Level of evidence of promising subgroup findings

Julien Tanniou
CIC Inserm 1412, University Hospital Brest

Work performed in collaboration with
Prof. dr. K.C.B. Roes, S. Smid,
dr. S. Teerenstra, and dr. I. van der Tweel.

EFSPI – Recent Developments in Biomarkers and Subgroups in Drug Development – Gothenburg – March 20th 2019

Subgroup Analyses – Regulatory Perspectives[1,2]

- Investigate the consistency of treatment effect across subgroups of clinical importance
- Explore the treatment effect across different subgroups within an overall non-statistically significant trial
- Evaluate safety profiles limited to one or a few subgroup(s)
- Establish efficacy in the targeted subgroup when included in a confirmatory testing strategy of a single trial

Subgroup Analyses – Regulatory Perspectives[3,4]

- In case the overall trial is *not* statistically significant, no further confirmatory testing is possible (type I error is exhausted).

Any step further can only be based on a case by case decision:

- A pharmacological rational
- External evidence that the subgroup is a "well-known" entity
- Stratification of the randomisation as an indicator
- Convincing p-value
- Overall outcome should claim that no harm is introduced by the experimental treatment. At least a trend toward superior efficacy of the experimental treatment for superiority trial
- Good overall safety and subgroup safety

OR

- Replication of promising subgroup findings in an independent trial

Dexamethasone for Cardiac Surgery (DECS) Study[5]

- Prophylactic corticosteroids in cardiac surgery to attenuate the inflammatory response to cardiopulmonary bypass and surgical trauma
- Randomised controlled trial in 4494 patients undergoing cardiac surgery with cardiopulmonary bypass
- Single high dose Dexamethasone (1 mg/kg) versus placebo
- Primary outcome: composite of death, myocardial infarction, stroke, renal failure, or respiratory failure, within 30 days of randomisation

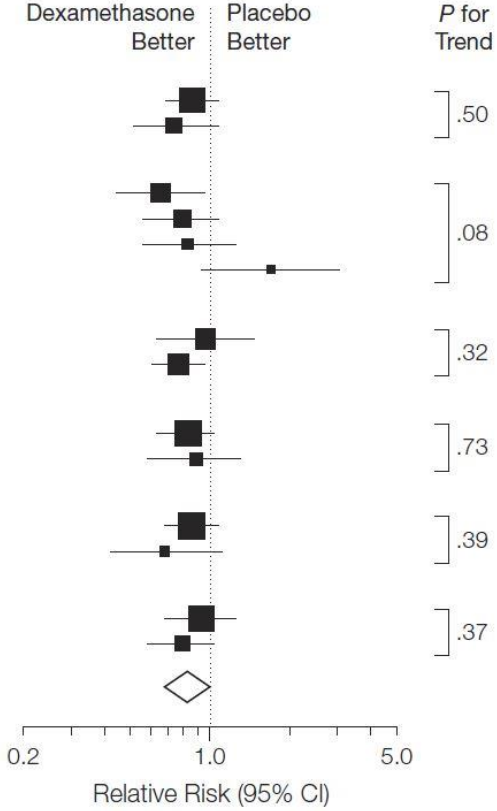
DECS Study – Primary Analysis

	No. (%) of Patients		
	Dexamethasone (n = 2235)	Placebo (n = 2247)	Relative Risk (95% CI)
Primary study end point[a]	157 (7.0)	191 (8.5)	0.83 (0.67-1.01)
Components of the primary study end point			
Death	31 (1.4)	34 (1.5)	0.92 (0.57-1.49)
Myocardial infarction	35 (1.6)	39 (1.7)	0.90 (0.57-1.42)
Stroke	29 (1.3)	32 (1.4)	0.91 (0.55-1.50)
Renal failure	28 (1.3)	40 (1.8)	0.70 (0.44-1.14)
Respiratory failure	67 (3.0)	97 (4.3)	0.69 (0.51-0.94)

[a] Primary study end point was a composite of death, myocardial infarction, stroke, renal failure, or respiratory failure, within 30 days after surgery.

