HW6

Zhuodiao Kuang zk2275

2023-11-04

Question 1

In order to understand the dose-response relationship in abecma discussed in Slides 16-19, what would be your approaches to address the reviewers' questions?

- a. Potential confounding factors for the observed dose-response in efficacy

There are more than 100 baseline characteristics, according to the information given in the slides. Therefore, first, I will identify potential confounding factors and reduce the number of influential characteristics, using the method of hypothesis testing or step-wise variable selection.

Second, I would identify prognostic factors, which actually predicted the outcomes. Additionally, each factor was tested individually by controlling the dose, and significant factors were kept as the confounding factors.

Third, I would control all the selected confounding factors, because some of their dose-response relationships were more significant for certain endpoints.

Last, multiplicity issues should be discussed for multiple tests.

- b. If there were any dose by baseline interactions

The interactions can exist in any forms, but I think the simplest method to test them is to explore the correlation, make visual inspection of the data, and refer to the reality and the definition of each factor.

Once the calculation or visualization feels it, we can apply statistical analysis to get statistically significant interactions. However, We may not finally get the exact form of interactions, even we use state-of-art methods.

Question 2

a) Please estimate the survival time that 90% subjects survived for baseline survival function.

Obs	id	time	status	trt	age	sex	ascites	hepato	spiders	edema	bili	chol	albumin	copper	alk.phos	ast	trig	platelet	protime	stage
1	1	400	2	1	58.7652293	f	1	1	1	1	14.5	261	2.6	156	1718	137.95	172	190	12.2	4
2	2	4500	0	1	56.44626968	f	0	1	1	0	1.1	302	4.14	54	7394.8	113.52	88	221	10.6	3
3	3	1012	2	1	70.07255305	m	0	0	0	0.5	1.4	176	3.48	210	516	96.1	55	151	12	4
4	4	1925	2	1	54.74058864	f	0	1	1	0.5	1.8	244	2.54	64	6121.8	60.63	92	183	10.3	4
5	5	1504	1	2	38.10540726	f	0	1	1	0	3.4	279	3.53	143	671	113.15	72	136	10.9	3
6	6	2503	2	2	66.2587269	f	0	1	0	0	0.8	248	3.98	50	944	93	63	.	11	3
7	7	1832	0	2	55.53456537	f	0	1	0	0	1	322	4.09	52	824	60.45	213	204	9.7	3
8	8	2466	2	2	53.0568104	f	0	0	0	0	0.3	280	4	52	4651.2	28.38	189	373	11	3
9	9	2400	2	1	42.50787132	f	0	0	1	0	3.2	562	3.08	79	2276	144.15	88	251	11	2
10	10	51	2	2	70.55989049	f	1	0	1	1	12.6	200	2.74	140	918	147.25	143	302	11.5	4
11	11	3762	2	2	53.71389459	f	0	1	1	0	1.4	259	4.16	46	1104	79.05	79	258	12	4
12	12	304	2	2	59.137577	f	0	0	1	0	3.6	236	3.52	94	591	82.15	95	71	13.6	4
13	13	3577	0	2	45.68925394	f	0	0	0	0	0.7	281	3.85	40	1181	88.35	130	244	10.6	3
14	14	1217	2	2	56.22176591	m	1	1	0	1	0.8	.	2.27	43	728	71	.	156	11	4
15	15	3584	2	1	64.64613279	f	0	0	0	0	0.8	231	3.87	173	9009.8	127.71	96	295	11	3
16	16	3672	0	2	40.44353183	f	0	0	0	0	0.7	204	3.66	28	685	72.85	58	198	10.8	3
17	17	769	2	2	52.183436	f	0	1	0	0	2.7	274	3.15	159	1533	117.8	128	224	10.5	4
18	18	131	2	1	53.9301848	f	0	1	1	1	11.4	178	2.8	588	961	280.55	200	283	12.4	4
19	19	4232	0	1	49.56057495	f	0	1	0	0.5	0.7	235	3.56	39	1881	93	123	209	11	3
20	20	1356	2	2	59.95345654	f	0	1	0	0	5.1	374	3.51	140	1919	122.45	135	322	13	4
21	21	3445	0	2	64.1889117	m	0	1	1	0	0.6	252	3.83	41	843	65.1	83	336	11.4	4

