Exercise 4

"Do the sample size assumptions for a trial addressing the following question:

"Among couples with unexplained infertility does a program of up to three IVF cycles compared with up to three FSH/IUI cycles increase the live birth rate?"

Can you?

What information do you need?

You will need

a. Hypothesized treatment difference (magnitude of effect)
b. Baseline live birth rate in the control group
c. Allocation ratio
d. Type I error probability (alpha value)
e. Type II error probability (beta value)
f. One or two-sided hypothesis
g. Type of statistical test used

Type I error probability (or alpha value):

the probability of rejecting a null hypothesis when it is actually true or the probability to perform a "false positive" error

Type II error probability (or beta value):

the probability of failing to reject a null hypothesis when in fact we should have rejected the null hypothesis or the probability of performing a "false negative" error

Power:

the probability of righteously rejecting the null hypothesis *(to be able to detect a difference that is actually present in the population)*

Scenario 1

Rate difference: **20%**

Live birth rate in the control group: **21%**

Alpha value (type I error probability): **5%**

Beta value (type II error probability): **20%**

Allocation ratio: **1**

Two-sided Fisher's Exact test

Scenario 1

Calculate the sample size needed

and

write it down !

Scenario 2

Rate difference: **10%**

Live birth rate in the control group: **21%**

Alpha value (type I error probability): **5%**

Beta value (type II error probability): **20%**

Allocation ratio: **1**

Two-sided Fisher's Exact test

Scenario 2

Calculate the sample size needed

and

write it down !

Scenario 3

Rate difference: **10%**

Live birth rate in the control group: **12%**

Alpha value (type I error probability): **5%**

Beta value (type II error probability): **20%**

Allocation ratio: **1**

Two-sided Fisher's Exact test

Scenario 3

Calculate the sample size needed

and

write it down !

Scenario 4

Rate difference: **10%**

Live birth rate in the control group: **21%**

Alpha value (type I error probability): **5%**

Beta value (type II error probability): **10%**

Allocation ratio: **1**

Two-sided Fisher's Exact test

Scenario 4

Calculate the sample size needed

and

write it down !

Discussion

	Scenario 1	Scenario 2	Scenario 3	Scenario 4
RD	20%	10%	10%	10%
Baseline Rate	21%	21%	12%	21%
Alpha	*5%*	*5%*	*5%*	*5%*
Beta	*20%*	*20%*	*20%*	*10%*
Power (1-beta)	80%	80%	80%	90%
Estimated Sample size in each group	82	300	217	401

What do you observe?

Comments on power analysis/sample size determination

1. Larger differences/ greater effect sizes need smaller sample size to be detected
2. Not only the actual rate difference, but also the expected baseline rate, determines sample size
3. Increased statistical power (1-beta) demands larger number of participants

Exercise 5

"Outline your analysis plan for the IVF-IUI trial"

"Among couples with unexplained infertility does a program of up to three IVF cycles compared with up to three FSH / IUI cycles increase the live birth rate?"

What are the available options?

What is the primary outcome? **Live Birth**

What is the primary outcome **measure**?

1) Live birth rate per strategy (up to 3 cycles)
2) Cumulative live birth rate?
3) Live birth rate per cycle?

What are the available options?

1) Live birth rate per strategy (up to 3 cycles)

- It assumes that all drop-outs have a 0% of live birth after dropping-out
- It represents the most conservative, yet the most realistic option
- It resembles real-life practice

What is the chance of having a live birth if you start the 3-cycle program

What are the available options?

2) Cumulative live birth rate?

- It assumes that in all drop-outs, the reason for dropping-out was irrelevant to the outcome/treatment
- It resembles, the proportion of patients that would have a live birth if they stayed on treatment for the maximum time (3 cycles)
- It tends to overestimate the cumulative probability of live birth because many times dropping out is quite relevant to ART treatment and outcome

What is the chance of having a live birth if you complete the 3-cycle program

What are the available options?

3) Live birth rate per cycle?

- It takes into account the drop-outs
- It distributes the probability of live birth evenly to each cycle performed
- It does not reflect the actual probability of live birth for each sequential ART cycle (1>2>3…)
- It violates the assumption of independency of observations→ spurious statistical findings

What are the available options?

	IUI (n=100)	IVF (n=100)
Cycle 1	18 (DR:4)	30 (DR:6)
Cycle 2	13 (DR:12)	16 (DR:16)
Cycle 3	9	12
Total live births	40	58
Live birth rate per strategy	40%	58%
Pregnancy rate per cycle	17.3%	29.6%
Cumulative live birth rate	53%	78%

Exercise 6

"Outline your analysis plan for the smoking meta-analysis"

After systematically reviewing the literature you end-up with 14 studies examining the association between smoking and live birth after IVF.

