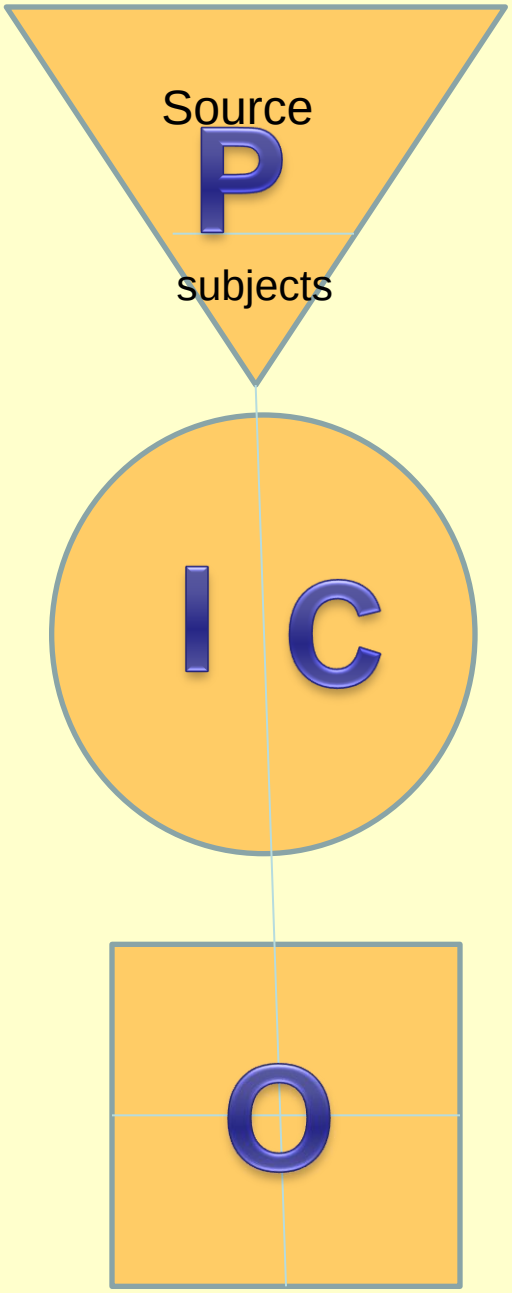
R A M M b o

Recruitment Representative

- افرادی که وارد مطالعه می شوند نماینده جمعیت مورد نظر باشند.
- انتخاب تصادفی از کل جامعه آماری
- در مقاله جمعیت مرجع بطور شفاف توضیح داده شود
- انتخاب تعداد مناسب و کافی نمونه (حجم نمونه کافی)
- بررسی مطالعه نمونه!!!!

RECRUITMENT

In cooperation with the university administration, we drew a random sample of 13 000 full-time undergraduates aged 17 to 24 years in March 2007 using survey recruitment procedures described in detail elsewhere.[21,22] In summary, 4 weeks after the start of the first semester, all students were sent a letter by mail, followed by an e-mail containing a hyperlink to a Web questionnaire. Students were informed that the study concerned "the experiences of tertiary students with alcohol, and their views about certain aspects of drinking." Up to 3 reminder e-mails were sent in the following weeks. Students were offered the opportunity to win 1 of 40 A$100 gift vouchers for participating.



آیا افراد انتخاب شده نماینده گروه هدف هستند؟

Recruitment
Representative

تخصیص به گروهها تصادفی و مخفی بوده است؟
(Random and cocealed)

Allocation
Adjustment

بجز مداخله مورد نظر ایاا در تمام موارد گروهها یکسان بوده اند؟

Maintenance

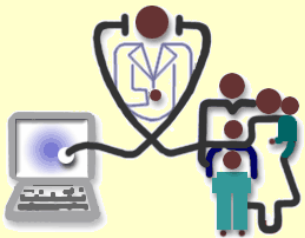
Measurement
Blinding
Objective

آیا سنجش پیامد مورد انتظار بصورت عینی و blindبوده؟

R A M M b o

Allocation

- گروهها از هر نظر بجز مداخله مورد نظر با هم همسان باشند
- اگر چنین نباشد هر تفاوتی در نتیجه می تواند ناشی از غیر همسانی متغیرهای مخدوش کننده باشد نه مداخله
- با تخصیص تصادفی به گروهها تا حدی می توان به این همسانی نزدیک شد



Is allocation concealed?

Methods: A randomized controlled trial was conducted at an Australian university in 2007. Invitations were sent to 13 000 undergraduates (age range, 17-24 years) to complete a Web-based Alcohol Use Disorders Identification Test. Of 7237 students who responded, 2435 scored in the hazardous/harmful range (≥8) and were randomized, and 2050

RANDOMIZATION AND INTERVENTIONS

Respondents who scored 8 or more on the AUDIT and had exceeded the Australian National Health and Medical Research Council's guideline for acute risk (binge drinking: 4 standard drinks for women, 6 for men) in the last 4 weeks were considered to have screened positive for unhealthy alcohol use. They were randomly assigned by the Web server software to the control group (screening only) or to the intervention group.

Adjustment

- Characteristics of the study groups
- Sample paper!!!!

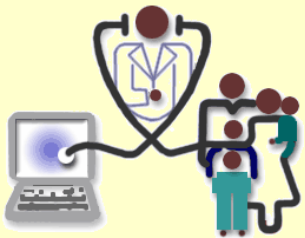
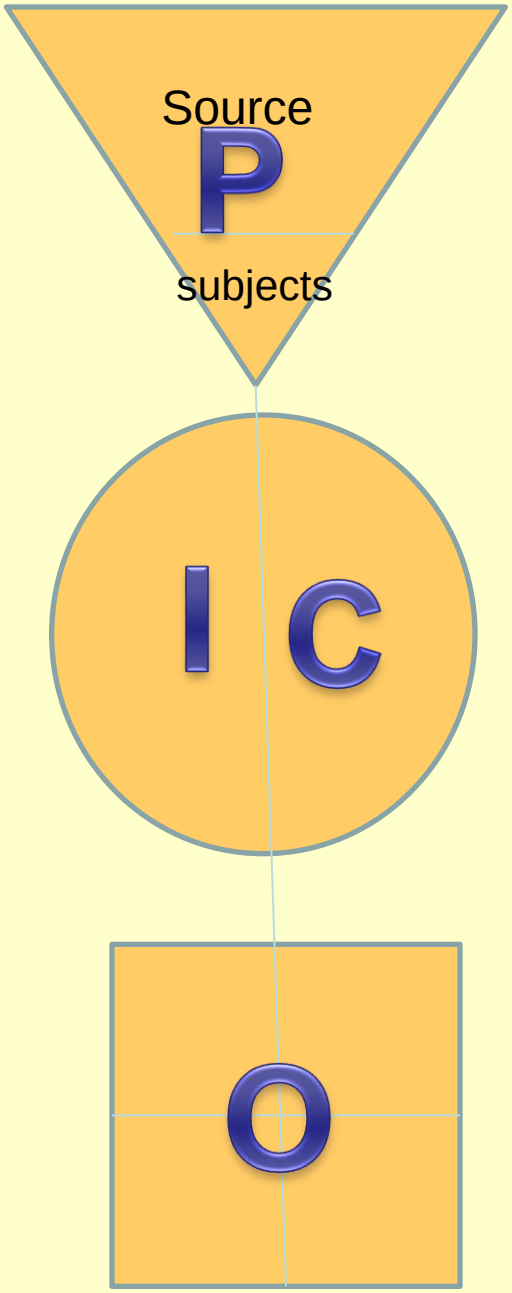


Table 1. Demographic Characteristics and Alcohol Use of the Study Groups at Baseline

	Control (n=1184)	Intervention (n=1251)
Women, %	45.5	45.1
Age, mean (SD), y	19.7 (1.8)	19.7 (1.8)
Living arrangement, %		
With a parent or guardian	64.3	66.4
Shared house/student housing	29.7	27.2
Other	5.9	6.3
AUDIT score, mean (SD)	14.3 (5.1)	14.2 (5.1)
Drinking summary data[a]		
Drinks alcohol 2 or more times per week, %	62.2	59.3
No. of standard drinks per typical drinking occasion, mean (SD)	8.5 (4.6)	8.5 (5.2)
Alcohol dependence subscale score, mean (SD)[b]	1.8 (1.8)	1.8 (1.8)

Abbreviation: AUDIT, Alcohol Use Disorders Identification Test.
[a] AUDIT items 1 and 2.
[b] Sum of scores for AUDIT items 4 through 6.