DECS Study – Subgroup Analyses

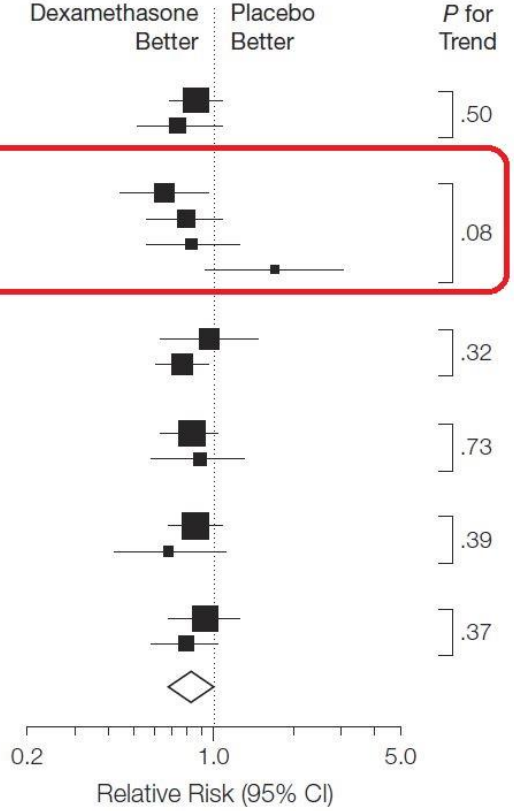
Subgroup	No. of Patients		No. of Events		Relative Risk (95% CI)	*P* for Trend
	Dexamethasone	Placebo	Dexamethasone	Placebo		
Sex						
Male	1622	1628	112	130	0.86 (0.68-1.10)	.50
Female	613	619	45	61	0.74 (0.52-1.08)	
Age, y						
<65	913	962	38	62	0.65 (0.44-0.96)	.08
65-74	818	781	57	70	0.78 (0.56-1.09)	
75-79	344	336	38	44	0.84 (0.56-1.27)	
≥80	159	168	24	15	1.69 (0.92-3.10)	
EuroScore						
0-4	1025	1060	43	46	0.97 (0.64-1.45)	.32
≥5	1180	1147	113	142	0.77 (0.61-0.98)	
Diabetes						
No	1816	1805	115	142	0.81 (0.64-1.02)	.73
Yes	415	436	41	49	0.88 (0.59-1.30)	
COPD						
No	1987	1978	133	153	0.87 (0.69-1.08)	.39
Yes	243	266	24	38	0.69 (0.43-1.12)	
Prolonged CPB						
No	1674	1659	85	92	0.92 (0.69-1.22)	.37
Yes	536	563	71	96	0.78 (0.59-1.03)	
Total	2235	2247	157	191	0.83 (0.67-1.01)	



COPD indicates chronic obstructive pulmonary disease; CPB, cardiopulmonary bypass. The effect estimates for the primary study end point in the subgroup analyses are shown. The size of each data marker correlates with the total number of patients in that subgroup.

DECS Study – Subgroup Analyses

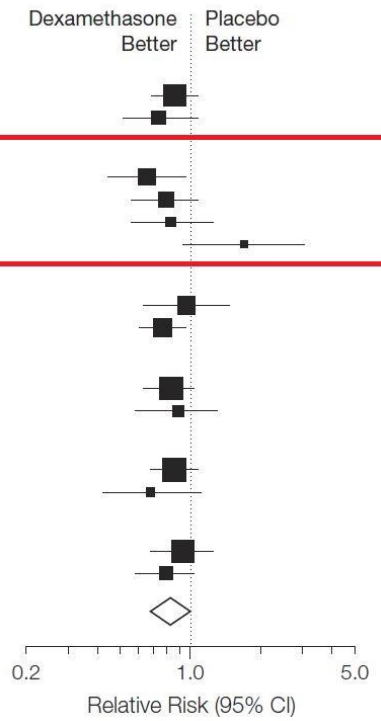
Subgroup	No. of Patients		No. of Events		Relative Risk (95% CI)	*P* for Trend
	Dexamethasone	Placebo	Dexamethasone	Placebo		
Sex						
Male	1622	1628	112	130	0.86 (0.68-1.10)	.50
Female	613	619	45	61	0.74 (0.52-1.08)	
Age, y						
<65	913	962	38	62	0.65 (0.44-0.96)	.08
65-74	818	781	57	70	0.78 (0.56-1.09)	
75-79	344	336	38	44	0.84 (0.56-1.27)	
≥80	159	168	24	15	1.69 (0.92-3.10)	
EuroScore						
0-4	1025	1060	43	46	0.97 (0.64-1.45)	.32
≥5	1180	1147	113	142	0.77 (0.61-0.98)	
Diabetes						
No	1816	1805	115	142	0.81 (0.64-1.02)	.73
Yes	415	436	41	49	0.88 (0.59-1.30)	
COPD						
No	1987	1978	133	153	0.87 (0.69-1.08)	.39
Yes	243	266	24	38	0.69 (0.43-1.12)	
Prolonged CPB						
No	1674	1659	85	92	0.92 (0.69-1.22)	.37
Yes	536	563	71	96	0.78 (0.59-1.03)	
Total	2235	2247	157	191	0.83 (0.67-1.01)	



