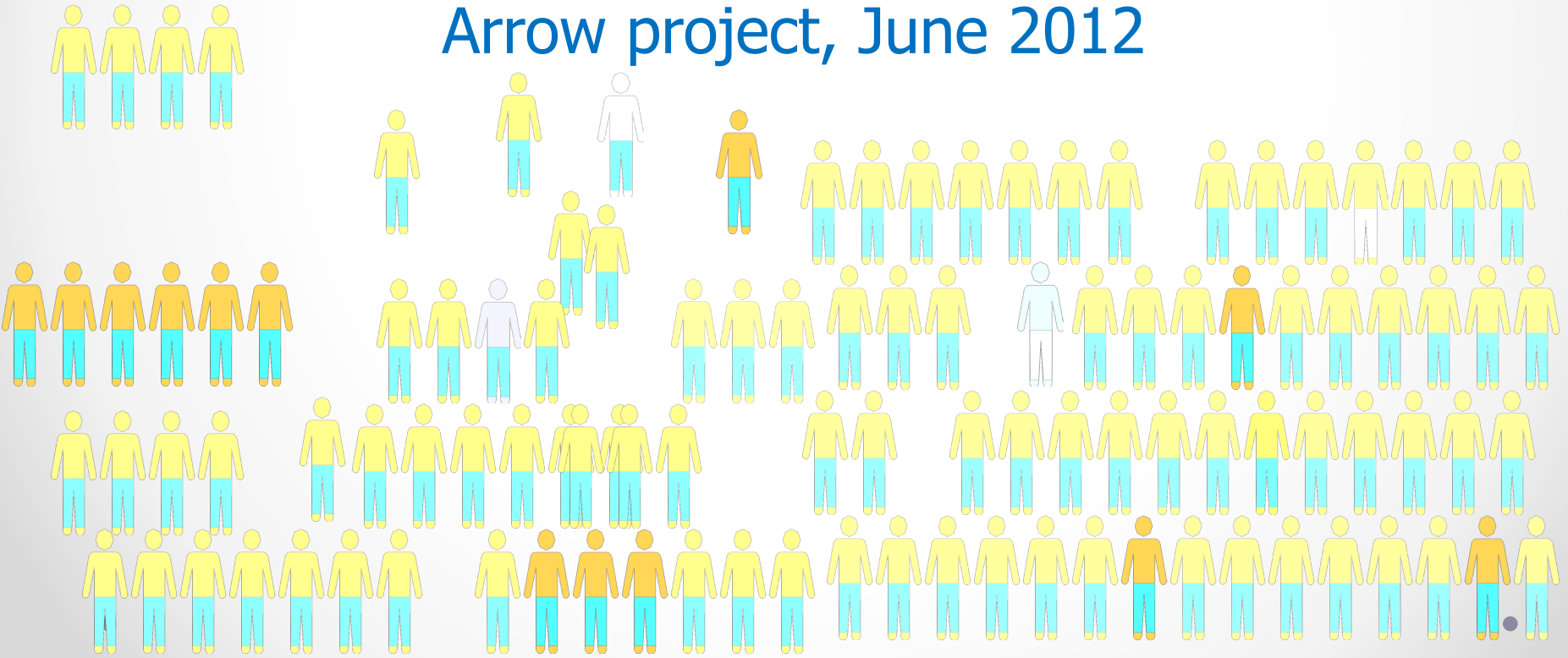
מכון גרטנר לחקר אפידמיולוגיה ומדיניות בריאות
THE GERTNER INSTITUTE FOR EPIDEMIOLOGY AND HEALTH POLICY RESEARCH