Data extraction from these studies results in the following table

Study	Smokers		Non-smokers	
	Live births	Total IVF cycles	Live births	Total IVF cycles
Study 01	17	77	22	79
Study 02	45	221	67	247
Study 03	111	454	99	401
Study 04	54	202	114	371
Study 05	67	297	111	407
Study 06	22	119	67	289
Study 07	33	188	49	183
Study 08	108	479	211	898
Study 09	27	101	31	117
Study 10	44	193	33	131
Study 11	12	33	10	24
Study 12	59	277	99	373
Study 13	74	499	101	511
Study 14	121	554	201	855

Outline the analysis plan:

1) Which effect measure will you use and why

(Odds ratio, Risk difference, Relative risks)?

PRESENTING RISK AND NNT

In Vitro Fertilization with Preimplantation Genetic Screening. Mastenbroek et al, NEJM July 2007; 357 (1) 9-17

- PGS resulted in lower live birth rate*
  - 24% [49 of 206] vs. 35% [71 of 202]
  - absolute risk diff 11%
  - NNH 10 (1/absolute difference)
  - Rate ratio (Relative Risk) 0.68
  - 95% CI, 0.5 to 0.92
  - p=0.01

PRESENTING MEASURE OF PROBABILITY: RISK VS ODDS

- Risk of drawing a spade from 52 cards

= 13/52 = 1/4 =25%

- Odds of a spade from 52 cards, 1:3

= 13/39 = 1/3 = 33%

RELATIVE RISK AND ODDS RATIOS

	Birth	No birth	Total
PGS	a 49	b 157	206
No PGS	c 71	d 131	202
	120	288	408

RR= a/a+b / c/c+d

= 49/206 / 71/202

RR = 0.68

OR = a x d / c x b

= 49 x 131 / 71 x 157

OR = 0.58

RELATIVE RISK AND ODDS RATIOS

	Birth	No birth	Total
PGS	a 88	b 118	206
No PGS	c 128	d 74	202
	216	192	408

RR= a/a+b / c/c+d

= 88/206 / 128/202

RR = 0.68

OR = a x d / c x b

= 88 x 74 / 128 x 118

OR = 0.43

Outline the analysis plan:

1) Which effect measure will you use and why (Odds ratio, Risk difference, Relative risks)?

- Odds ratios are sometimes difficult for the clinician to understand and for that reason their use should be discouraged
- Relative risks can sometimes be misleading
- Rate difference usually can reflect the effect size and help in the calculation of NNT

Outline the analysis plan:

2) Fixed or Random Effects model and why?

Fixed Effects:

In a fixed effect analysis we assume that all the included studies share a common effect size

Random Effects:

Rather than assume that there is one true effect, we allow that there is a distribution of true effect sizes. The combined effect therefore cannot represent the one common effect, but instead represents the mean of the population of true effects

Outline the analysis plan:

2) Fixed or Random Effects model and why?

- Depends on the presence of clinical and/or statistical heterogeneity
- It should be included in the protocol
- If there are doubts, a sensitivity analysis can be performed

Perform the Data input

and

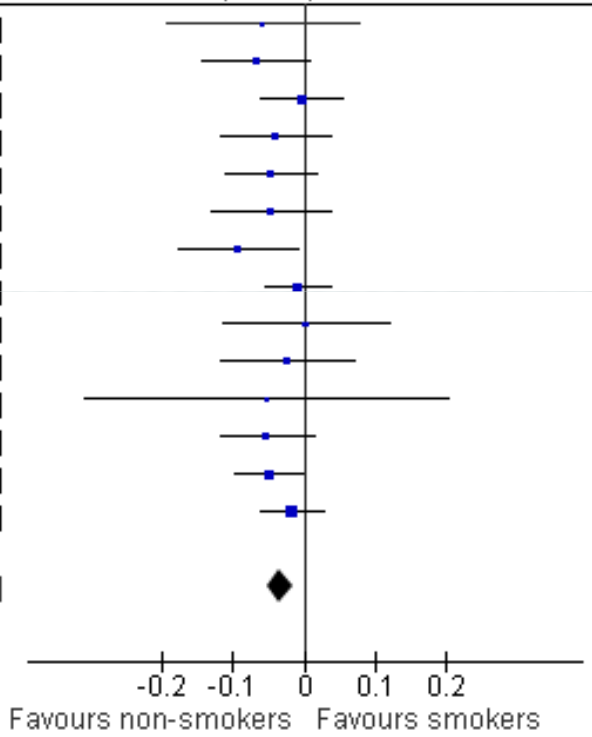
complete the analysis !

The magic of the Forest plot !