COPD indicates chronic obstructive pulmonary disease; CPB, cardiopulmonary bypass. The effect estimates for the primary study end point in the subgroup analyses are shown. The size of each data marker correlates with the total number of patients in that subgroup.

DECS Study – Subgroup Analyses (2)

Subgroup	No. of Patients		No. of Events		Relative Risk (95% CI)	P for Trend
	Dexamethasone	Placebo	Dexamethasone	Placebo		
Sex						
Male	1622	1628	112	130	0.86 (0.68-1.10)	.50
Female	613	619	45	61	0.74 (0.52-1.08)	
Age, y						
<65	913	962	38	62	0.65 (0.44-0.96)	.08
65-74	818	781	57	70	0.78 (0.56-1.09)	
75-79	344	336	38	44	0.84 (0.56-1.27)	
≥80	159	168	24	15	1.69 (0.92-3.10)	
EuroScore						
0-4	1025	1060	43	46	0.97 (0.64-1.45)	.32
≥5	1180	1147	113	142	0.77 (0.61-0.98)	
Diabetes						
No	1816	1805	115	142	0.81 (0.64-1.02)	.73
Yes	415	436	41	49	0.88 (0.59-1.30)	
COPD						
No	1987	1978	133	153	0.87 (0.69-1.08)	.39
Yes	243	266	24	38	0.69 (0.43-1.12)	
Prolonged CPB						
No	1674	1659	85	92	0.92 (0.69-1.22)	.37
Yes	536	563	71	96	0.78 (0.59-1.03)	
Total	2235	2247	157	191	0.83 (0.67-1.01)	



COPD indicates chronic obstructive pulmonary disease; CPB, cardiopulmonary bypass. The effect estimates for the primary study end point in the subgroup analyses are shown. The size of each data marker correlates with the total number of patients in that subgroup.

In patients younger than 65 years, the RR for mortality was 0.42 (95% CI, 0.13-1.34; p=0.13), but it gradually increased with age to 3.87 (95% CI, 1.10-13.6; p=0.02) in patients aged 80 years or older (p for trend=0.05).

DECS Study – Subgroup Analyses (3)

Questions raised:

- How convincing the results concerning patients younger than 65 years are?
- Does the evidence support a decision to apply Dexamethasone in the young but not in the elderly?

Practical context:

- As we will never be able to foresee and pre-plan everything, how to deal with observed data post-hoc?

