answer for ex1.2

Jiaqi Wang

2025-09-15

Question 2：

1.This phase 3, open-label, global randomized trial compared ponatinib with imatinib. Both drugs were combined with reduced-intensity chemotherapy. It included adults aged 18 years or older with newly diagnosed Ph+ ALL and ECOG performance status ≤2. Patients were randomized in a 2:1 ratio to ponatinib or imatinib plus chemotherapy, then treated with single-agent therapy. The primary outcome was MRD-negative complete remission at the end of cycle 3. This need sustained remission with central laboratory–confirmed MRD negativity. The result showed that 34.4% in the ponatinib group achieved MRD-negative complete remission and 16.7% in the imatinib group, proved that ponatinib is better. 2. Answer the following questions about Table 1 in the article by Jabbour.

a. Which variables are continuous? Age

b. Which variables are unordered categorical variables?

   Sex (Male, Female); Central nervous system disease / extramedullary disease (Yes/No); BCR::ABL1 dominant isoform (p190, p210, atypical, undetermined/not tested); Cardiovascular comorbidities (Yes/No) Hypertension (Yes/No) Diabetes (Yes/No) Obesity (Yes/No) Dyslipidemia (Yes/No) History of smoking (Yes/No)

c. Which variables are ordered categorical variables?

   Eastern Cooperative Oncology Group performance status score (0, 1, 2) age group stratification (others, ≥60 years) Number of cardiovascular comorbidities (≥1, ≥2)

d. Why does the table not have an “overall” column that shows descriptive statistics for all 245 patients?

   Because its purpose is to compare baseline characteristics between the randomized treatment groups (ponatinib vs imatinib). Showing an overall column for all 245 patients would not help assess whether the groups are balanced after randomization.

e. Why are there no p-values in Table 1?

   For randomized trials, We are testing whether differences across treatments in certain characteristics are significant when we designed the study such that this should not be true nonsensical. p-values do not indicate whether or not any group imbalance will actually affect the analysis results.

Question 3：

1. In the CONSORT flowchart for this study, a total of 330 patients agreed to be screened; of these, 85 were excluded. The primary reasons included: failure to meet inclusion criteria (74 patients), withdrawal after providing informed consent (6 patients), adversely events occurring during screening (4 patients), and other reasons (1 patient). Finally, 245 patients were randomized to have 164 assigned to the ponatinib group and 81 to the imatinib group. During follow-up, some patients withdrew due to HSCT, undesirable events, lack of response, disease progression, or withdrawal of consent. Finally, 154 patients (ponatinib group) and 78 patients (imatinib group) were included in the molecular response analysis for the primary endpoint. The final analyzable population often are patients in better condition who can tolerate treatment. This may lead to trial results overestimating the drug’s efficacy and safety. The study population included adult patients aged 18 years older, with Ph+ ALL and an ECOG

performance status ≤2. However, in real-world clinical practice, patients with ECOG >2, co-infected with severe cardiovascular disease, or experiencing early adverse events are excluded from actual analyses. Therefore, when extrapolating study results to broader clinical populations, efficacy and safety may not be the same as observed it in the trial.

2. No statistical tests are required in Table 1 of the randomized trial. Randomized grouping would have resulted in some random variations, which are due to randomness and do not represent systematic bias. We are testing whether differences across treatments in certain characteristics are significant when we designed the study such that this should not be true – nonsensical. p-values do not indicate whether or not any group imbalance will actually affect the analysis results

3. In urtra-running, compare these features is to identify potential selection bias. This determines whether the final analysis population can represents the overall sample. In this context, using p-values is meaningless. In terms of confounding, P-value are not testing what is actuallyimportant in terms of choosing a confounder

Question 4：

```
dat <- read.csv("/Users/seventh/Desktop/All class-25 fall/702/ex1.2 consort/data.cs
v")
dat$Gender <- factor(dat$Gender, labels = c("Male", "Female"))
dat$ECOG <- factor(dat$ECOG, labels = c("0","1", "2"))
dat$CNS <- factor(dat$CNS, labels = c("No", "Yes"))
dat$BCR <- factor(dat$BCR, labels = c("p210", "p190"))
dat$CV1 <- factor(dat$CV1, labels = c("No", "Yes"))
dat$CV2 <- factor(dat$CV2, labels = c("No", "Yes"))
dat$Hypertension <- factor(dat$Hypertension, labels = c("No", "Yes") )
dat$Diabetes <- factor(dat$Diabetes, labels = c("No","Yes"))
dat$Obesity <- factor(dat$Obesity, labels = c("No", "Yes") )
dat$Dyslipidemia <- factor(dat$Dyslipidemia, labels = c("No", "Yes") )
dat$Smoking <- factor(dat$Smoking, labels = c("No", "Yes") )

dat$Treatment <- factor(dat$Treatment, labels = c("Ponatinib", "Imatinib"))

library(labelled)
var_label(dat$Age) <- "Age (years)"
var_label(dat$Gender) <- "Gender"
var_label(dat$ECOG) <- "ECOG Performance Status"
var_label(dat$CNS) <- "CNS/Extramedullary disease"
var_label(dat$BCR) <- "BCR::ABL1 Isoform"
var_label(dat$CV1) <- ">=1 Cardivascular Comorbidity"
var_label(dat$CV2) <- ">=2 Caradivascular Comorbidty"
var_label(dat$Hypertension) <- "Hypertension"
var_label(dat$Diabetes) <- "Diabetes"
var_label(dat$Obesity) <- "Obesity"
var_label(dat$Dyslipidemia) <- "Dyslipidemia"
var_label(dat$Smoking) <- "History of Smoking"

library(tableone)

vars <- c("Age", "Gender", "ECOG", "CNS", "BCR",
          "CV1", "CV2", "Hypertension", "Diabetes",
          "Obesity", "Dyslipidemia", "Smoking")
tab1 <- CreateTableOne(vars = vars, strata = "Treatment",
                       data = dat, test = TRUE, smd = TRUE)
library(kableExtra)
kableone(tab1, smd = TRUE, varLabels = TRUE,
         showAllLevels = TRUE, printToggle = FALSE) %>%
  kable_styling(full_width = FALSE)
```