Research Methods

Sample Size Calculation

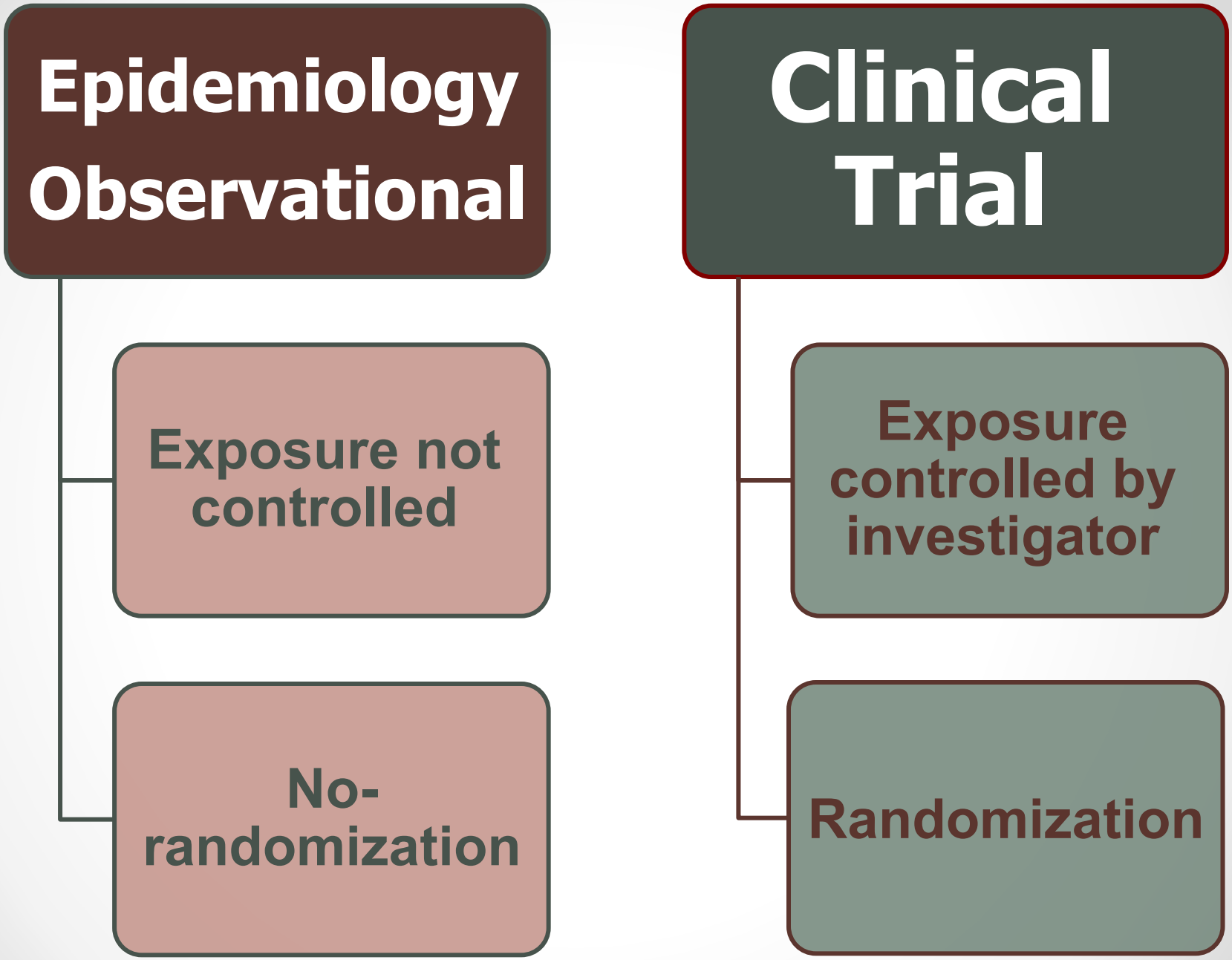
Liat Lerner-Geva M.D., Ph.D

Arrow project, June 2012



STUDY DESIGN

- **Analytic**
  - **Epidemiology Observational**
  - **Clinical Trial**
- **Descriptive**
  - **Prevalence/ Incidence**


Epidemiology Observational
Exposure not controlled
No-randomization
Clinical Trial
Exposure controlled by investigator
Randomization

Analytic Studies

Exposed & Outcome Simultaneously

RR

Ex	Dis +	Dis -
+	a	b
-	c	d

$$RR=\frac{a/a+b}{c/c+d}$$

Cases Vs. Controls

OR

Ex	Dis +	Dis -
+	a	b
-	c	d

$$OR=\frac{a*d}{c*b}$$

Exposed Vs. Unexposed

RR

Ex	Dis +	Dis -
+	a	b
-	c	d

$$RR=\frac{a/a+b}{c/c+d}$$

Exposed Vs. Unexposed

RR

Ex	Dis +	Dis -
+	a	b
-	c	d

$$RR=\frac{a/a+b}{c/c+d}$$

Levels of Evidence

Box 2. Hierarchies of evidence for questions of therapy, prevention, aetiology or harm[26]

Level 1a	Systematic review (with homogeneity) of randomised controlled trials (RCTs)
Level 1b	Individual RCT (with narrow confidence interval)
Level 1c	All-or-none studies
Level 2a	Systematic review (with homogeneity) of cohort studies
Level 2b	Individual cohort study (including low quality RCT; eg <80% follow-up)
Level 2c	'Outcomes' research; ecological studies
Level 3a	Systematic reviews (with homogeneity) of case-control studies
Level 3b	Individual case-control study
Level 4	Case series (and poor quality cohort and case-control studies)
Level 5	Expert opinion without explicit critical appraisal, or based on physiology, bench research or 'first principles'