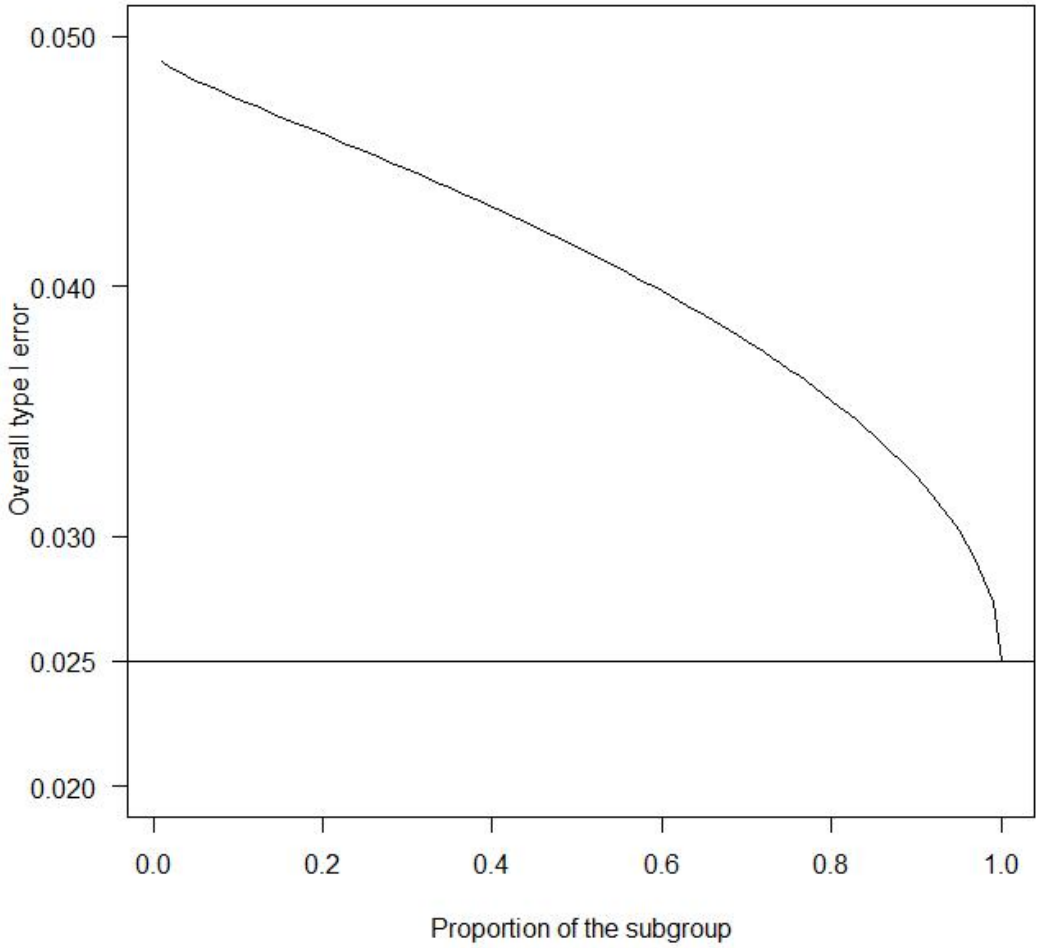
Overall Type I Error[6]

Under the null hypothesis, $P_0[\bar{X}_{T_0} - \bar{X}_{C_0} < c \; and \; \bar{X}_{T_1} - \bar{X}_{C_1} > d]$ can be derived from the joint distribution of $\bar{X}_{T_0} - \bar{X}_{C_0}$ and $\bar{X}_{T_1} - \bar{X}_{C_1}$ which is N (0, Σ) with

$$\Sigma = \begin{bmatrix} \frac{2}{n} & \frac{2}{n} \\ \frac{2}{n} & \frac{2}{p \cdot n} \end{bmatrix}$$

From this result, the correlation between $\bar{X}_{T_0} - \bar{X}_{C_0}$ and $\bar{X}_{T_1} - \bar{X}_{C_1}$ is $\sqrt{p}$. Therefore, it is expected that a higher p will result in a smaller $P_0[\bar{X}_{T_0} - \bar{X}_{C_0} < c \; and \; \bar{X}_{T_1} - \bar{X}_{C_1} > d]$, and hence a smaller overall type I error.

Theoretical Overall Type I Error (one subgroup)



Overall Type I Error (more than one subgroup)[7]

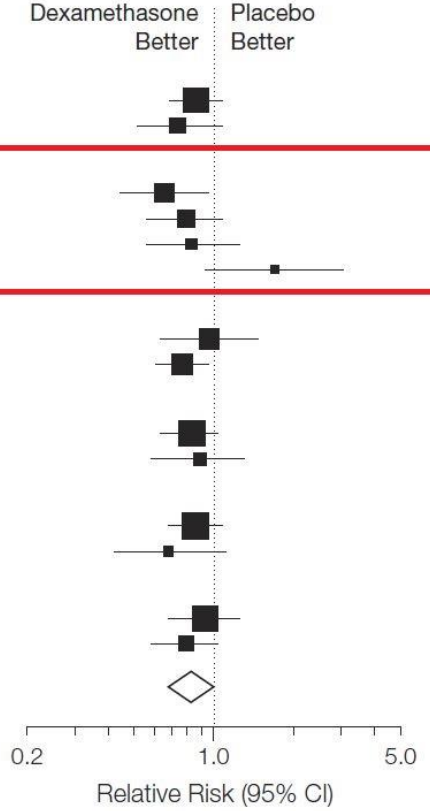
Number of subgroup(s)	Proportion of the subgroup(s) of interest								
	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
1	0.047	0.046	0.045	0.043	0.042	0.040	0.038	0.035	0.032
2	0.068	0.065	0.062	0.060	0.057	0.053	0.049	0.044	0.038
3	0.088	0.084	0.080	0.077	0.071	0.065	0.059	0.052	0.043
5	0.129	0.121	0.111	0.106	0.095	0.086	0.075	0.065	0.051
10	0.214	0.201	0.180	0.163	0.146	0.124	0.107	0.087	0.064
20	0.358	0.312	0.275	0.247	0.209	0.176	0.147	0.113	0.077