آیا افراد انتخاب شده نماینده گروه هدف هستند؟

Recruitment
Representative

تخصیص به گروهها تصادفی و مخفی بوده است؟
(Random and cocealed)

Allocation
Adjustment

بجز مداخله مورد نظر ایاا در تمام موارد گروهها یکسان بوده اند؟

Maintenance

Measurement
Blinding
Objective

آیا سنجش پیامد مورد انتظار بصورت عینی و blindبوده؟

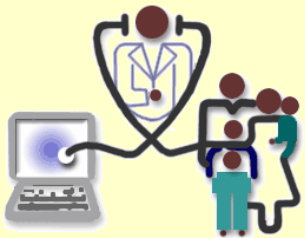
R A M M b o

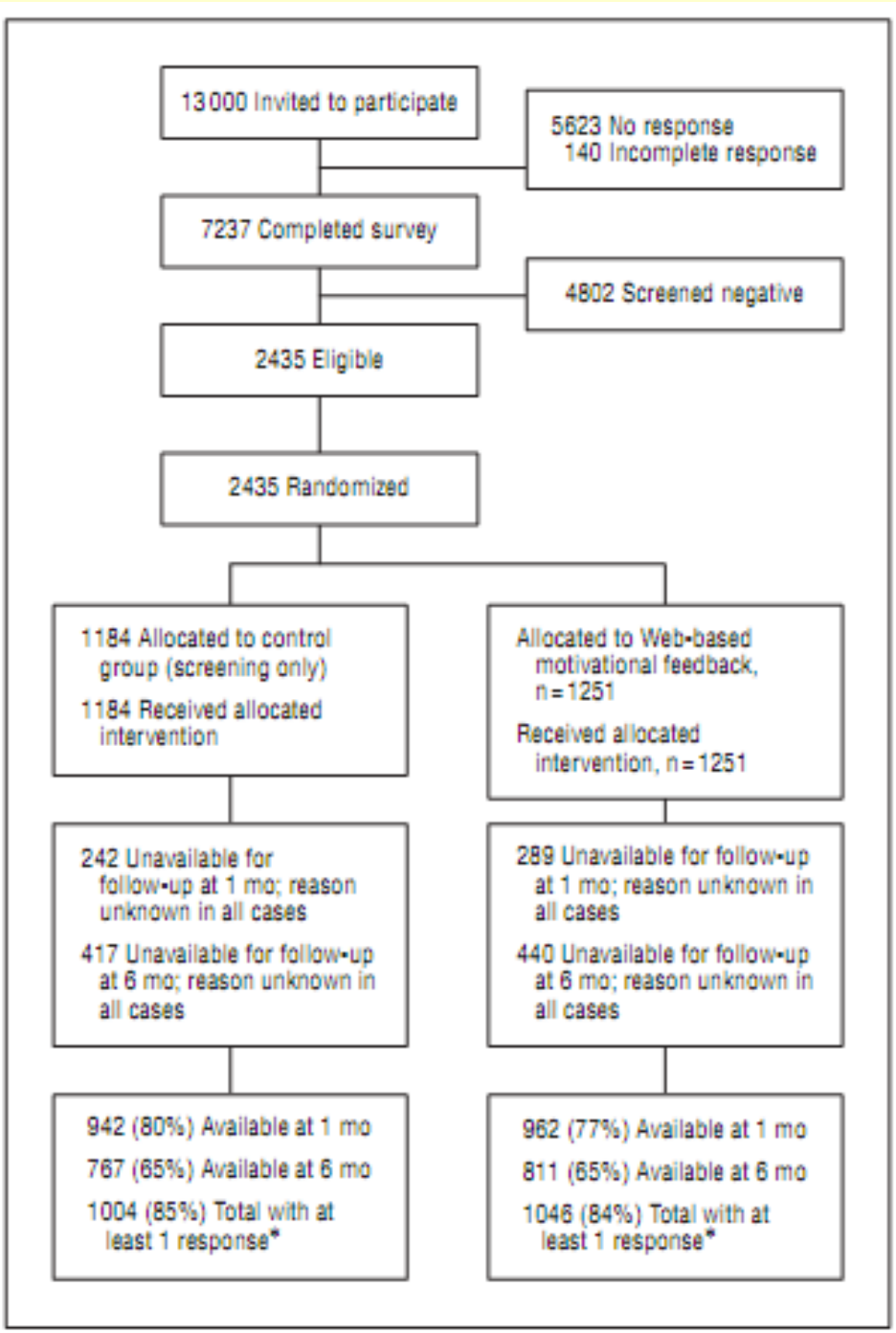
Maintenance

- آیا شرایط قابل مقایسه دوگروه تا آخر مطالعه حفظ شده است؟
- آیا یک گروه درمان یا مداخله اضافی دیگری دریافت نکرده است؟

- Adequate follow-up
- تعداد افراد تا پایان مطالعه
- قرار گرفتن در همان گروه
- مدت پیگیری(تا بروز پیامد مورد نظر)