Study or Subgroup	Smokers Events	Total	Non-Smokers Events	Total	Weight	Risk Difference IV, Fixed, 95% CI
Study 01	17	77	22	79	1.8%	-0.06 [-0.19, 0.08]
Study 02	45	221	67	247	5.5%	-0.07 [-0.14, 0.01]
Study 03	111	454	99	401	9.7%	-0.00 [-0.06, 0.06]
Study 04	54	202	114	371	5.5%	-0.04 [-0.12, 0.04]
Study 05	67	297	111	407	7.9%	-0.05 [-0.11, 0.02]
Study 06	22	119	67	289	4.5%	-0.05 [-0.13, 0.04]
Study 07	33	188	49	183	4.6%	-0.09 [-0.18, -0.01]
Study 08	108	479	211	898	15.0%	-0.01 [-0.06, 0.04]
Study 09	27	101	31	117	2.4%	0.00 [-0.12, 0.12]
Study 10	44	193	33	131	3.6%	-0.02 [-0.12, 0.07]
Study 11	12	33	10	24	0.5%	-0.05 [-0.31, 0.20]
Study 12	59	277	99	373	7.5%	-0.05 [-0.12, 0.01]
Study 13	74	499	101	511	15.1%	-0.05 [-0.10, -0.00]
Study 14	121	554	201	855	16.4%	-0.02 [-0.06, 0.03]
Total (95% CI)		**3694**		**4886**	**100.0%**	**-0.03 [-0.05, -0.02]**
Total events	794		1215			

Heterogeneity: Chi² = 6.91, df = 13 (P = 0.91); I² = 0%

Test for overall effect: Z = 3.68 (P = 0.0002)



What is your conclusion?

When you constructed the protocol of this analysis you aimed at exploring whether smoking at the time of the IVF cycle is more detrimental than just being a past smoker.

As you examine the study data you realize that some studies include in the "Smokers" group only Current Smokers while other studies have included Current and Past Smokers.

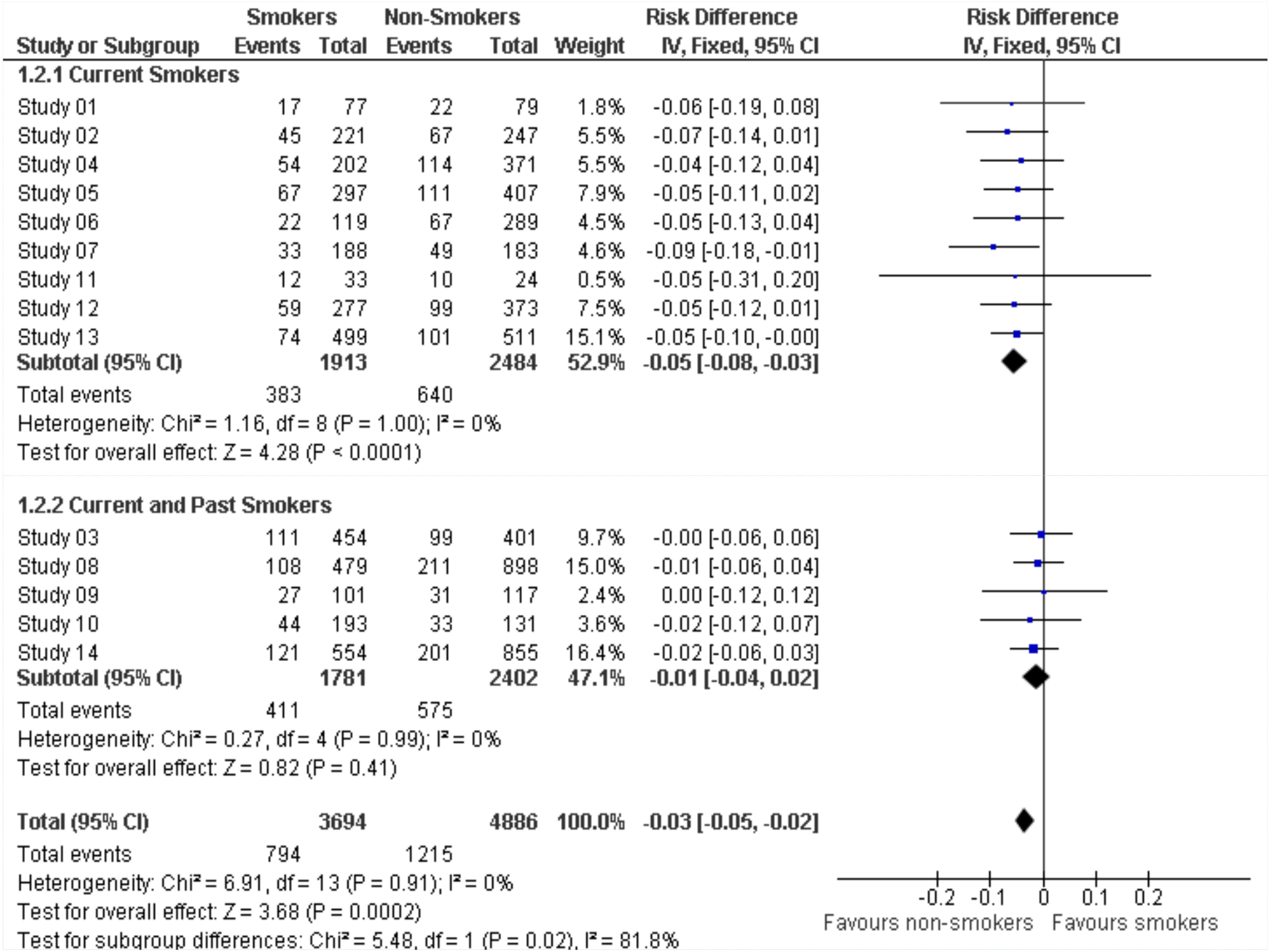
Re-analyze the study data appropriately in order to make inferences regarding this question