Figure 1: Crude Survival Data

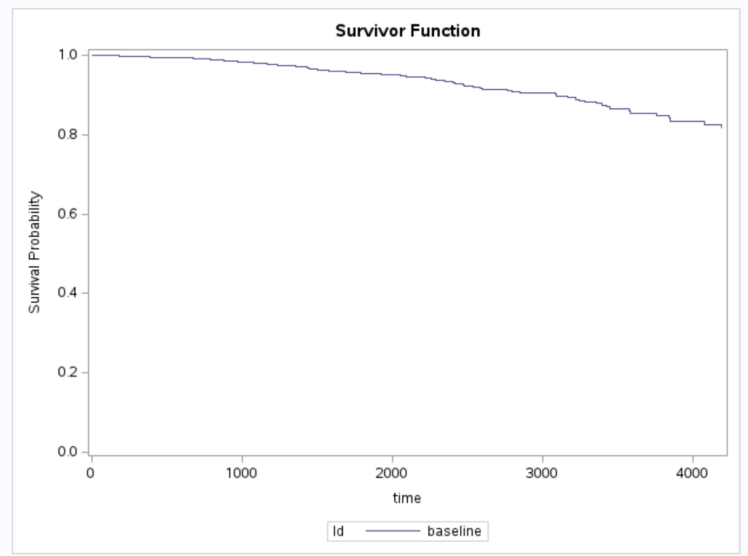

Figure 2: Survival Fucntion for Baseline Covariates

Based on the results above, it is projected that 90% of individuals survived for 3090 days during the baseline survival function.

Obs	Id	sexn	edema	bili	albumin	copper	ast	protime	stage	time	Survival	StdErrSurvival	LowerSurvival	UpperSurvival
120	baseline	0	0	0	0	0	0	0	0	2689	0.91255	0.12054	0.70440	1
121	baseline	0	0	0	0	0	0	0	0	2769	0.91004	0.12389	0.69691	1
122	baseline	0	0	0	0	0	0	0	0	2796	0.90748	0.12730	0.68934	1
123	baseline	0	0	0	0	0	0	0	0	2847	0.90485	0.13079	0.68161	1
124	baseline	0	0	0	0	0	0	0	0	3086	0.90192	0.13469	0.67306	1
125	baseline	0	0	0	0	0	0	0	0	3090	0.89898	0.13858	0.66456	1
126	baseline	0	0	0	0	0	0	0	0	3092	0.89588	0.14269	0.65565	1
127	baseline	0	0	0	0	0	0	0	0	3170	0.89267	0.14693	0.64652	1
128	baseline	0	0	0	0	0	0	0	0	3222	0.88934	0.15130	0.63718	1
129	baseline	0	0	0	0	0	0	0	0	3244	0.88598	0.15570	0.62782	1
130	baseline	0	0	0	0	0	0	0	0	3282	0.88248	0.16027	0.61818	1

Figure 3: Targeted interval for Baseline

b) Please estimate the survival time that 90% subjects survived for male subjects who had edema and stage=4 and taking median for all other covariates in the model.

The MEANS Procedure

Variable	Median
bili	1.4000000
albumin	3.5300000
copper	73.0000000
ast	114.7000000
protime	10.6000000

Figure 4: Taking Median for All Other Covariates

Based on the results, the estimated survival time for male subjects with edema and stage = 4, after adjusting for medians of all other covariates, was 348 days, for 90% of the subjects.