	level	Ponatinib	Imatinib	p	test	SMD
n		81	164			
Age (years) (mean (SD))		55.05 (18.30)	53.75 (20.24)	0.626		0.067
Gender (%)	Male	38 (46.9)	97 (59.1)	0.094		0.247
	Female	43 (53.1)	67 (40.9)			

	level	Ponatinib	Imatinib	p	test	SMD
ECOG Performance Status (%)	0	31 (38.3)	42 (25.6)	0.115		0.280
	1	22 (27.2)	58 (35.4)			
	2	28 (34.6)	64 (39.0)			
CNS/Extramedullary disease (%)	No	39 (48.1)	83 (50.6)	0.821		0.049
	Yes	42 (51.9)	81 (49.4)			
BCR::ABL1 Isoform (%)	p210	44 (54.3)	81 (49.4)	0.555		0.099
	p190	37 (45.7)	83 (50.6)			
>=1 Cardivascular Comorbidity (%)	No	37 (45.7)	90 (54.9)	0.223		0.185
	Yes	44 (54.3)	74 (45.1)			
>=2 Caradivascular Comorbidty (%)	No	42 (51.9)	84 (51.2)	1.000		0.013
	Yes	39 (48.1)	80 (48.8)			
Hypertension (%)	No	42 (51.9)	86 (52.4)	1.000		0.012
	Yes	39 (48.1)	78 (47.6)			
Diabetes (%)	No	34 (42.0)	89 (54.3)	0.094		0.248
	Yes	47 (58.0)	75 (45.7)			
Obesity (%)	No	40 (49.4)	84 (51.2)	0.893		0.037
	Yes	41 (50.6)	80 (48.8)			
Dyslipidemia (%)	No	43 (53.1)	86 (52.4)	1.000		0.013
	Yes	38 (46.9)	78 (47.6)			
History of Smoking (%)	No	44 (54.3)	88 (53.7)	1.000		0.013
	Yes	37 (45.7)	76 (46.3)			

Question 5：

Answer the following questions about the Table 1 that you generated.

1. Would you reject the null hypothesis of no difference between the treatment arms for any of the baseline characteristics in Table 1 (using a 5% alpha level)?

   All p>0.5. I don’t reject null hypothesis

2. If you were to interpret the p-value as continuous, rather than using a threshold value like 5%, then are there any factors for which the p-value is suggesting the evidence leans in favor of the alternative hypothesis (i.e., that the distribution of a factor is actually different in patients assigned to Ponatinib

vs. Imatinib)? Discuss why this doesn't make any sense in this context. Refer to Altman's paper as a guide as you think about this.

If we consider the p-values as continuous indicators, we observe that Gender (p=0.071) and Diabetes (p=0.071) have relatively small values. This suggests a tendency for some differences. However, in randomized studies, such interpretations are meaningless because these differences from naturally from random assignment and should not be tested using p-values.

3. Are there any factors that are imbalanced between the groups when you look at the standardized mean differences (use 0.2 as a threshold)? List the factors and describe the imbalance that you see.

   Gender(SMD=0.246): The proportion of females was higher in the Imatinib group ECOG(SMD=0.207) : Patients in the Ponatinib group had worse overall status. Diabetes(SMD=0.247):The proportion of patients with diabetes was higher in the Imatinib group.

4. Why do you think the standardize mean differences are a more satisfactory approach to assessing imbalance than using p-values?

   SMD measures the magnitude of differences and better reflects clinically meaningful imbalances. In contrast, the p-value depends not only on the magnitude of differences but also on sample size. Therefore, SMD is more suitable as a measure of balance.

5. Suppose that the factors you identified as being imbalanced were also strongly associated with the primary outcome of the study. Using Altman's argument, what might you consider doing if you were the statistician for the study?

   Statisticians can add covariates to regression analysis models.

6. Suppose that the factors you identified as being imbalanced had no association with the primary outcome of the study. Again, referring to Altman's paper, is the imbalance a concern and what, if anything, would you consider doing about it?

   If these variables are unrelated to the primary outcome, their imbalance doe not lead a substantive issue. Such imbalance don't require extra handling. Only need to transparently presented in the results report.