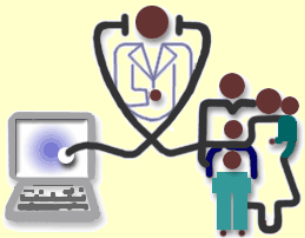
Intention to treat analysis

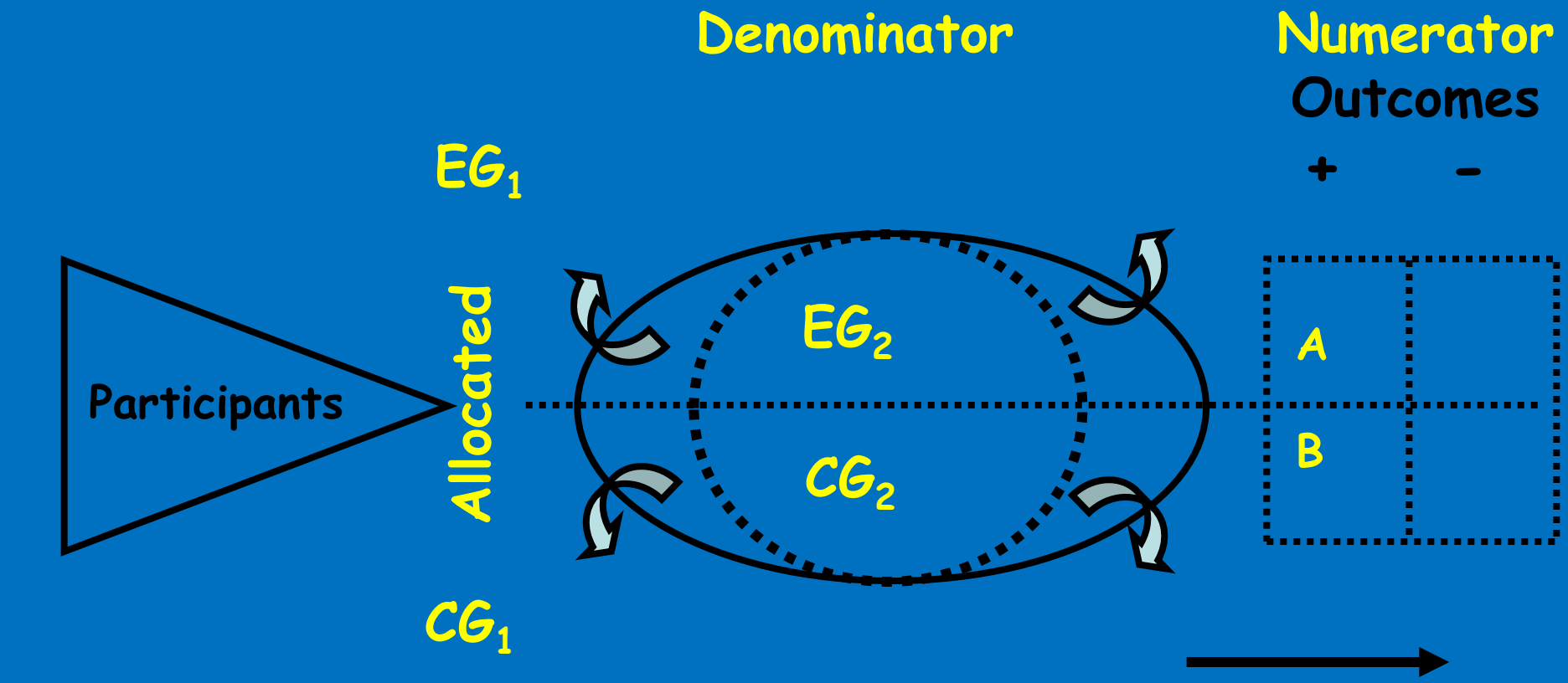


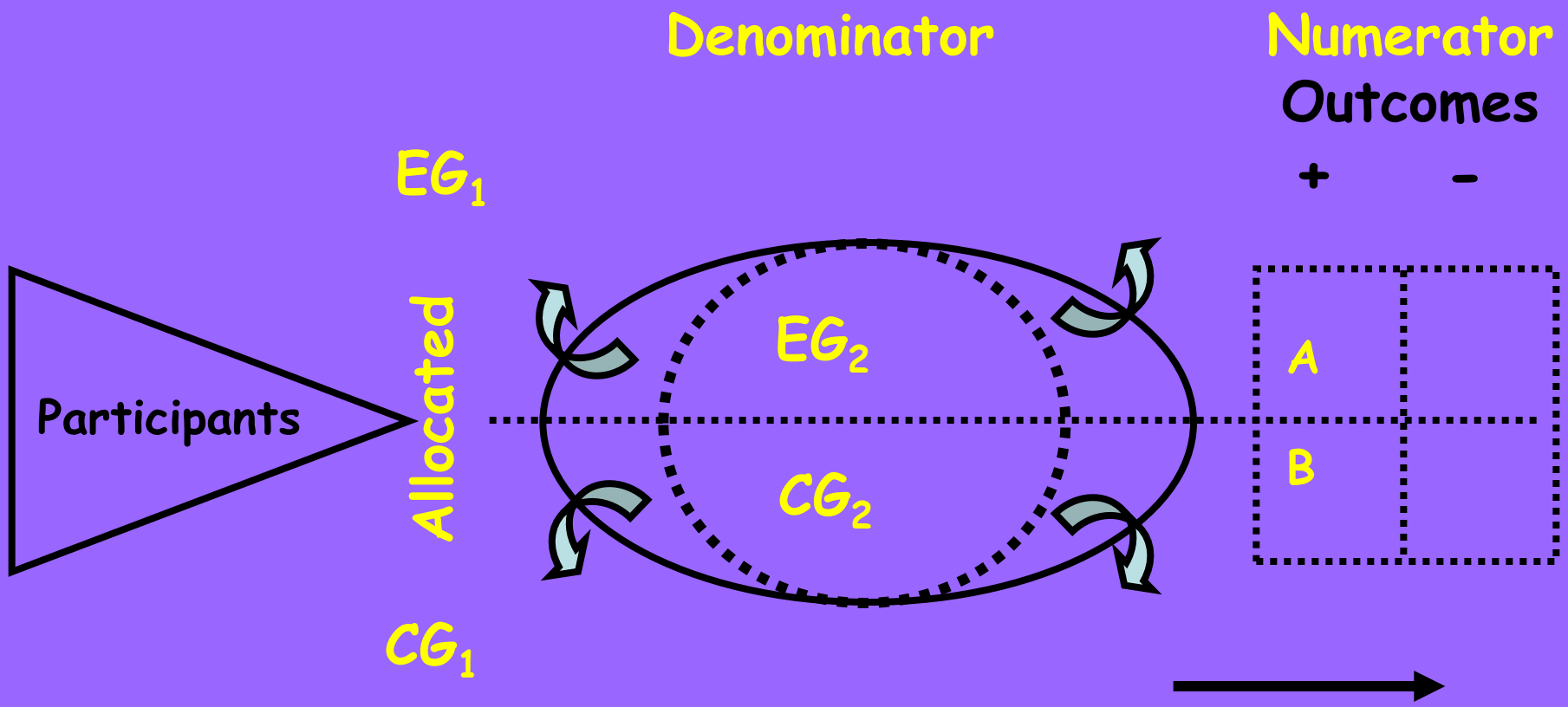

13 000 Invited to participate
5623 No response
140 Incomplete response
7237 Completed survey
4802 Screened negative
2435 Eligible
2435 Randomized
1184 Allocated to control group (screening only)
1184 Received allocated intervention
Allocated to Web-based motivational feedback, n = 1251
Received allocated intervention, n = 1251
242 Unavailable for follow-up at 1 mo; reason unknown in all cases
417 Unavailable for follow-up at 6 mo; reason unknown in all cases
289 Unavailable for follow-up at 1 mo; reason unknown in all cases
440 Unavailable for follow-up at 6 mo; reason unknown in all cases
942 (80%) Available at 1 mo
767 (65%) Available at 6 mo
1004 (85%) Total with at least 1 response*
962 (77%) Available at 1 mo
811 (65%) Available at 6 mo
1046 (84%) Total with at least 1 response*

Intention to treat analysis

- افراد در همان گروهی مورد تجزیه و تحلیل قرار می گیرند که از همان ابتدا در آن گروه بودند.




Denominator
Numerator
Outcomes
+
-
EG1
Participants
Allocated
EG2
CG2
CG1
A
B

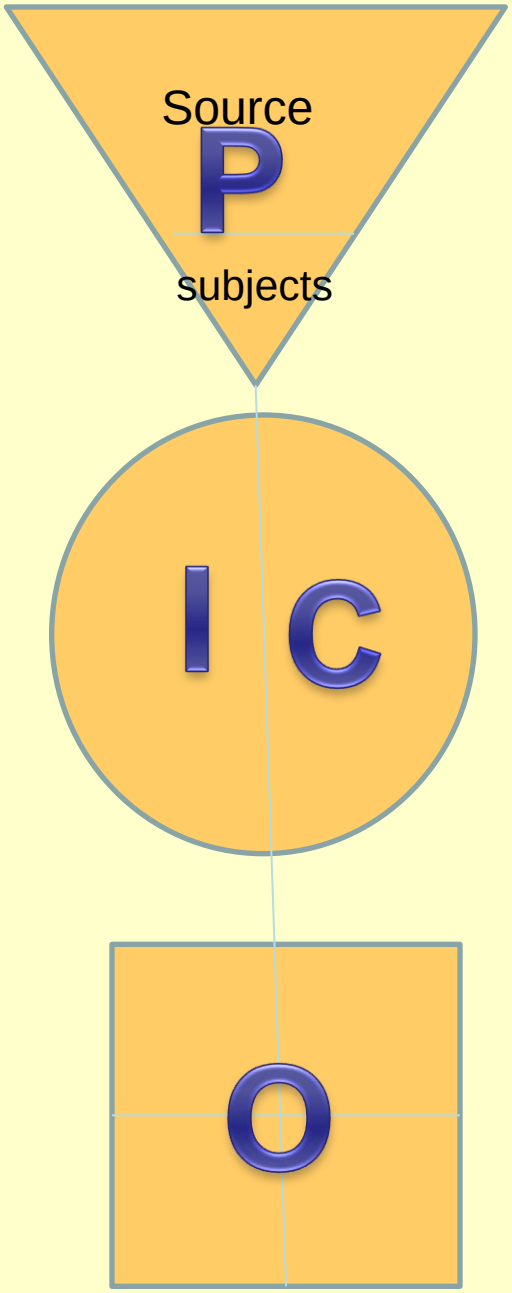


EGO = A/EG_1 (Intention To Treat [ITT] analyses)
or EGO = A/EG_2 (On Treatment [OT] analyses)

STATISTICAL ANALYSIS

The primary outcomes and the AREAS scores were analyzed with negative binomial regression for panel data using the Stata 10.1 xtnbreg procedure (StataCorp, College Station, Texas).[28] The APS scores were analyzed with linear regression after log transformation using xtmixed (StataCorp). For the proportions of students exceeding the Australian National Health and Medical Research Council's drinking guidelines, we used generalized linear mixed models with the xtlogit procedure (StataCorp).[29,30] All models included a random intercept to account for clustering within participant[29] as well as fixed effects for group, follow-up assessment, and their interaction.[31] The interaction term allowed differences in the intervention effect between follow-up assessments. The results are presented as rate ratios, difference in regression coefficients, and odds ratios, respectively, along with overall tests of intervention with 2 *df*.

Participants were analyzed in the group to which they were randomized (intention to treat), and all participants were followed up regardless of their compliance with the intervention. We describe patterns of missing values and compare those observed and those missing in terms of baseline characteristics. In this study, some participants were missing at the 1 month follow-up asssessment, others at 6 months, and some at both time points (no postrandomization data available). We compared baseline AUDIT scores, age, sex, and treatment ob-



آیا افراد انتخاب شده نماینده گروه هدف هستند؟

Recruitment
Representative

تخصیص به گروهها تصادفی و مخفی بوده است؟
(Random and cocealed)

Allocation
Adjustment

بجز مداخله مورد نظر ایاا در تمام موارد گروهها یکسان بوده اند؟

Maintenance

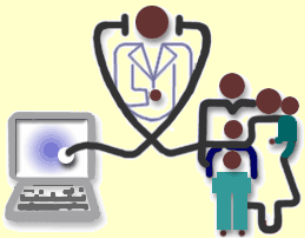
Measurement
Blinding
Objective

آیا سنجش پیامد مورد انتظار بصورت عینی و blindبوده؟

R A M M b o

Measurement (Blinding,Objective)

- Double blind
- Objective outcomes
- Sample paper!!!!



Were patients, clinicians, and study personnel kept blind to treatment?