RR/OR > 1

Positive association- Risk factor

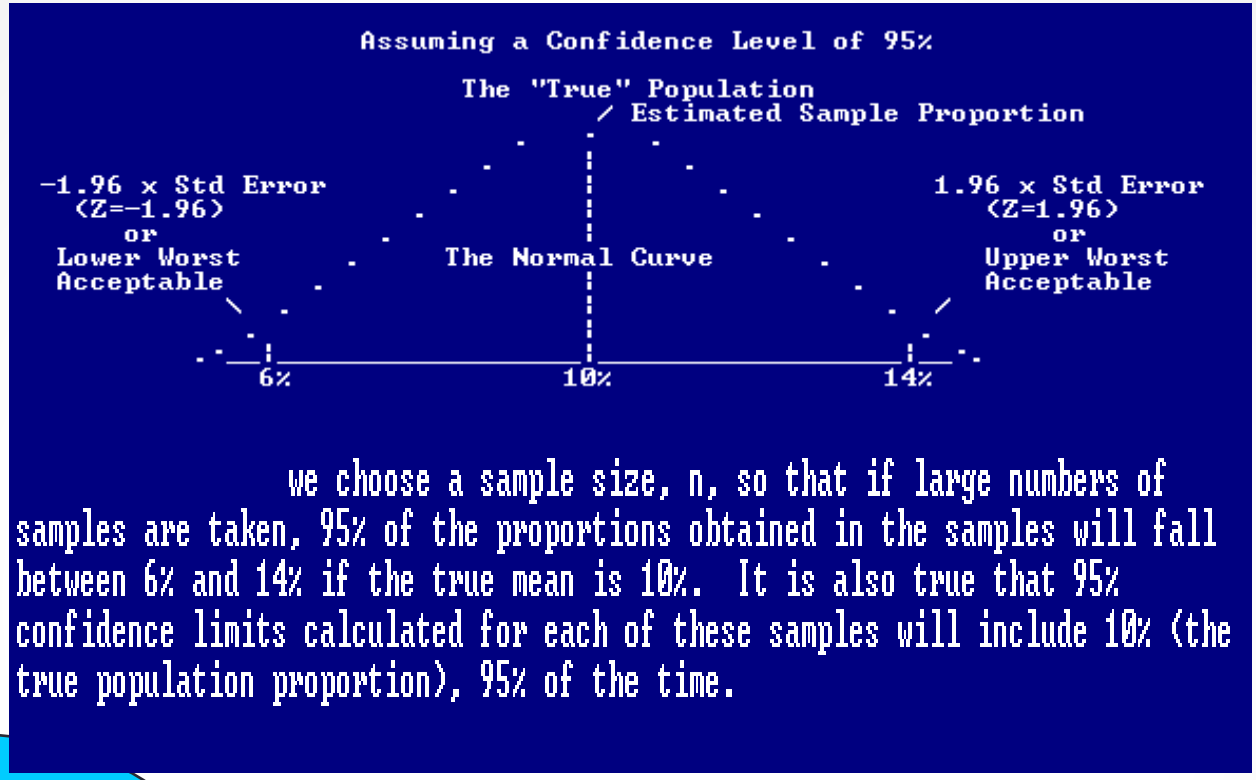
RR/OR = 1

No association

RR/OR < 1

Negative association- Protective factor

Confidential Interval



95% CI for P

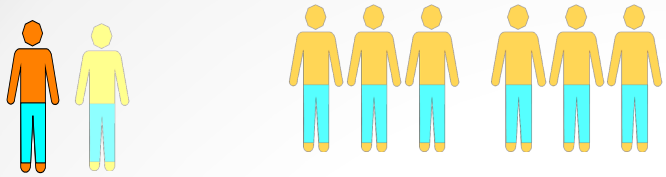
A 95% confidence level means that 95% of the intervals would include the parameter;

p

Sample

95%CI

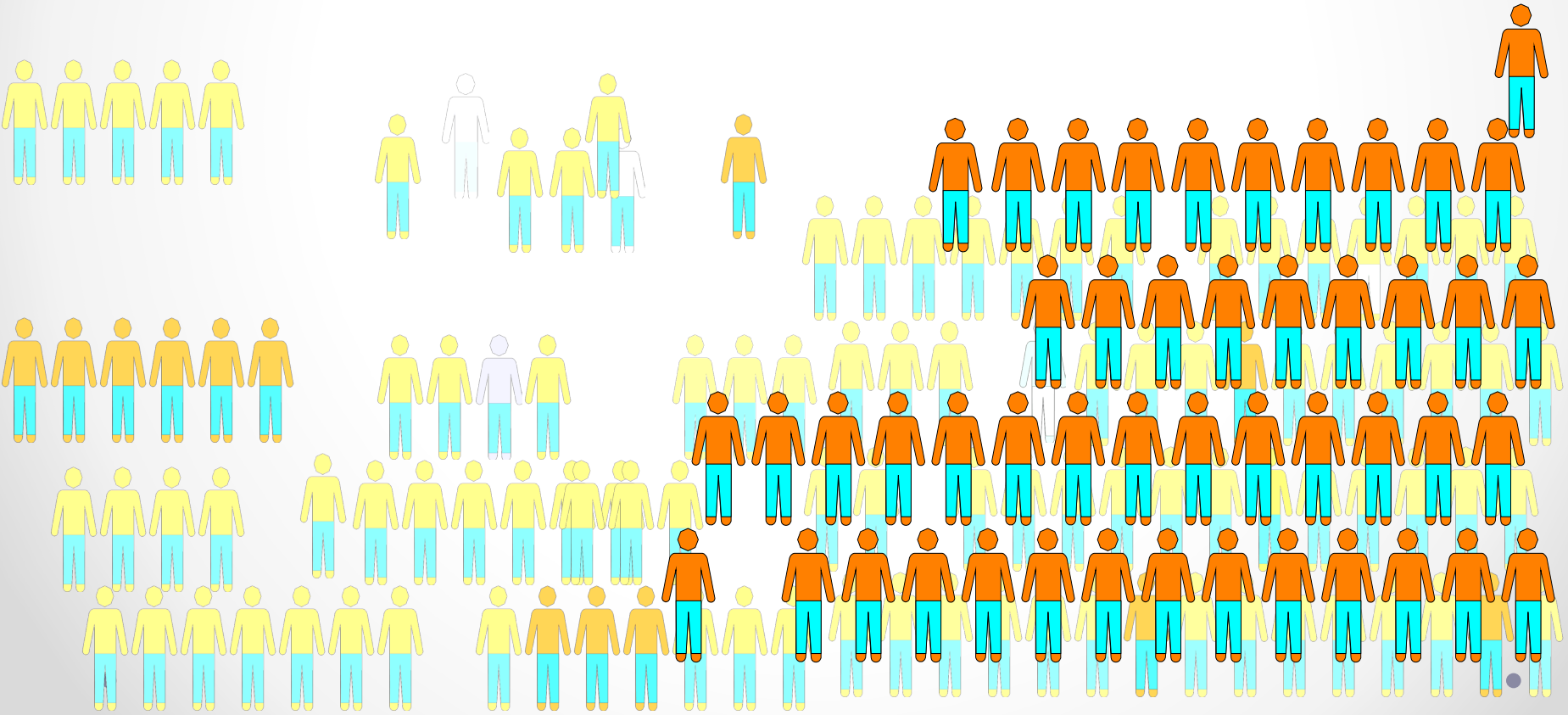
- RR=1.2 95%CI 0.9-1.9
- RR=1.2 95%CI 1.1-18.7
- RR=1.2 95%CI 1.1-2.3



