Planned subgroup analyses in randomized clinical trials: A comparison between early 2000s, 2012 and 2016

Ala Taji Heravi[1], Dmitry Gryaznov[1], Stefan Schandelmaier[1,2], Matthias Briel[1,2] and the ASPIRE study group.
1 Department of Clinical Research, Basel Institute for Clinical Epidemiology and Biostatistics, University Hospital Basel and University of Basel, Basel, Switzerland
2 Department of Health Research Methods, Evidence, and Impact, McMaster University, Hamilton, Canada

Background

Clearly structured and comprehensive protocols are important to guide the conduct and analysis of randomized clinical trials (RCTs). Subgroup analysis (SGA) is playing a role in identifying consistency or differences in subgroups of patients to ultimately better target clinical recommendations and clinical decision making. However, in a previous study with RCT protocols approved between 2000 and 2003, our group found that almost 30% of protocols included one or more planned subgroup analyses, which were insufficiently described. In the present study, we investigated if the prevalence and description of subgroup analyses in RCT protocols from 2012 and 2016, differ from those approved in the early 2000s.

Methods

We used existing databases of about 1500 RCT protocols (894 RCTs: early 2000s; 257 RCTs: 2012; 292 RCTs: 2016) that were approved by research ethics committees (RECs) in Switzerland, Germany (Freiburg) and Canada (Hamilton) in early 2000s, 2012 and 2016. To build these databases, investigators trained in clinical research methodology recorded RCT characteristics alongside by whether any subgroup analysis is mentioned in the protocol and, if so, we extracted some information for the credibility of subgroup analyses, e.g. whether a clear hypothesis for a subgroup effect is prespecified. We descriptively summarized RCT characteristics with respect to planned subgroup analyses using R version 4.0.2.

Trial characteristics	2000-2003			2012			2016		
	SGA not planned n=642	SGA planned n=252 (28%)	All trials n= 894	SGA not planned n=164	SGA planned n= 93 (36%)	All trials n= 257	SGA not planned n=196	SGA planned n=96 (32%)	All trials n= 292
Target sample size									
Median	200	521	260	165	600	300	164	303	199
Q1, Q3	80 , 471	229 , 1030	100 , 610	71.5 , 432	354,1500	100,720	74.8, 416	150 , 600	100,490
Center status									
Multicenter	500 (77.9)	241 (95.6)	741 (82.9)	119 (72.6)	91 (97.8)	210 (81.7)	131 (66.8)	84 (87.5)	215(73.6)
Single center	139 (21.7)	10(4.0)	149 (16.7)	45 (27.4)	2 (2.2)	47 (18.3)	65 (33.2)	12 (12.5)	77 (26.4)
Unclear	3(0.5)	1(0.4)	4(0.4)	0	0	0	0	0	0
Sponsorship									
Industry	356 (55.5)	195 (77.4)	551 (61.6)	69 (42.1)	69 (74.2)	138 (53.7)	73 (37.2)	57 (59.4)	130 (44.5)
Investigator	286 (44.5)	57 (22.6)	343 (38.4)	95 (57.9)	24 (25.8)	119 (46.3)	123 (62.8)	39 (40.6)	162 (55.5)
Clinical area									
Oncology	113 (17.6)	42 (16.7)	155 (17.3)	22 (13.4)	25 (26.9)	47 (18.3)	27 (13.8)	24 (25.0)	51 (17.5)
Cardiovascular	59 (9.2)	49 (19.5)	108 (12.1)	8 (4.9)	19 (20.4)	27 (10.5)	15 (7.7)	20 (20.8)	35 (12.0)
Infectious diesease	60 (9.3)	27 (10.8)	87 (9.7)	6 (3.7)	3 (3.2)	9 (3.5)	4 (2.0)	3 (3.1)	7 (2.4)
Endocrinology	47 (7.3)	15 (6.0)	62 (6.9)	4 (2.4)	1 (1.1)	5 (1.9)	6 (3.1)	3 (3.1)	9 (3.1)
Neurology	37 (5.8)	24 (9.6)	61 (6.8)	13 (7.9)	7 (7.5)	20 (7.8)	17 (8.7)	5 (5.2)	31 (10.6)
Surgery	75 (11.7)	18 (7.2)	93 (10.4)	27 (16.5)	10 (10.8)	37 (14.4)	21 (10.7)	10 (10.4)	22(7.5)
Pediatrics	34 (5.3)	11 (4.4)	45 (5.0)	11 (6.7)	3 (3.2)	14 (5.4)	11 (5.6)	8 (8.3)	19 (6.5)
Other	217 (33.8)	66(26.3)	283 (31.7)	73 (44.5)	25 (26.9)	98 (38.1)	95 (48.5)	23 (24.0)	118 (40.4)

Results

About a third of RCT protocols included one or more planned subgroup analyses in 2012 and 2016, which remained unchanged comparing to early 2000s (28.2% in early 2000s; 36.19% in 2012; 32.88% in 2016) (**Table1**). RCTs that planned at least one subgroup analysis were typically industry-sponsored, multicenter trials with larger sample size. In protocols of oncology and cardiovascular RCTs planned subgroup analyses are most prevalent (50% or more include subgroup analyses). A clear subgroup hypothesis was provided only in 6.7%, 9.7% and 16.7% and an anticipated direction of a potential subgroup effect was provided in 4%, 9.7% and 14.7% of protocols planning at least one subgroup analysis in early 2000s, 2012 and 2016, respectively. An appropriate interaction test was specified in about a third of RCT protocols with planned subgroup analyses in all three time periods, and the median number of planned subgroup analyses increased from the early 2000s to 2012 and 2016 from 3 (interquartile range [IQR], 1-6) to 6 (IQR, 3-13).

Conclusions

The proportion of RCT protocols with planned subgroup analyses in RCTs appeared to remain stable at about a third. The reporting of important characteristics of planned subgroup analyses has slightly improved over time in terms of providing clear hypotheses and an anticipated direction of a potential subgroup effect, but is still unsatisfactory. We did not find any improvement over time in terms of the specification of interaction tests, and the doubling of the median number of planned subgroup analyses from 3 to 6 further raises concerns about the credibility of potential subgroup findings.

Ala Taji Heravi
Department of Clinical Research,
Basel Institute for Clinical Epidemiology & Biostatistics
University Hospital Basel, Switzerland
ala.tajiheravi@usb.ch **alaheravi**