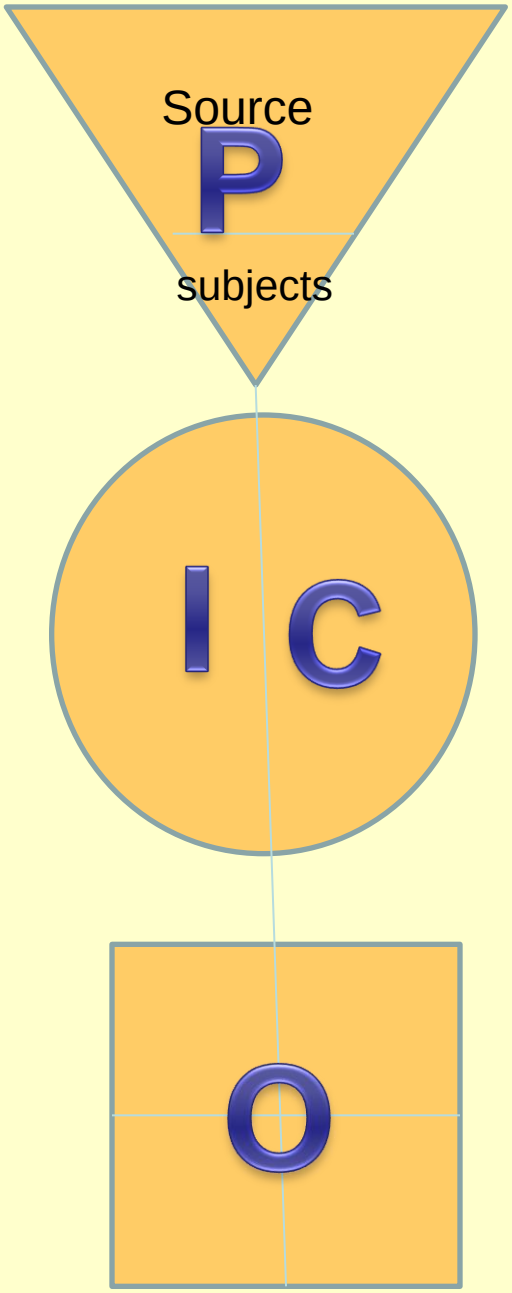
Fig. 3 **A double-blind placebo-controlled clinical trial for CAM therapies.**

baseline, a form of booster intervention.

Participants were blind to the true nature of the study, which was presented as a series of surveys, in accordance with ethical approval, provided by the research ethics committee of Curtin University, Perth, Australia. Researchers were blind to participants' group allocation.

OUTCOMES AND FOLLOW-UP

Methods: A randomized controlled trial was conducted at an Australian university in 2007. Invitations were sent to 13 000 undergraduates (age range, 17-24 years) to complete a Web-based Alcohol Use Disorders Identification Test. Of 7237 students who responded, 2435 scored in the hazardous/harmful range (≥8) and were randomized, and 2050 (84%) completed at least 1 follow-up assessment. Intervention was 10 minutes of Web-based motivational assessment and personalized feedback. Controls received only screening. Follow-up assessments were conducted at 1 and 6 months with observers and participants blinded to allocation. Outcome measures were drinking frequency, typical occasion quantity, overall volume, number of personal problems, an academic problems score, prevalence of binge drinking, and prevalence of heavy drinking.



آیا افراد انتخاب شده نماینده گروه هدف هستند؟

Recruitment
Representative

تخصیص به گروهها تصادفی و مخفی بوده است؟
(Random and cocealed)