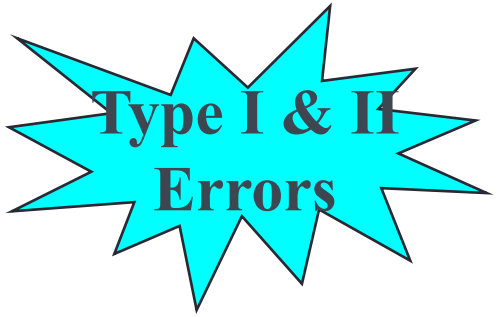
"How Many Subjects Do We Have To Study ?"

The time to answer this question is ***before*** the study is done.

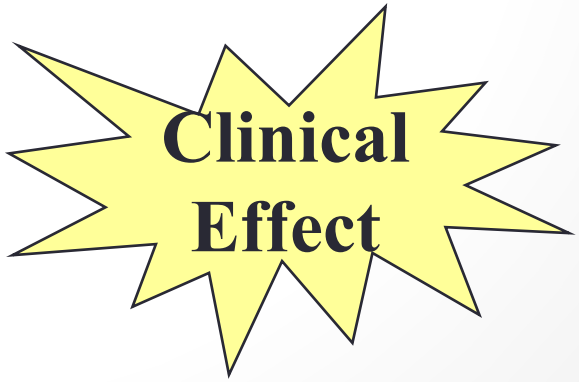
Gordis L. Epidemiology 2nd edition. W.B Saunders Company, Philadelphia, 2000.



Sample Size



- Type I error
- Type II error



- Prevalence
- Expected difference

Type I & II Errors

Type I Error / α

Probability that if the two **SAMPLES** differ this reflects a true difference in the two **POPULATIONS**

confidence level or 1- **α : 95%**

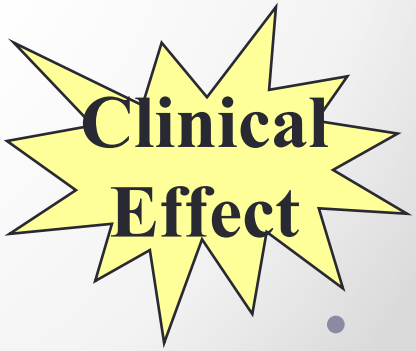
Type II Error/ β Power/1-β

Type I & II Errors

- No effect observed, when in fact there is a real effect
- If the test hypothesis is false, but is not rejected, the **incorrect** decision to **not reject** is called *type-II* error, or beta error.
- POWER=1- **β**

PREVALENCE & EXPECTED DIFFERENCES

- Literature
- Pilot studies
- Expert opinion
- Clinical significance
- Basic science evaluation
- .....