Overall Type I Error (more than one subgroup)[7]

Number of subgroup(s)	Indep Tests	Proportion of subgroup(s) of interest								
		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
1	0.049	0.047	0.046	0.045	0.043	0.042	0.040	0.038	0.035	0.032
2	0.073	0.068	0.065	0.062	0.060	0.057	0.053	0.049	0.044	0.038
3	0.096	0.088	0.084	0.080	0.077	0.071	0.065	0.059	0.052	0.043
5	0.141	0.129	0.121	0.111	0.106	0.095	0.086	0.075	0.065	0.051
10	0.243	0.214	0.201	0.180	0.163	0.146	0.124	0.107	0.087	0.064
20	0.412	0.358	0.312	0.275	0.247	0.209	0.176	0.147	0.113	0.077

DECS Study – Monotonic Trend

Subgroup	No. of Patients		No. of Events		Relative Risk (95% CI)	*P* for Trend
	Dexamethasone	Placebo	Dexamethasone	Placebo		
Sex						
Male	1622	1628	112	130	0.86 (0.68-1.10)	.50
Female	613	619	45	61	0.74 (0.52-1.08)	
Age, y						
<65	913	962	38	62	0.65 (0.44-0.96)	.08
65-74	818	781	57	70	0.78 (0.56-1.09)	
75-79	344	336	38	44	0.84 (0.56-1.27)	
≥80	159	168	24	15	1.69 (0.92-3.10)	
EuroScore						
0-4	1025	1060	43	46	0.97 (0.64-1.45)	.32
≥5	1180	1147	113	142	0.77 (0.61-0.98)	
Diabetes						
No	1816	1805	115	142	0.81 (0.64-1.02)	.73
Yes	415	436	41	49	0.88 (0.59-1.30)	
COPD						
No	1987	1978	133	153	0.87 (0.69-1.08)	.39
Yes	243	266	24	38	0.69 (0.43-1.12)	
Prolonged CPB						
No	1674	1659	85	92	0.92 (0.69-1.22)	.37
Yes	536	563	71	96	0.78 (0.59-1.03)	
Total	2235	2247	157	191	0.83 (0.67-1.01)	



COPD indicates chronic obstructive pulmonary disease; CPB, cardiopulmonary bypass. The effect estimates for the primary study end point in the subgroup analyses are shown. The size of each data marker correlates with the total number of patients in that subgroup.

Monotonic Trend – Observed[7]

Scenarios 1-3: observed trend in the data across subgroups:

1. RR1 < RR2 < RR3;
2. OR1 < OR2 < OR3;
3. RD1 < RD2 < RD3.

Monotonic Trend – Observed + Interaction

Scenarios 4-6: observed trend in the data across subgroups supported by a statistically significant interaction test:

4. RR1 < RR2 < RR3 and a statistically significant interaction (Poisson regression analysis);
5. OR1 < OR2 < OR3 and a statistically significant interaction (logistic regression analysis);
6. RD1 < RD2 < RD3 and a statistically significant interaction (binomial regression analysis).

Monotonic Trend – Observed + Subgroup

Scenarios 7-9: observed trend in the data across subgroups supported by a statistically significant test for the effect in the most promising subgroup:

7. RR1 < RR2 < RR3 and a statistically significant effect for the most promising subgroup;
8. OR1 < OR2 < OR3 and a statistically significant effect for the most promising subgroup;
9. RD1 < RD2 < RD3 and a statistically significant effect for the most promising subgroup.