Study	Smokers		Non-smokers	
	Live births	Total IVF cycles	Live births	Total IVF cycles
Study 01	17	77	22	79
Study 02	45	221	67	247
Study 03	111	454	99	401
Study 04	54	202	114	371
Study 05	67	297	111	407
Study 06	22	119	67	289
Study 07	33	188	49	183
Study 08	108	479	211	898
Study 09	27	101	31	117
Study 10	44	193	33	131
Study 11	12	33	10	24
Study 12	59	277	99	373
Study 13	74	499	101	511
Study 14	121	554	201	855

Rearrange the Data

and

do the analysis !


Smokers
Non-Smokers
Risk Difference
Risk Difference
Study or Subgroup
Events
Total
Events
Total
Weight
IV, Fixed, 95% CI
IV, Fixed, 95% CI
1.2.1 Current Smokers
Study 01 17 77 22 79 1.8% -0.06 [-0.19, 0.08]
Study 02 45 221 67 247 5.5% -0.07 [-0.14, 0.01]
Study 04 54 202 114 371 5.5% -0.04 [-0.12, 0.04]
Study 05 67 297 111 407 7.9% -0.05 [-0.11, 0.02]
Study 06 22 119 67 289 4.5% -0.05 [-0.13, 0.04]
Study 07 33 188 49 183 4.6% -0.09 [-0.18, -0.01]
Study 11 12 33 10 24 0.5% -0.05 [-0.31, 0.20]
Study 12 59 277 99 373 7.5% -0.05 [-0.12, 0.01]
Study 13 74 499 101 511 15.1% -0.05 [-0.10, -0.00]
Subtotal (95% CI) 1913 2484 52.9% -0.05 [-0.08, -0.03]
Total events 383 640
Heterogeneity: Chi² = 1.16, df = 8 (P = 1.00); I² = 0%
Test for overall effect: Z = 4.28 (P < 0.0001)
1.2.2 Current and Past Smokers
Study 03 111 454 99 401 9.7% -0.00 [-0.06, 0.06]
Study 08 108 479 211 898 15.0% -0.01 [-0.06, 0.04]
Study 09 27 101 31 117 2.4% 0.00 [-0.12, 0.12]
Study 10 44 193 33 131 3.6% -0.02 [-0.12, 0.07]
Study 14 121 554 201 855 16.4% -0.02 [-0.06, 0.03]
Subtotal (95% CI) 1781 2402 47.1% -0.01 [-0.04, 0.02]
Total events 411 575
Heterogeneity: Chi² = 0.27, df = 4 (P = 0.99); I² = 0%
Test for overall effect: Z = 0.82 (P = 0.41)
Total (95% CI) 3694 4886 100.0% -0.03 [-0.05, -0.02]
Total events 794 1215
Heterogeneity: Chi² = 6.91, df = 13 (P = 0.91); I² = 0%
Test for overall effect: Z = 3.68 (P = 0.0002)
Test for subgroup differences: Chi² = 5.48, df = 1 (P = 0.02), I² = 81.8%
-0.2 -0.1 0 0.1 0.2
Favours non-smokers
Favours smokers

What is your conclusion?

What would be the best way to test the aforementioned hypothesis?

What is your conclusion?

The groups which also contain past smokers seem to have smaller and not significant effect sizes

Smoking during the IVF cycle seems to be detrimental for IVF outcome

While

There is no evidence that being a past smoker is detrimental for IVF outcome

What would be the best way to test the aforementioned hypothesis?

Direct comparison between the two groups
Current vs. Past smokers

Or
Indirect meta-analysis between studies which compare

Current vs. No smokers
AND
Past smokers vs. No smokers