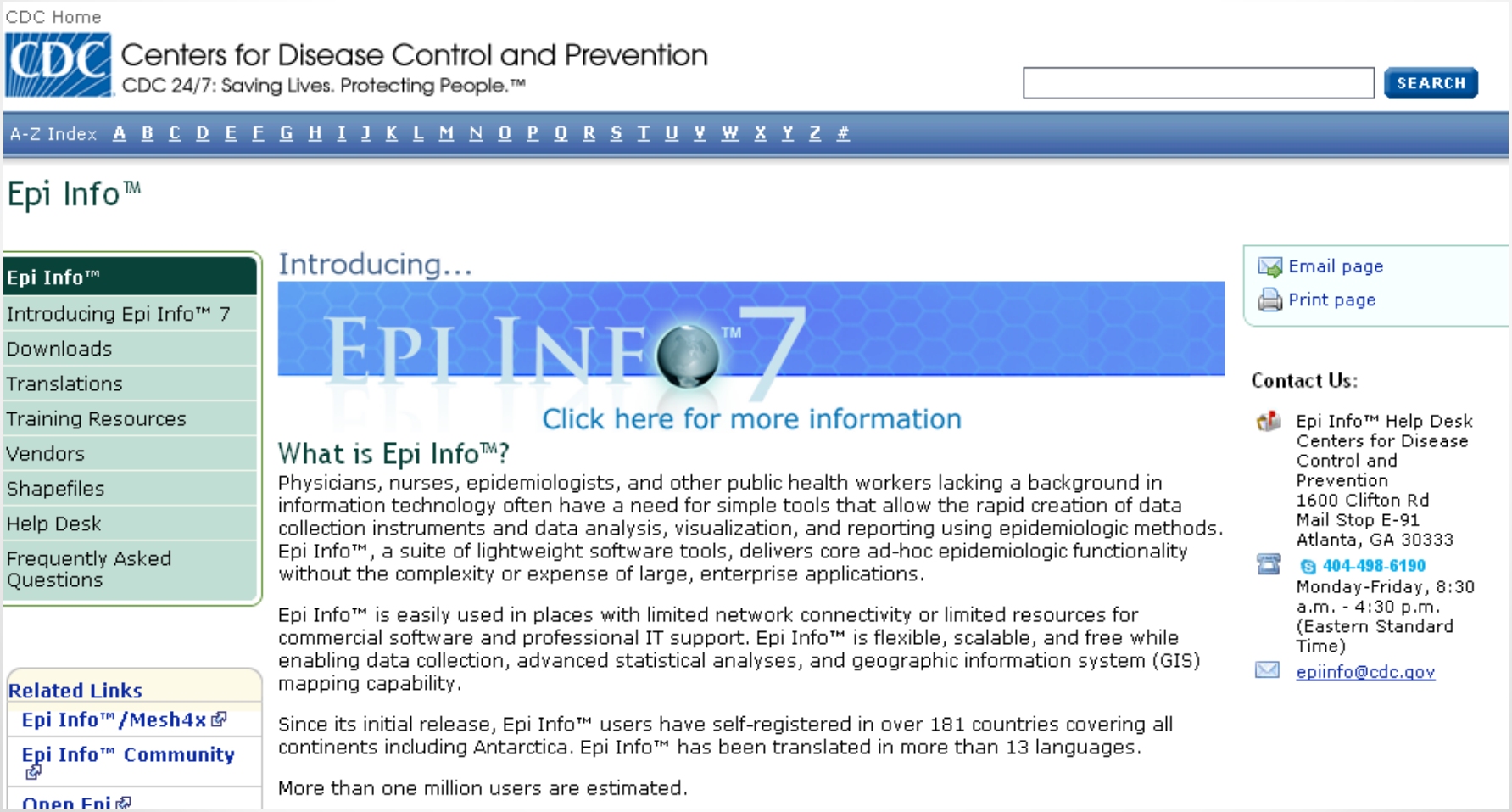
The effect of prevalence

Clinical Effect

RR=2

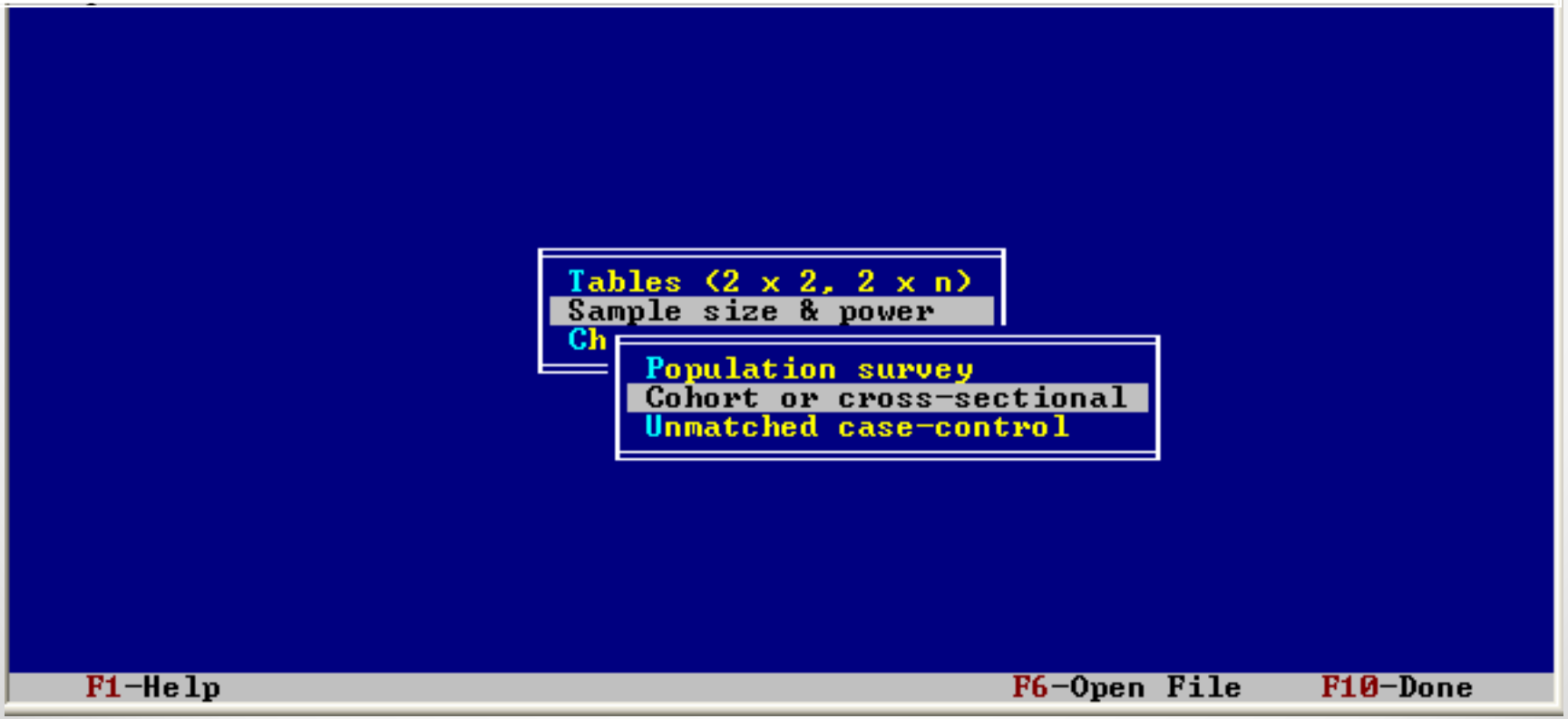
Frequent event 20%	Rare event 2%	
20/100 → 40/100	2/100 → 4/100	Small sample
200/1000 → 400/1000	20/1000 → 40/1000	Large sample