Monotonic Trend – Simulations

Parameters	Values		
Significance level of the overall test	0.05 (two-sided)		
Power of the overall test	0.80		
Proportions of responders, effect size, and sample size per arm	$\pi_c = 0.2$ $\pi_t = 0.25$	$h = 0.12$	$N = 1091$
	$\pi_c = 0.5$ $\pi_t = 0.625$	$h = 0.25$	$N = 244$
	$\pi_c = 0.2$ $\pi_t = 0.4$	$h = 0.44$	$N = 79$

Note: π_c: proportion of responders control arm; π_t: proportion of responders treatment arm; h: Cohen's effect size; N: sample size per arm.

Monotonic Trend – Simulated Proportions

k	$\pi_c = 0.2; \pi_t = 0.25;$ $N = 1091$	$\pi_c = 0.5; \pi_t = 0.625;$ $N = 244$	$\pi_c = 0.2; \pi_t = 0.4;$ $N = 79$
0	$p_{G_{1T}} = 0.2$ $p_{G_{2T}} = 0.2$ $p_{G_{3T}} = 0.2$ $p_{G_{1C}} = p_{G_{2C}} = p_{G_{3C}} = 0.2$	$p_{G_{1T}} = 0.5$ $p_{G_{2T}} = 0.5$ $p_{G_{3T}} = 0.5$ $p_{G_{1C}} = p_{G_{2C}} = p_{G_{3C}} = 0.5$	$p_{G_{1T}} = 0.2$ $p_{G_{2T}} = 0.2$ $p_{G_{3T}} = 0.2$ $p_{G_{1C}} = p_{G_{2C}} = p_{G_{3C}} = 0.2$
0.5	$p_{G_{1T}} = 0.225$ $p_{G_{2T}} = 0.2125$ $p_{G_{3T}} = 0.2$ $p_{G_{1C}} = p_{G_{2C}} = p_{G_{3C}} = 0.2$	$p_{G_{1T}} = 0.5625$ $p_{G_{2T}} = 0.53125$ $p_{G_{3T}} = 0.5$ $p_{G_{1C}} = p_{G_{2C}} = p_{G_{3C}} = 0.5$	$p_{G_{1T}} = 0.3$ $p_{G_{2T}} = 0.25$ $p_{G_{3T}} = 0.2$ $p_{G_{1C}} = p_{G_{2C}} = p_{G_{3C}} = 0.2$
1	$p_{G_{1T}} = 0.25$ $p_{G_{2T}} = 0.225$ $p_{G_{3T}} = 0.2$ $p_{G_{1C}} = p_{G_{2C}} = p_{G_{3C}} = 0.2$	$p_{G_{1T}} = 0.625$ $p_{G_{2T}} = 0.5625$ $p_{G_{3T}} = 0.5$ $p_{G_{1C}} = p_{G_{2C}} = p_{G_{3C}} = 0.5$	$p_{G_{1T}} = 0.4$ $p_{G_{2T}} = 0.3$ $p_{G_{3T}} = 0.2$ $p_{G_{1C}} = p_{G_{2C}} = p_{G_{3C}} = 0.2$
1.5	$p_{G_{1T}} = 0.275$ $p_{G_{2T}} = 0.2375$ $p_{G_{3T}} = 0.2$ $p_{G_{1C}} = p_{G_{2C}} = p_{G_{3C}} = 0.2$	$p_{G_{1T}} = 0.6875$ $p_{G_{2T}} = 0.59375$ $p_{G_{3T}} = 0.5$ $p_{G_{1C}} = p_{G_{2C}} = p_{G_{3C}} = 0.5$	$p_{G_{1T}} = 0.5$ $p_{G_{2T}} = 0.35$ $p_{G_{3T}} = 0.2$ $p_{G_{1C}} = p_{G_{2C}} = p_{G_{3C}} = 0.2$
2	$p_{G_{1T}} = 0.3$ $p_{G_{2T}} = 0.25$ $p_{G_{3T}} = 0.2$ $p_{G_{1C}} = p_{G_{2C}} = p_{G_{3C}} = 0.2$	$p_{G_{1T}} = 0.75$ $p_{G_{2T}} = 0.625$ $p_{G_{3T}} = 0.5$ $p_{G_{1C}} = p_{G_{2C}} = p_{G_{3C}} = 0.5$	$p_{G_{1T}} = 0.6$ $p_{G_{2T}} = 0.4$ $p_{G_{3T}} = 0.2$ $p_{G_{1C}} = p_{G_{2C}} = p_{G_{3C}} = 0.2$

Note: π_c: proportion of responders control arm; π_t: proportion of responders treatment arm; N: sample size per arm; k: parameter related to the simulated proportion of responders for each subgroup (under the null: $k = 0$; under the alternative: $k = \{0.5, 1, 1.5, 2\}$).