Allocation
Adjustment

بجز مداخله مورد نظر ایاا در تمام موارد گروهها یکسان بوده اند؟

Maintenance

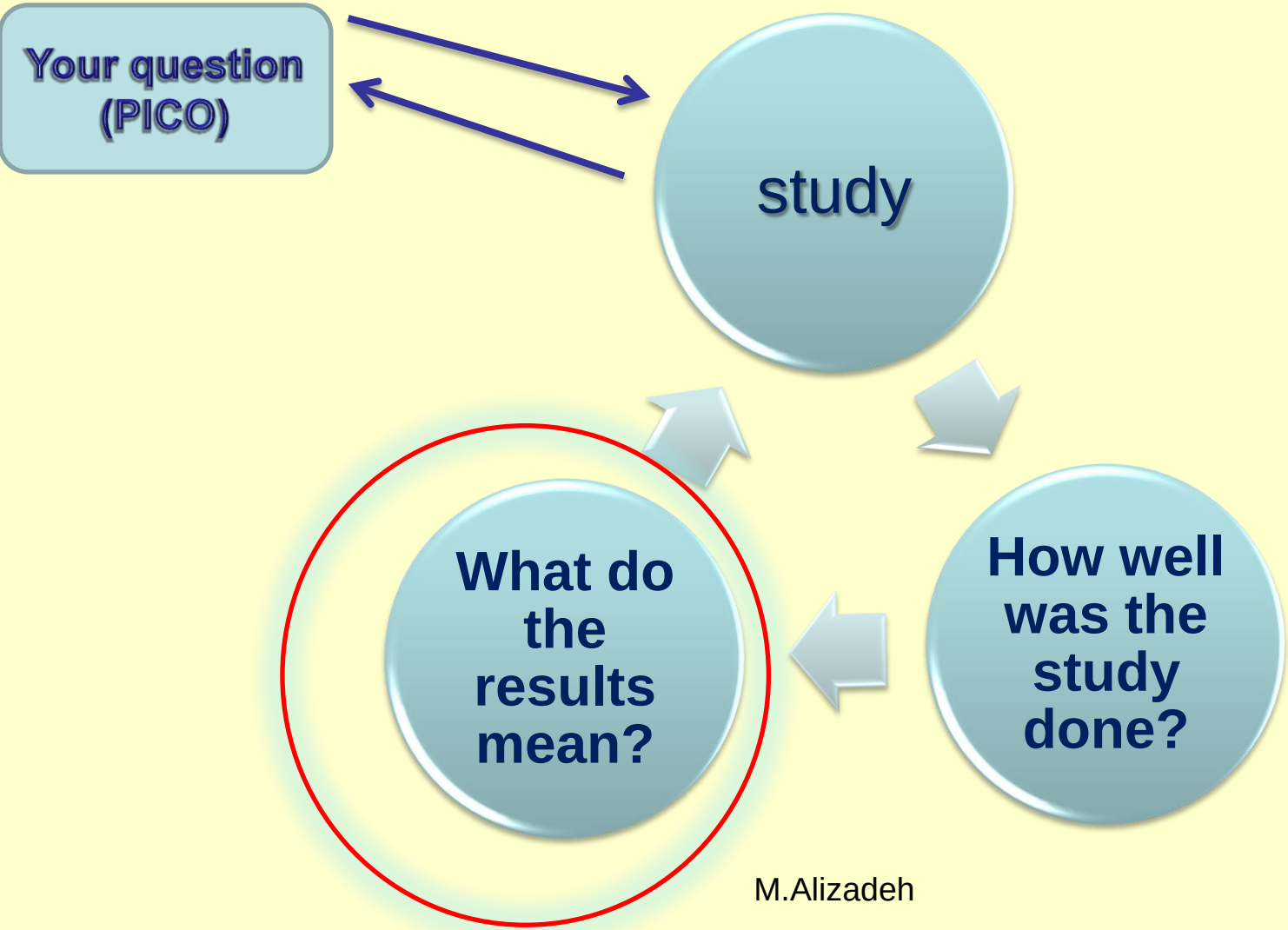
Measurement
Blinding
Objective

آیا سنجش پیامد مورد انتظار بصورت عینی و blindبوده؟

R A M M b o


KSB

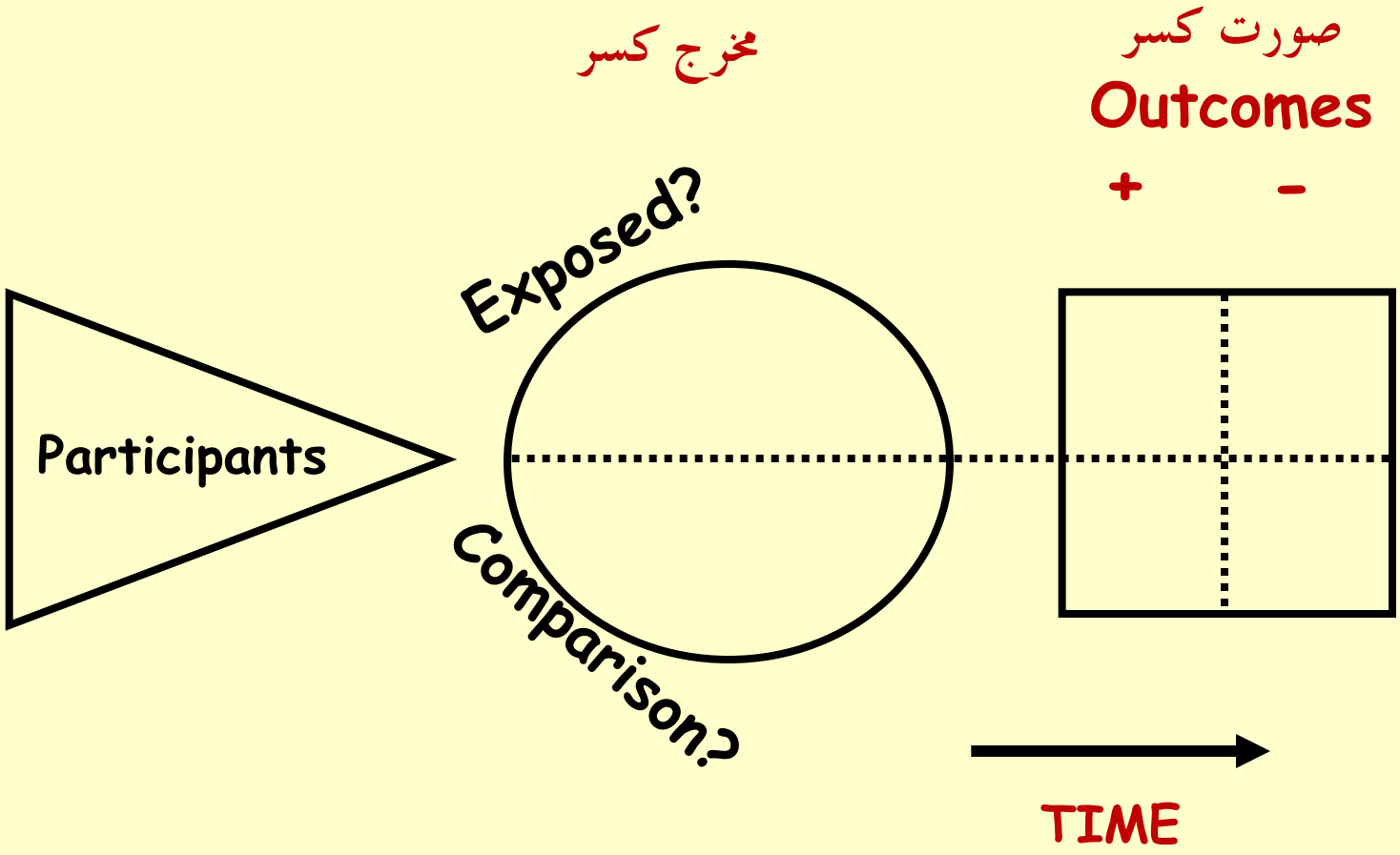
Critical appraisal



OCCURRENCE =

Outcomes / Population / Time

Numerator/Denominator/Time


مخرج کسر
صورت کسر
Outcomes
+
-
Exposed?
Participants
Comparison?
TIME

Relative Risk(RR)

- خطر نسبی
- احتمال بروز یک پیامد یا نتیجه در گروه مداخله چند برابر گروه مقایسه است؟
- RR=1 تفاوتی بین دو گروه وجود ندارد
- RR<1 درمان احتمال بروز پیامد را کم می کند
- RR>1 درمان احتمال بروز پیامد را بیشتر می کند

$$RR=\frac{\text{ریسک بروز پیامد در گروه مداخله}}{\text{ریسک بروز پیامد در گروه مقایسه}}$$

Variable	Treatment Effects: RR[a] (95% CI) Intervention/Control	
	Without Multiple Imputation for Missing Values	With Multiple Imputation for Missing Values[b]
Primary outcomes		
Frequency of drinking		
1 mo	0.89 (0.83 to 0.94)	0.90 (0.85 to 0.96)
6 mo	0.91 (0.85 to 0.97)	0.91 (0.86 to 0.96)
Typical occasion quantity	0.93 (0.88 to 0.98)	0.93 (0.88 to 0.99)
1 mo	0.96 (0.91 to 1.02)	0.94 (0.89 to 0.99)
6 mo		
Volume consumed	0.83 (0.78 to 0.90)	0.85 (0.79 to 0.92)
1 mo	0.89 (0.82 to 0.96)	0.86 (0.81 to 0.92)
6 mo		

Absolute Risk Reduction(ARR)
Absolute Risk difference

- تفاوت خطر
- ARR=0 تفاوتی بین دو گروه از نظر تاثیر درمان وجود ندارد
- ARR مثبت: درمان مفید است
- ARR منفی: درمان مضر است
- ریسک پیامد در گروه مقایسه منهای ریسک پیامد در گروه مداخله

0.15-0.10=0.05(5%) ۵ درصد کاهش در میزان مرگ ناشی از درمان

- If the experimental treatment increased the risk of a good event, we can use this same equation to calculate the absolute benefit increase (ABI).
- Or, if the experimental treatment increases the risk of an adverse event, we can use the equation to calculate the absolute risk increase (ARI)

Relative Risk reduction(RRR)

- کاهش میزان بروز پیامد در گروه مواجهه نسبت به گروه مقایسه
- 1- RR
- ARR/risk of event in control group
- 0.05/0.15= 0.33(33%)
- **بروز عوارض یا مرگ در گروه مداخله ۳۳٪ کمتر از گروه مقایسه است**

- If the experimental treatment increases the risk of a good event, we can use this same equation to calculate the relative benefit increase (RBI).
- Similarly, if the experimental treatment increases the risk of an adverse event we can use the equation to calculate the relative risk increase (RRI).

Number needed to treat(NNT)

- 1/ARR
- 1/0.05=20
- **تعداد بیمارانیکه باید درمان شوند تا از یک مورد مرگ یا عوارض پیشگیری شود**
- **۲۰ بیمار به مدت ....سال باید درمان شوند تا از یک مورد مرگ جلوگیری شود**

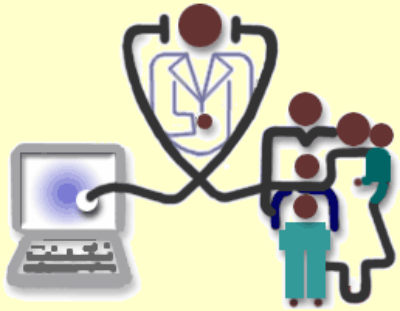
- In a randomised controlled trial (Stroke 1997; 28:1861-6) looking into the long-term outcome for stroke patients treated in stroke units (SU) compared with patients treated in general wards (GW),
- the mortality rate 5 years after the onset of stroke was 59.1% in the patients treated in SU and 70.9% in those treated in the GW. How many patients need to be treated in stroke units to prevent one additional death?
- NNT = _____

- ARR = |CER - EER| = |0.709 - 0.591| = 0.118
  NNT = 1/ARR = 1/0.118 = 9

- Nine patients would need to be treated in stroke units to prevent one additional death.

- When the treatment increases the risk of adverse events, we can calculate the number of patients that we'd need to treat with this therapy to cause one additional bad event and this term is called the number needed to harm (NNH). The NNH is calculated as 1/ARI.

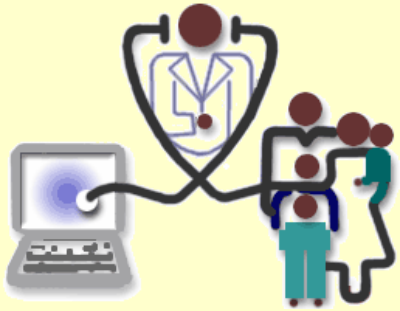
Important example



Event: stroke
Mean follow-up: 5 years

Control rate	Experimental rate	RRR CER-EER/CER	ARR CER-EER	NNT 1/ARR
5.7%	4.3%	(5.7%-4.3%)/5.7%=25%	5.7%-4.3%=0.014	1/1.4%=72