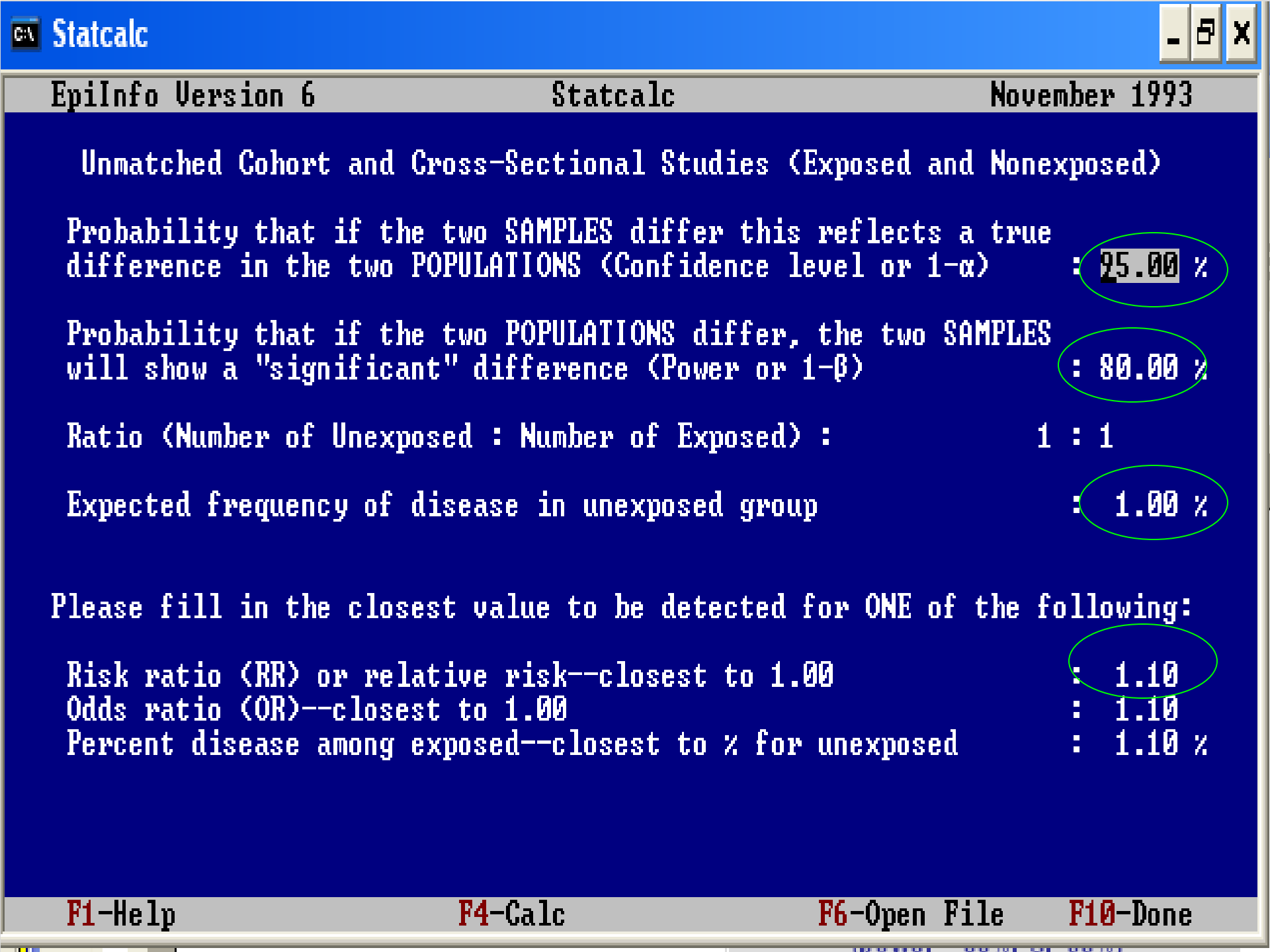
WWW.CDC.EPIINFO



COHORT

(power = 80%, α=0.05)




Statcalc
EpiInfo Version 6 Statcalc November 1993
Unmatched Cohort and Cross-Sectional Studies (Exposed and Nonexposed)
Probability that if the two SAMPLES differ this reflects a true difference in the two POPULATIONS (Confidence level or 1-α) : 95.00 %
Probability that if the two POPULATIONS differ, the two SAMPLES will show a "significant" difference (Power or 1-β) : 80.00 %
Ratio (Number of Unexposed : Number of Exposed) : 1 : 1
Expected frequency of disease in unexposed group : 1.00 %
Please fill in the closest value to be detected for ONE of the following:
Risk ratio (RR) or relative risk--closest to 1.00 : 1.10
Odds ratio (OR)--closest to 1.00 : 1.10
Percent disease among exposed--closest to % for unexposed : 1.10 %
F1-Help F4-Calc F6-Open File F10-Done

Statcalc

EpiInfo Version 6 — Statcalc — November 1993

Unmatched Cohort and Cross-Sectional Studies (Exposed and Nonexposed)
Sample Sizes for 1.00 % Disease in Unexposed Group

Conf.	Power	Unex:Exp	Disease in Exposed	Risk Ratio	Odds Ratio	Sample Size Unexp.	Exposed	Total
95.00 %	80.00 %	1:1	1.10 %	1.10	1.10	165,089	165,089	330,178
90.00 %	"	"				130,462	130,462	260,924
95.00 %	"	"				165,089	165,089	330,178
99.00 %	"	"				244,678	244,678	489,356
99.90 %	"	"				356,798	356,798	713,596
95.00 %	80.00 %	"				165,089	165,089	330,178
"	90.00 %	"				220,332	220,332	440,664
"	95.00 %	"				272,017	272,017	544,034
"	99.00 %	"				383,763	383,763	767,526
"	80.00 %	4:1				408,048	102,012	510,060
"	"	3:1				327,069	109,023	436,092
"	"	2:1				246,086	123,043	369,129
"	"	1:2				124,576	249,153	373,729
"	"	1:3				111,068	333,205	444,273
"	"	1:4				104,313	417,252	521,565

Change values for inputs as desired, then press F4 to recalculate.