Monotonic Trend – Results

			$N = 1091$		$N = 244$		$N = 79$	
Strategy	**Metric**	k	$r_{G_1} = 0.33$	$r_{G_1} = 0.5$	$r_{G_1} = 0.33$	$r_{G_1} = 0.5$	$r_{G_1} = 0.33$	$r_{G_1} = 0.5$
Observed	RR	0	0.22	0.20	0.21	0.20	0.20	0.19
		1	0.46	0.43	0.47	0.44	0.46	0.42
		2	0.74	0.69	0.80	0.75	0.75	0.68
	OR	0	0.22	0.20	0.21	0.20	0.20	0.18
		1	0.44	0.43	0.46	0.43	0.42	0.38
		2	0.72	0.68	0.78	0.73	0.69	0.61
	RD	0	0.21	0.19	0.21	0.19	0.18	0.17
		1	0.44	0.42	0.44	0.42	0.42	0.39
		2	0.73	0.68	0.75	0.71	0.70	0.64

Note: N: sample size per arm; r_{G_1}: proportion (over the total sample) of the subgroup of interest G_1; k: parameter related to the simulated proportion of responders for each subgroup (under the null: $k = 0$; under the alternative: $k = \{0.5, 1, 1.5, 2\}$); NA: not applicable; OR: odds ratio; RD: risk difference; RR: relative risk.

Monotonic Trend – Results

			$N = 1091$		$N = 244$		$N = 79$	
Strategy	**Metric**	k	$r_{G_1} = 0.33$	$r_{G_1} = 0.5$	$r_{G_1} = 0.33$	$r_{G_1} = 0.5$	$r_{G_1} = 0.33$	$r_{G_1} = 0.5$
Observed	RR	0	0.22	0.20	0.21	0.20	0.20	0.19
		1	0.46	0.43	0.47	0.44	0.46	0.42
		2	0.74	0.69	0.80	0.75	0.75	0.68
	OR	0	0.22	0.20	0.21	0.20	0.20	0.18
		1	0.44	0.43	0.46	0.43	0.42	0.38
		2	0.72	0.68	0.78	0.73	0.69	0.61
	RD	0	0.21	0.19	0.21	0.19	0.18	0.17
		1	0.44	0.42	0.44	0.42	0.42	0.39
		2	0.73	0.68	0.75	0.71	0.70	0.64
Observed + interaction ($p < 0.10$)	RR	0	0.09	0.08	0.09	0.09	0.09	0.08
		1	0.24	0.24	0.26	0.25	0.26	0.24
		2	0.62	0.59	0.71	0.67	0.62	0.59
	OR	0	0.09	0.09	0.09	0.08	0.09	0.08
		1	0.24	0.22	0.25	0.23	0.22	0.20
		2	0.58	0.55	0.69	0.66	0.54	0.48
	RD	0	0.09	0.08	0.09	0.08	0.09	0.08
		1	0.24	0.23	0.25	0.23	0.24	0.21
		2	0.60	0.57	0.66	0.64	0.59	0.54

Note: N: sample size per arm; r_{G_1}: proportion (over the total sample) of the subgroup of interest G_1; k: parameter related to the simulated proportion of responders for each subgroup (under the null: $k = 0$; under the alternative: $k = \{0.5, 1, 1.5, 2\}$); NA: not applicable; OR: odds ratio; RD: risk difference; RR: relative risk.