Important example



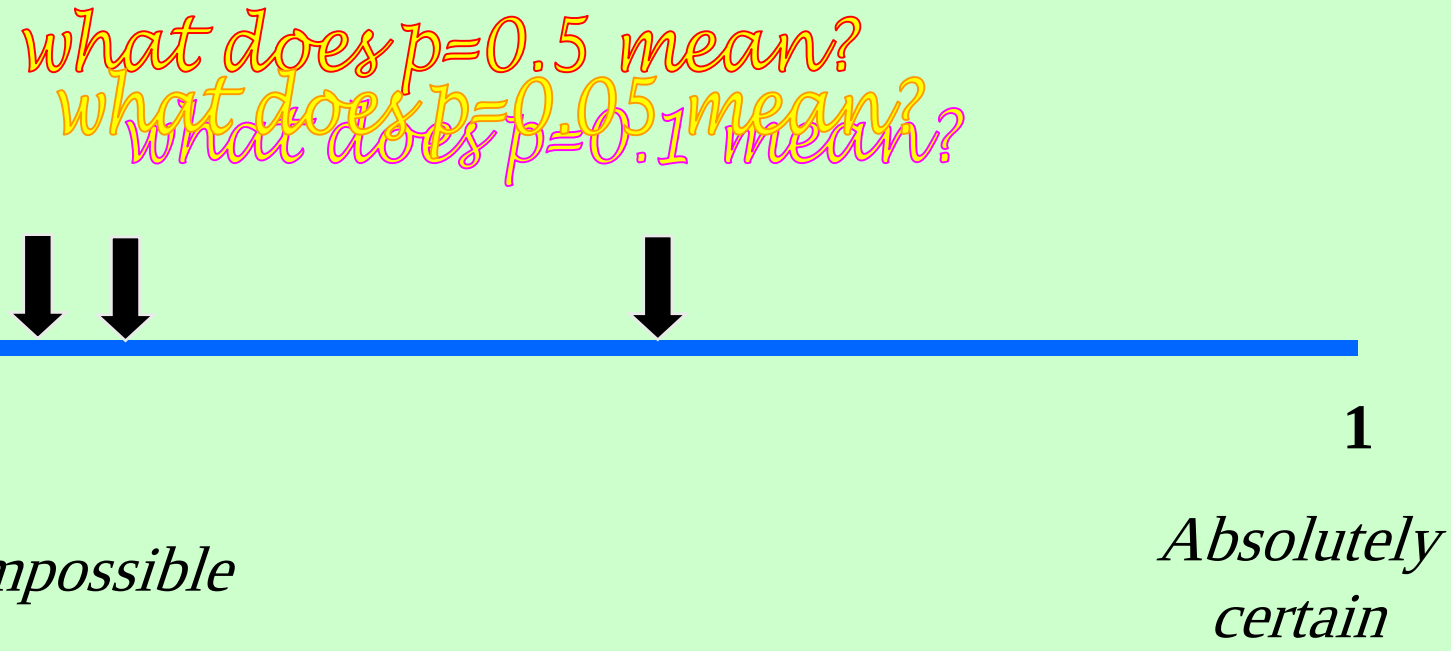
Event: stroke

Mean follow-up: 5 years

Control rate	Experimental rate	RRR CER-EEE/CER	ARR CER-EER	NNT 1/ARR
5.7%	4.3%	(5.7%-4.3%)/5.7%=25%	5.7%-4.3%=0.014	1/1.4%=72
0.000057%	0.000043%	(0.000057%-0.000043%)/0.000057%=25%	0.000057%-0.000043%=0.000014%	1/0.000014%=7142857

P value

چقدر احتمال دارد که نتایج بدست آمده بخاطر شانس باشد؟(شانسی باشد)

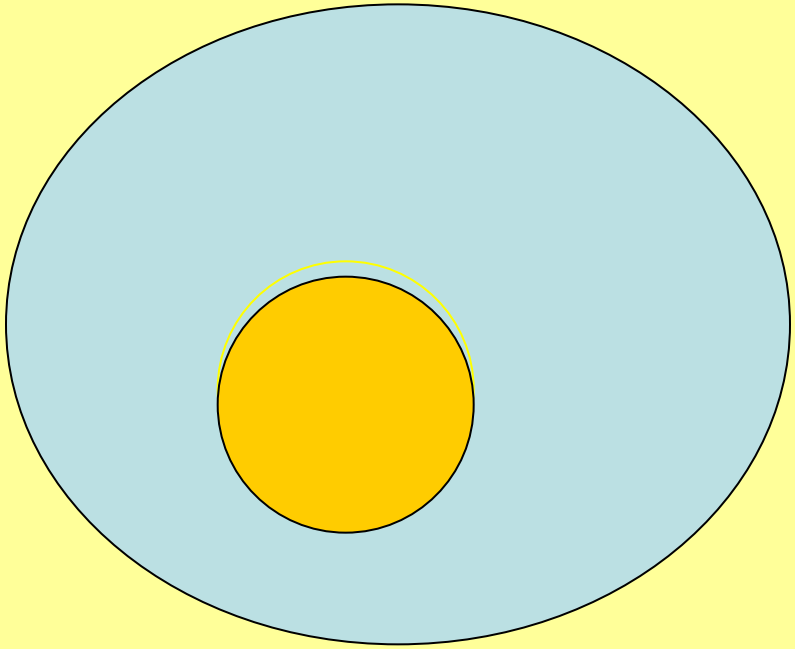


Number in treatment arm	**5**
Responders in treatment arm	**4**
Proportion responding in treatment arm	**0.8**
Number in control arm	**5**
Responders in control arm	**2**
Proportion responding in control arm	**0.4**
p-value	**0.29**

Number in treatment arm	**5**	**10**
Responders in treatment arm	**4**	**8**
Proportion responding in treatment arm	**0.8**	**0.8**
Number in control arm	**5**	**10**
Responders in control arm	**2**	**4**
Proportion responding in control arm	**0.4**	**0.4**
p-value	**0.29**	**0.09**

	Number in treatment arm	5	10	15	20	100
	Responders in treatment arm	4	8	12	16	80
	Proportion responding in treatment arm	0.8	0.8	0.8	0.8	0.8
	Number in control arm	5	10	15	20	100
	Responders in control arm	2	4	6	8	40
	Proportion responding in control arm	0.4	0.4	0.4	0.4	0.4
	p-value	0.29	0.09	0.03	0.01	<0.0001