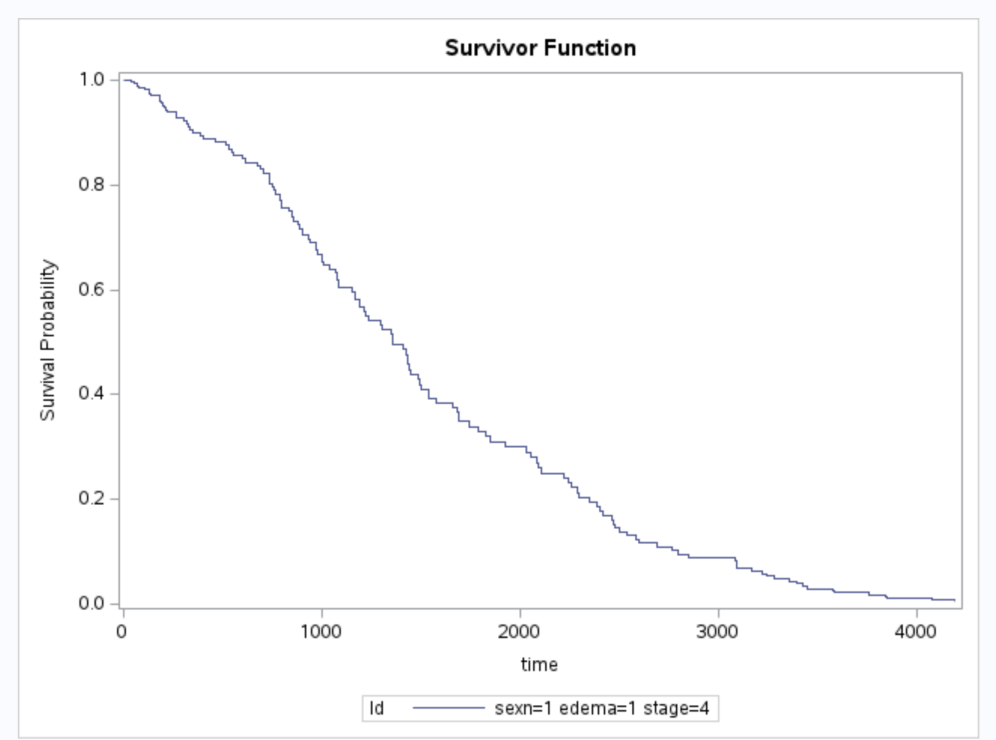

Figure 5: Survival Fucntion for Male Subjects

Obs	Id	sexn	edema	bili	albumin	copper	ast	protime	stage	time	Survival	StdErrSurvival	LowerSurvival	UpperSurvival
15	sexn=1 edema=1 stage=4	1	1	1.4	3.53	73	114.7	10.6	4	216	0.94322	0.026670	0.89237	0.99697
16	sexn=1 edema=1 stage=4	1	1	1.4	3.53	73	114.7	10.6	4	223	0.93848	0.028464	0.88431	0.99596
17	sexn=1 edema=1 stage=4	1	1	1.4	3.53	73	114.7	10.6	4	264	0.92890	0.032044	0.86817	0.99387
18	sexn=1 edema=1 stage=4	1	1	1.4	3.53	73	114.7	10.6	4	304	0.92314	0.034148	0.85858	0.99255
19	sexn=1 edema=1 stage=4	1	1	1.4	3.53	73	114.7	10.6	4	321	0.91739	0.036214	0.84909	0.99119
20	sexn=1 edema=1 stage=4	1	1	1.4	3.53	73	114.7	10.6	4	326	0.91164	0.038251	0.83967	0.98978
21	sexn=1 edema=1 stage=4	1	1	1.4	3.53	73	114.7	10.6	4	334	0.90589	0.040263	0.83031	0.98834
22	sexn=1 edema=1 stage=4	1	1	1.4	3.53	73	114.7	10.6	4	348	0.89994	0.042347	0.82065	0.98688
23	sexn=1 edema=1 stage=4	1	1	1.4	3.53	73	114.7	10.6	4	388	0.89396	0.044417	0.81101	0.98540
24	sexn=1 edema=1 stage=4	1	1	1.4	3.53	73	114.7	10.6	4	400	0.88798	0.046469	0.80142	0.98389
25	sexn=1 edema=1 stage=4	1	1	1.4	3.53	73	114.7	10.6	4	460	0.88171	0.048630	0.79137	0.98236

Figure 6: Targeted interval for Male Subjects

Question 3

Perform model selection using backward selection using PBC data.