F1-Help F5-Print F6-Open File F10-Done

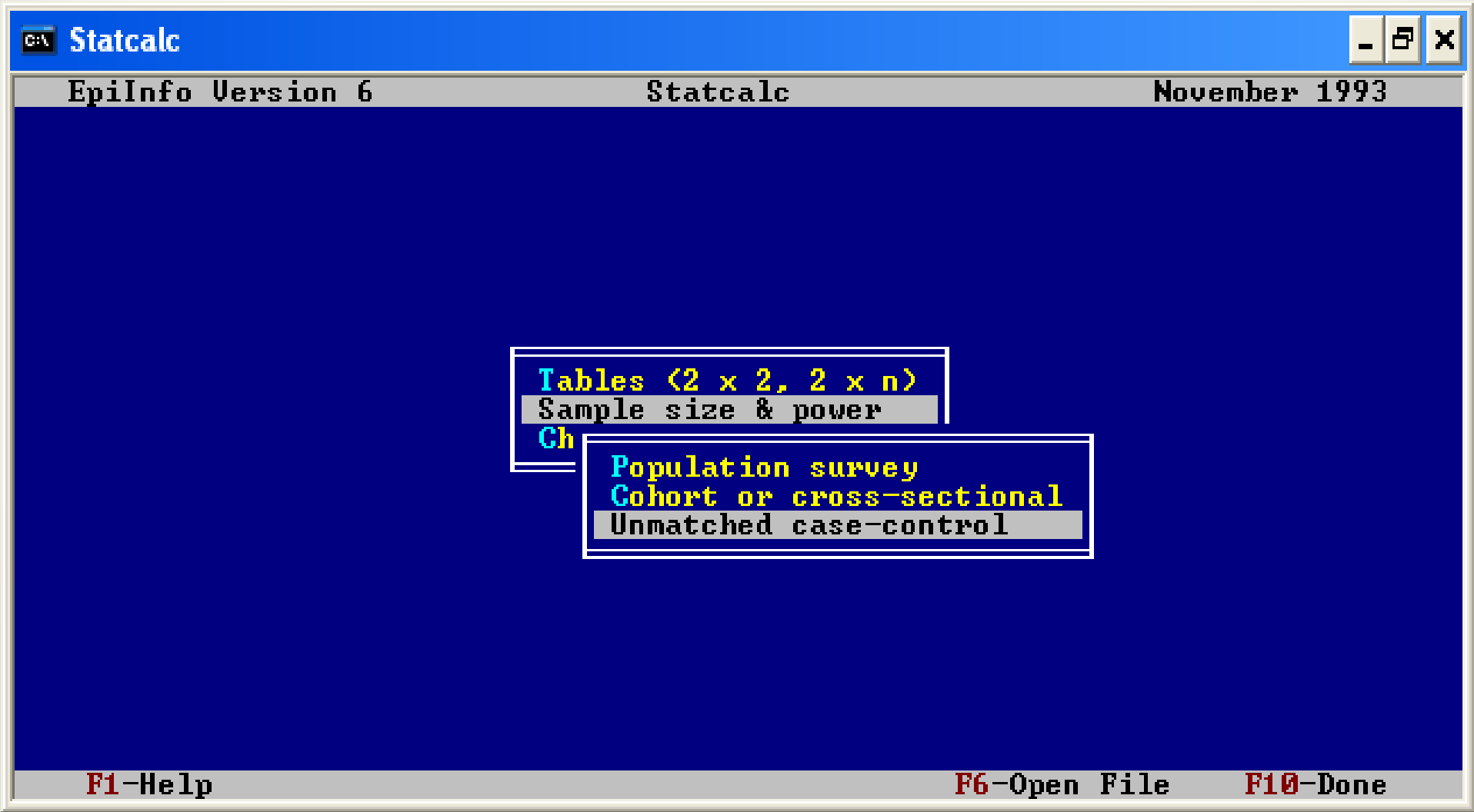
Cohort Study-Examples

(power = 80%, α=5%)

Unexposed (n)	Exposed (n)	Outcome among Unexposed	RR-minimal	Unexposed/ Exposed Ratio
165,089	165,089	1.0	1.1	1.0
16,389	16,389	0.5	1.5	1.0
8,145	8,145	1.0	1.5	1.0
5,065	5,065	0.5	2.0	1.0

Case-Control Study

(power = 80%, CI=95%)



Statcalc

EpiInfo Version 6 | Statcalc | November 1993

Unmatched Case-Control Study (Comparison of ILL and NOT ILL)

Probability that if the two SAMPLES differ this reflects a true difference in the two POPULATIONS (Confidence level or 1-α) : 95.00 %

Probability that if the two POPULATIONS differ, the two SAMPLES will show a "significant" difference (Power or 1-β) : 80.00 %