Confidence interval

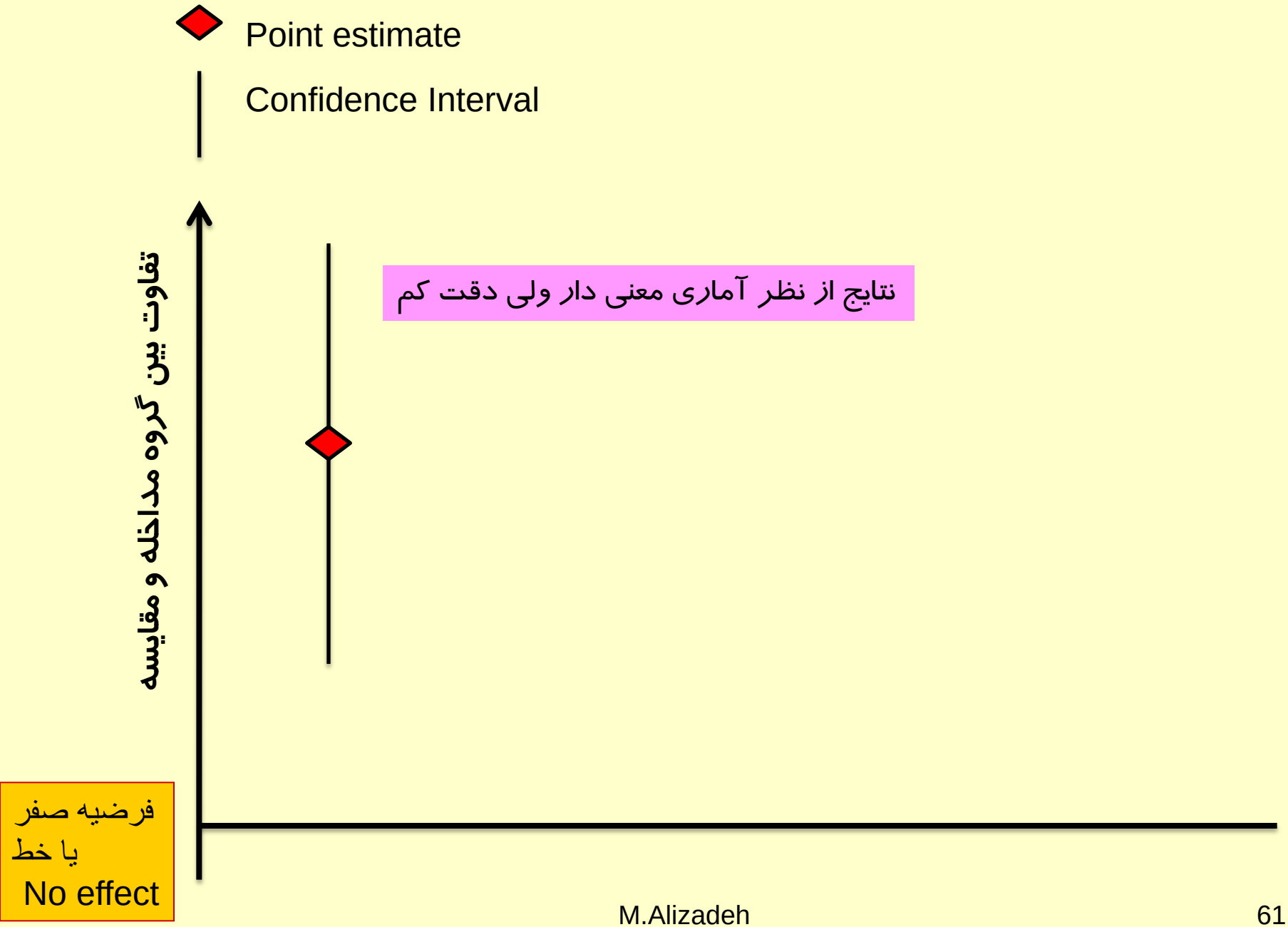


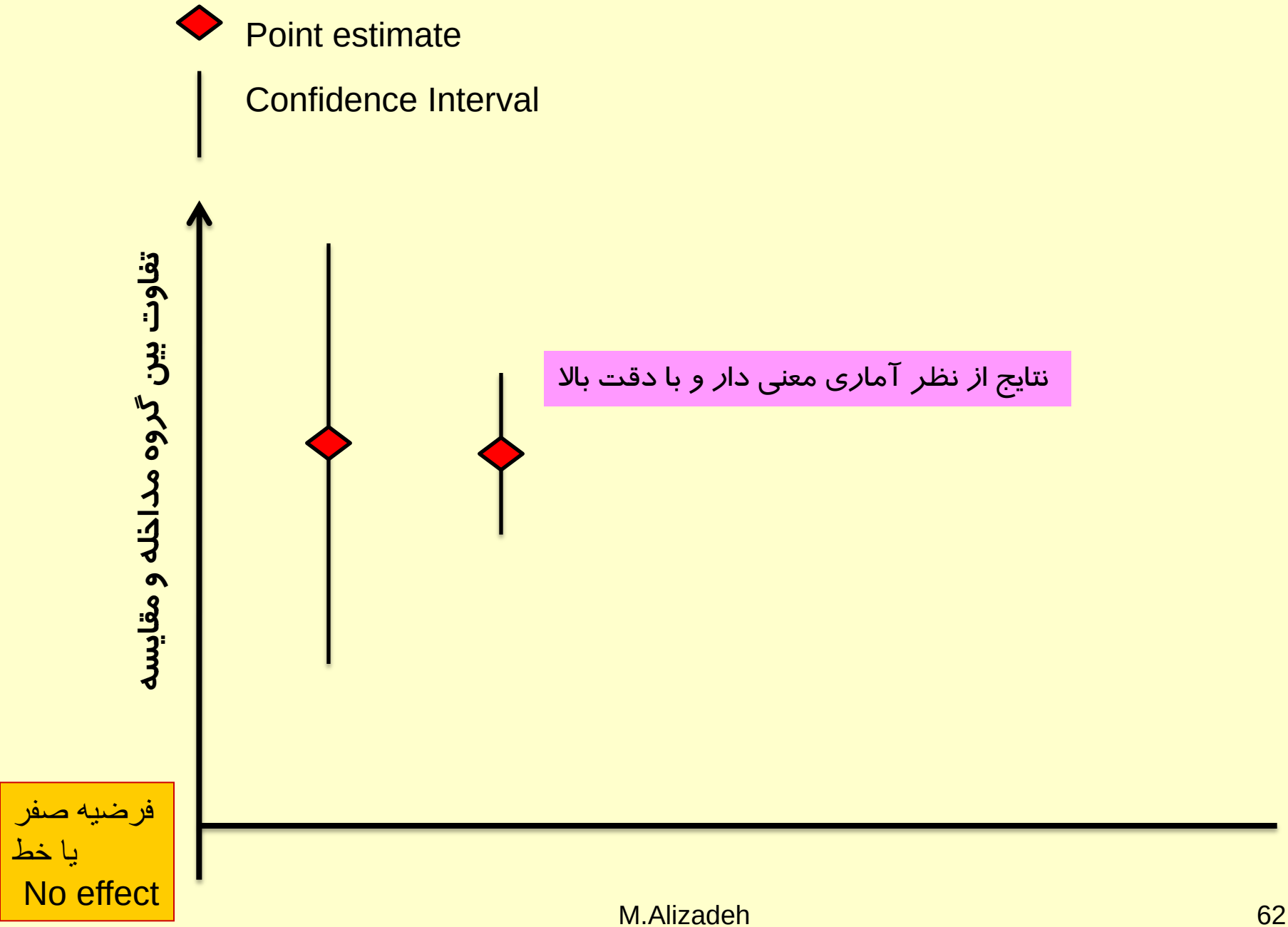
20%

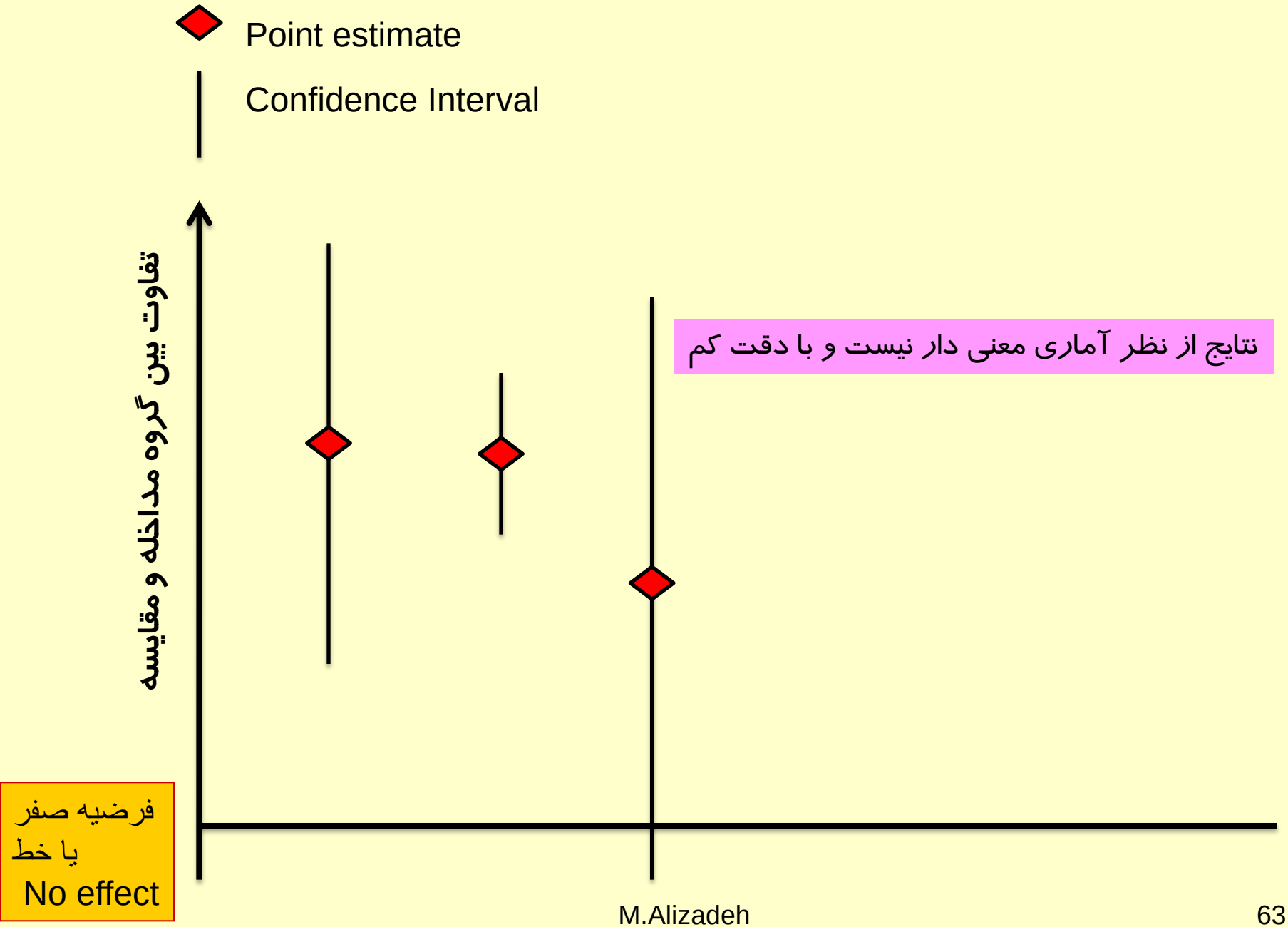
15%-25%

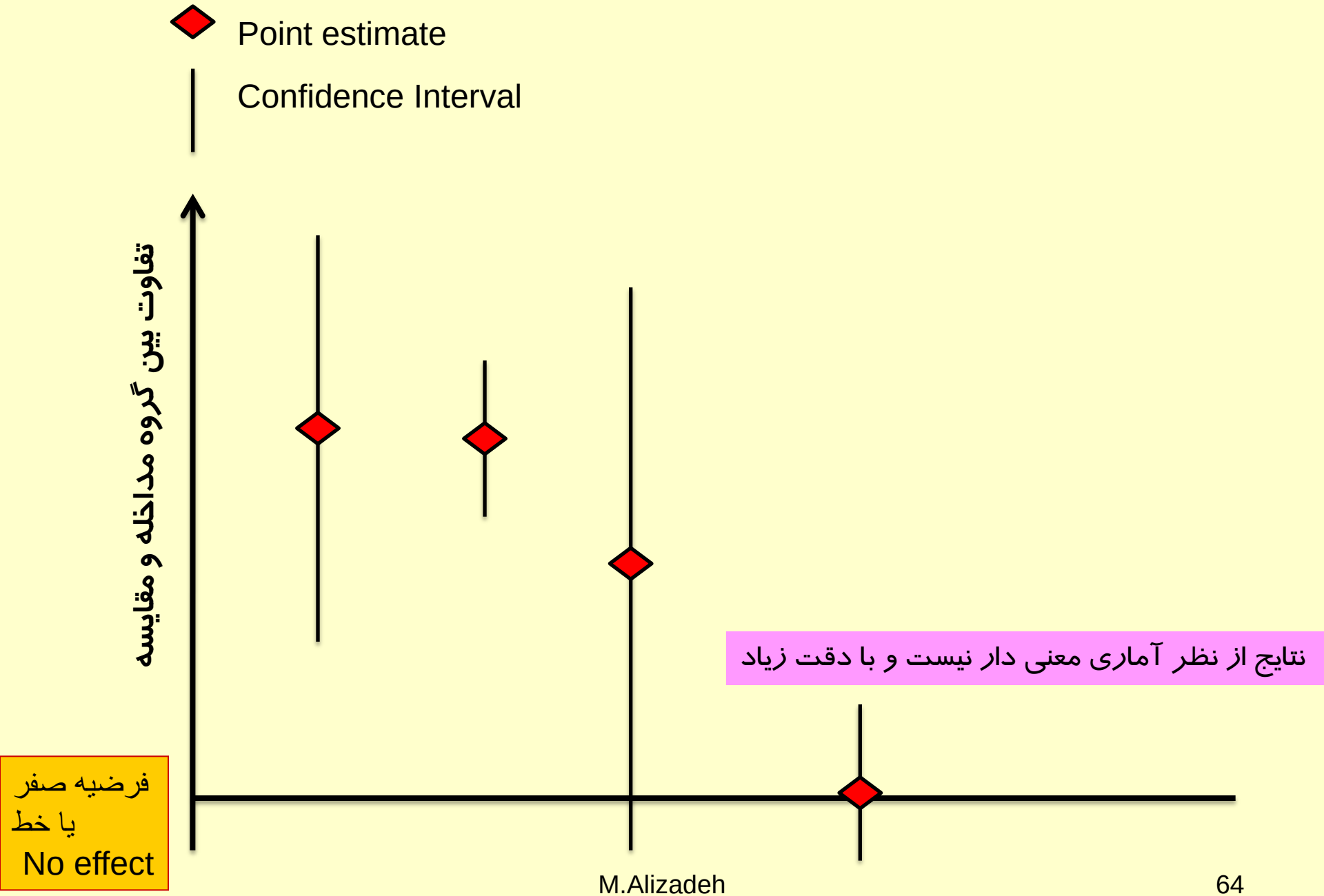
Confidence Interval

- An estimate of the range of values that are likely to include a real value.


Point estimate
Confidence Interval
تفاوت بین گروه مداخله و مقایسه
نتایج از نظر آماری معنی دار ولی دقت کم
فرضیه صفر
یا خط
No effect


Point estimate
Confidence Interval
تفاوت بین گروه مداخله و مقایسه
نتایج از نظر آماری معنی دار و با دقت بالا
فرضیه صفر
یا خط
No effect


Point estimate
Confidence Interval
تفاوت بین گروه مداخله و مقایسه
نتایج از نظر آماری معنی دار نیست و با دقت کم
فرضیه صفر
یا خط
No effect

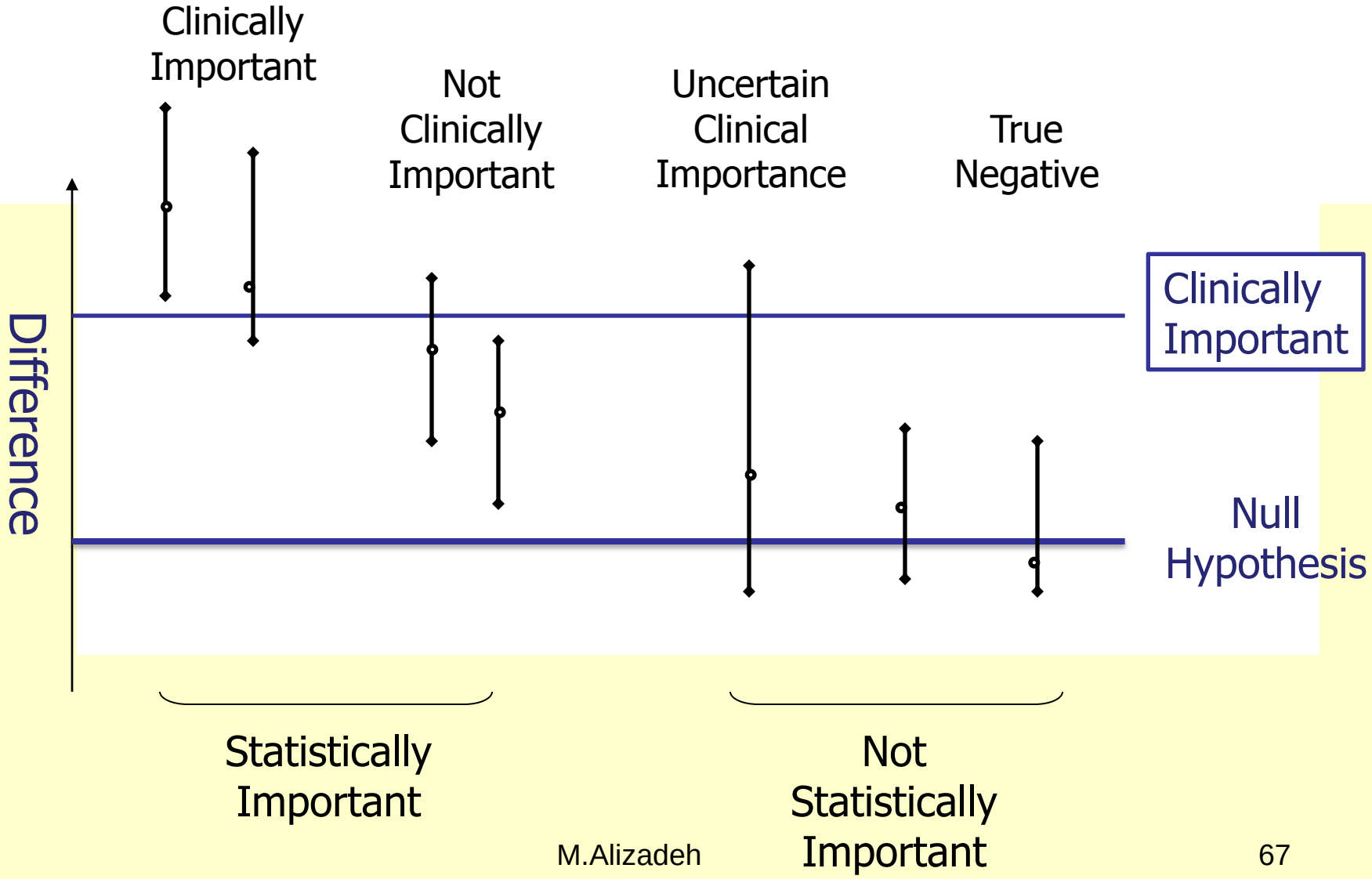

Point estimate
Confidence Interval
تفاوت بین گروه مداخله و مقایسه
نتایج از نظر آماری معنی دار نیست و با دقت زیاد
فرضیه صفر
یا خط
No effect

- a randomized placebo-controlled trial of acellular pertussis vaccine:
- 72/1670 (4.3%) infants developed pertussis among those receiving the vaccine
- 240/1665 (14.4%) did so among the control group.
- absolute risk reduction, is 10.1%.
- The SE of this difference is 0.99%,
- so that the 95% CI is 10.1% ± 1.96 × 0.99%, and therefore runs from 8.2 to 12.0.

- ***In a randomized trial to compare suturing and stapling for large-bowel anastomosis,***
- ***wound infection occurred in 10.9% and 13.5% of cases respectively (P = 0.30).***
- ***The 95% CI for this difference of 2.6% is −2 to +8.***

- ***Sung et al carried out a randomized trial to compare octreotide infusion and emergency sclerotherapy for acute variceal hemorrhage in 100 patients.***
- ***The observed rates of controlled bleeding were 84% in the octreotide group and 90% in the sclerotherapy group,***
- ***P = 0.56.***
- ***In this case, however, the 95% CI for the treatment difference of 6% is −7 to +19.***
- ***It is clear that the study cannot rule out a large difference in effectiveness, so that the authors' conclusion that "octreotide infusion and sclerotherapy are equally effective in controlling variceal haemorrhage" is certainly not valid.***

Distinction between statistical significance and clinical importance (Berry 1986)



Thank You

alizadehm@tbzmed.ac.ir