- Step 0. The model contains the following effects:

trt sexn ascites hepato spiders edema bili albumin copper alk.phos ast protime stage

- Step 1. Effect ascites is removed. The model contains the following effects:

trt sexn hepato spiders edema bili albumin copper alk.phos ast protime stage

- Step 2. Effect trt is removed. The model contains the following effects:

sexn hepato spiders edema bili albumin copper alk.phos ast protime stage

- Step 3. Effect spiders is removed. The model contains the following effects:

sexn hepato edema bili albumin copper alk.phos ast protime stage

- Step 4. Effect hepato is removed. The model contains the following effects:

sexn edema bili albumin copper alk.phos ast protime stage

- Step 5. Effect alk.phos is removed. The model contains the following effects:

sexn edema bili albumin copper ast protime stage
The final result is shown as below:

Analysis of Maximum Likelihood Estimates								
Parameter		DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio	Label
sexn	0	1	-0.44618	0.24370	3.3520	0.0671	0.640	sexn 0
edema		1	0.77035	0.29838	6.6655	0.0098	2.161	
bili		1	0.08657	0.01832	22.3404	<.0001	1.090	
albumin		1	-0.79440	0.23568	11.3617	0.0007	0.452	
copper		1	0.00303	0.0009021	11.2783	0.0008	1.003	
ast		1	0.00310	0.00153	4.0986	0.0429	1.003	
protime		1	0.20347	0.09459	4.6272	0.0315	1.226	
stage		1	0.47924	0.12643	14.3673	0.0002	1.615	

Summary of Backward Elimination					
Step	Effect Removed	DF	Number In	Wald Chi-Square	Pr > ChiSq
1	ascites	1	12	0.0024	0.9611
2	trt	1	11	0.2845	0.5938
3	spiders	1	10	0.2936	0.5879
4	hepato	1	9	0.9860	0.3207
5	alk.phos	1	8	2.4575	0.1170

Figure 7: Backward Result

SAS CODE

P2 a proc import out = PBC datafile = "/home/u63668108/PBC.csv" dbms = csv replace; getnames = yes; run;

data PBC; set PBC; if sex='f' then sexn=0; else sexn=1;

```
if chol="NA" then chol="";
if trig="NA" then trig="";
if platelet="NA" then platelet="";
chol_new = input(chol, 8.);
drop chol;
rename chol_new=chol;
trig_new = input(trig, 8.);
drop trig;
rename trig_new=trig;
platelet_new = input(platelet, 8.);
drop platelet;
rename platelet_new=platelet;
```

data covar; length Id $20; input Id $1-10 sexn edema bili albumin copper ast protime stage ; datalines; baseline 0 0 0 0 0 0 0 0 ; run;

proc phreg data=PBC plots(overlay)=survival; class sexn; model time*status(0) = sexn edema bili albumin copper ast protime stage; baseline covariates=covar out=Pred1 survival=*all*/rowid=Id; run;

proc print data=Pred1(FIRSTOBS = 120 OBS = 130); run;

P2 b proc import out = PBC datafile = "/home/u63668108/PBC.csv" dbms = csv replace; getnames = yes; run;

data PBC; set PBC; if sex='f' then sexn=0; else sexn=1;

```
if chol="NA" then chol="";
if trig="NA" then trig="";
if platelet="NA" then platelet="";
chol_new = input(chol, 8.);
drop chol;
rename chol_new=chol;
trig_new = input(trig, 8.);
drop trig;
rename trig_new=trig;
platelet_new = input(platelet, 8.);
drop platelet;
rename platelet_new=platelet;
```

proc means data=PBC median; var bili albumin copper ast protime; run;

data covar2; length Id$ 50; input sexn edema bili albumin copper ast protime stage Id $30-60; datalines; 1 1 1.4 3.53 73 114.7 10.6 4 sexn=1 edema=1 stage=4 ; run;

proc phreg data=PBC plots(overlay)=survival; class sexn; model time*status(0) = sexn edema bili albumin copper ast protime stage; baseline covariates=covar2 out=Pred2 survival=*all*/rowId=Id; run;

proc print data=Pred2(FIRSTOBS = 15 OBS = 25); run;

P3 proc phreg data=PBC; class trt sexn; model time*status(0) = trt sexn ascites hepato spiders edema bili albumin copper alk.phos ast protime stage/selection=backward sls=0.1; run;

SAS DOCUMENT

https://support.sas.com/documentation/cdl/en/statug/68162/HTML/default/viewer.htm#statug_glmselect_details07.htm

Github Link

https://github.com/zk2275/Survival_Analysis_HW6