Monotonic Trend – Results

Strategy	Metric	k	$N = 1091$		$N = 244$		$N = 79$	
			$r_{G_1} = 0.33$	$r_{G_1} = 0.5$	$r_{G_1} = 0.33$	$r_{G_1} = 0.5$	$r_{G_1} = 0.33$	$r_{G_1} = 0.5$
Observed	RR	0	0.22	0.20	0.21	0.20	0.20	0.19
		1	0.46	0.43	0.47	0.44	0.46	0.42
		2	0.74	0.69	0.80	0.75	0.75	0.68
	OR	0	0.22	0.20	0.21	0.20	0.20	0.18
		1	0.44	0.43	0.46	0.43	0.42	0.38
		2	0.72	0.68	0.78	0.73	0.69	0.61
	RD	0	0.21	0.19	0.21	0.19	0.18	0.17
		1	0.44	0.42	0.44	0.42	0.42	0.39
		2	0.73	0.68	0.75	0.71	0.70	0.64
Observed + interaction ($p < 0.10$)	RR	0	0.09	0.08	0.09	0.09	0.09	0.08
		1	0.24	0.24	0.26	0.25	0.26	0.24
		2	0.62	0.59	0.71	0.67	0.62	0.59
	OR	0	0.09	0.09	0.09	0.08	0.09	0.08
		1	0.24	0.22	0.25	0.23	0.22	0.20
		2	0.58	0.55	0.69	0.66	0.54	0.48
	RD	0	0.09	0.08	0.09	0.08	0.09	0.08
		1	0.24	0.23	0.25	0.23	0.24	0.21
		2	0.60	0.57	0.66	0.64	0.59	0.54
Observed + subgroup	RR	0	0.06	0.06	0.06	0.06	0.06	0.06
		1	0.17	0.19	0.18	0.19	0.17	0.16
		2	0.56	0.58	0.65	0.67	0.51	0.54
	OR	0	0.06	0.06	0.06	0.06	0.06	0.06
		1	0.17	0.19	0.17	0.19	0.17	0.17
		2	0.56	0.57	0.64	0.66	0.51	0.51
	RD	0	0.06	0.06	0.06	0.06	0.06	0.06
		1	0.17	0.19	0.17	0.19	0.17	0.16
		2	0.56	0.57	0.64	0.65	0.51	0.53

Note: N: sample size per arm; r_{G_1}: proportion (over the total sample) of the subgroup of interest G_1; k: parameter related to the simulated proportion of responders for each subgroup (under the null: $k = 0$; under the alternative: $k = \{0.5, 1, 1.5, 2\}$); NA: not applicable; OR: odds ratio; RD: risk difference; RR: relative risk.

Conclusions

- The purpose of subgroup investigations should be made clearly at the design stage[2,3,8]
- Subgroup investigations in overall non-statistically significant studies are exploratory only
- Concept of credibility is of paramount importance
- Overall type I error not as high as one would expect due to correlation between test statistics, especially when multiple subgroups are investigated
- When a covariate is composed of more than two subgroups, such as biomarker levels, several options exist

Conclusions

- An observed (monotonic) trend across subgroups together with a statistically significant treatment effect in the most promising subgroup might indicate the presence of a potential signal
- Non-statistical considerations such as the biological/pharmacological rational, together with (at least partial) replication of the subgroup finding(s) should be provided to strengthen the subgroup finding

References

[1] J.M. Grouin, M. Coste, and J. Lewis. Subgroup analyses in randomized clinical trials: statistical and regulatory issues. J Biopharm Stat, 15(5):869–882, 2005.

[2] J. Tanniou, I. van der Tweel, S. Teerenstra and K.C.B. Roes. Subgroup analyses in confirmatory clinical trials: time to be specific about their purposes. BMC Medical Research Methodology, 16:20, 2016.

[3] European Medicines Agency. Guideline on the investigation of subgroups in confirmatory clinical trials. 2019.

[4] A. Koch, T. Framke. Reliably basing conclusions on subgroups of randomized clinical trials. J Biopharm Stat. 2014;24:42–57.

[5] J.M. Dieleman, A.P. Nierich, P.M. Rosseel, et al. Intraoperative high-dose dexamethasone for cardiac surgery: a randomized controlled trial. JAMA, 308(17):1761–1767, 2012.

References (2)

[6] J. Tanniou, I. van der Tweel, S. Teerenstra, K.C.B. Roes. Level of evidence for promising subgroup findings in an overall non-significant trial. Stat Methods Med Res. 2016;25(5):2193–2213.

[7] J. Tanniou J, S.C. Smid, I. van der Tweel, S. Teerenstra, K.C.B. Roes. Level of evidence for promising subgroup findings: The case of trends and multiple subgroups. Statistics in Medicine. 2019; 1– 12.

[8] A. Dmitrienko, C. Muysers, A. Fritsch, I. Lipkovich. General guidance on exploratory and confirmatory subgroup analysis in late-stage clinical trials. J Biopharm Stat. 2016;26(1):71-98.