#NOT ILL/#ILL (1 means equal sample sizes) 1 : 1

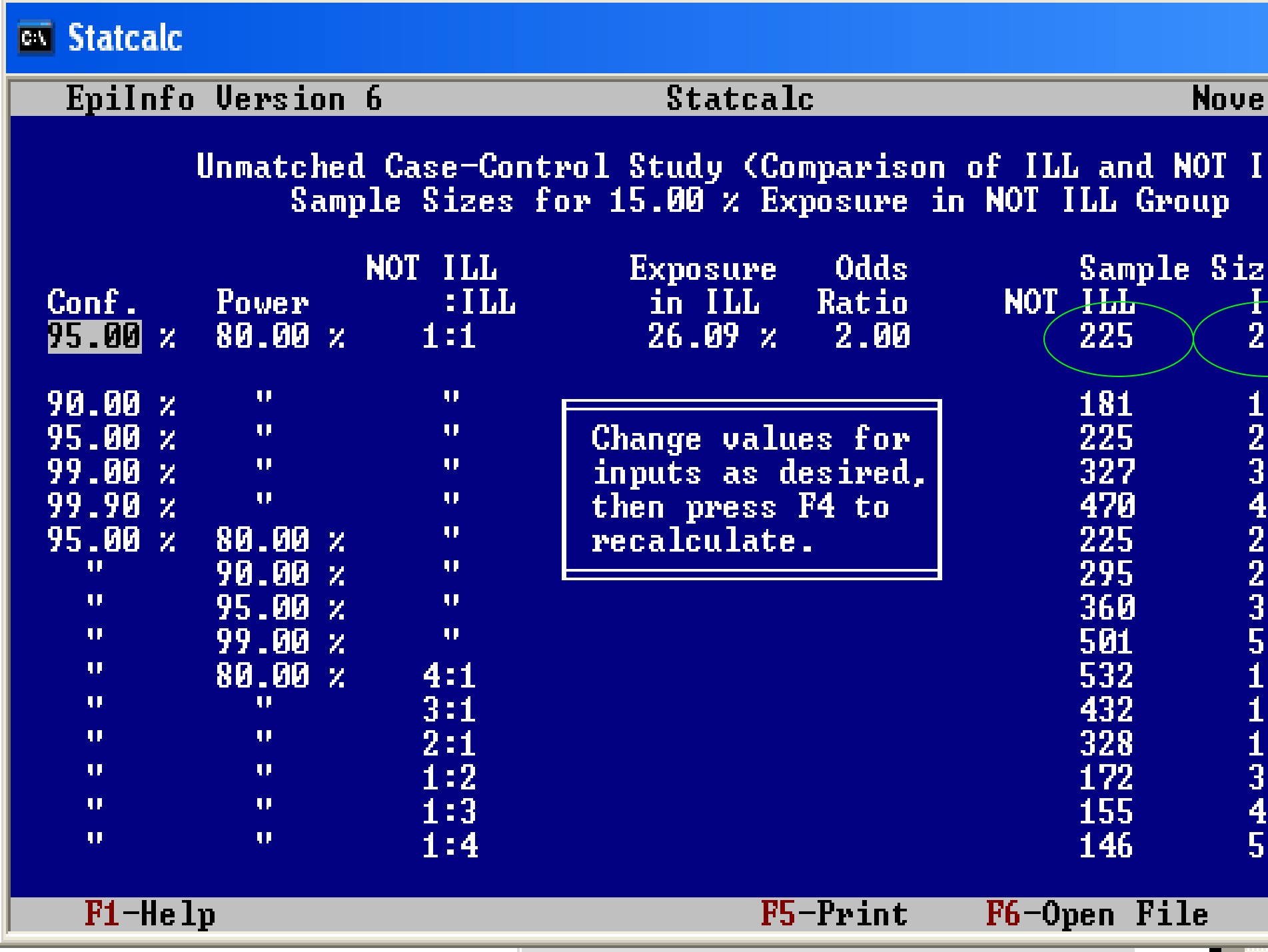
Expected frequency of exposure in NOT ILL group : 15.00 %

Please fill in the closest value to be detected for ONE of the following:

Odds ratio (OR)--closest to 1.00 : 2.00

Percent exposure among ILL group--closest to % for NOT ILL : 26.09 %

F1-Help F4-Calc F6-Open File F10-Done


Statcalc
EpiInfo Version 6
Statcalc
Nove
Unmatched Case-Control Study (Comparison of ILL and NOT I
Sample Sizes for 15.00 % Exposure in NOT ILL Group
Conf.
Power
NOT ILL :ILL
Exposure in ILL
Odds Ratio
Sample Siz
NOT ILL
I
95.00 % 80.00 % 1:1 26.09 % 2.00 225 2
90.00 % " " 181 1
95.00 % " " 225 2
99.00 % " " 327 3
99.90 % " " 470 4
95.00 % 80.00 % " 225 2
" 90.00 % " 295 2
" 95.00 % " 360 3
" 99.00 % " 501 5
" 80.00 % 4:1 532 1
" " 3:1 432 1
" " 2:1 328 1
" " 1:2 172 3
" " 1:3 155 4
" " 1:4 146 5
Change values for inputs as desired, then press F4 to recalculate.
F1-Help
F5-Print
F6-Open File

Case-Control Study

(power = 80%, CI=95%)

Controls (n)	Cases (n)	% Exposed in Controls	Minimal OR	Cases/Controls Ratio
225	225	15.0	2.0	1.0
691	691	15.0	1.5	1.0
447	447	30.0	1.5	1.0
676	338	30.0	1.5	2.0

In summary….

Sample Size	
↓	↑ Prevalence
↓	↑ Expected RR / OR