Assignment 2 suggested solution

(The R script is for reference only. Please also note that the numerical values may be slightly different due to different definition of the propensity score strata)

a) Calculate the crude mortality rates of patients receiving early and late oseltamivir treatment respectively. [2 marks]

The crude death mortality of patients was 5.9% and 5.6% in late treatment and early treatment of oseltamivir respectively.

```
h1n1 <- read.csv(file="h1n1pdm.csv")
head(h1n1)
with(h1n1,table(et,death))
round(with(h1n1, prop.table(table(et, death),1)),3)
```

b) Describe the patient characteristics across the two treatment groups. [3 marks]

Table: Numbers and percentages of patients by time of treatment

Factor	Late (n=1227)	Early (n=373)
Male	614 (50.0%)	183 (49.1%)
Age group		
Children	336 (27.4%)	54 (14.5%)
Adult	257 (20.9%)	35 (9.4%)
Elderly	634 (51.7%)	284 (76.1%)
With mechanic ventilation	441 (35.9%)	179 (48.0%)
Asthma	186 (15.2%)	65 (17.4%)
COPD	203 (16.5%)	102 (27.3%)

Patients receiving the early treatment tends to be older, more likely for women, more severe with the use of mechanic ventilation, more likely to have asthma condition, and more likely to have COPD condition.

```
ftable(with(h1n1, table(male, et)))
round(prop.table(ftable(with(h1n1, table(male, et))),2),3)
ftable(with(h1n1, table(agegp, et)))
round(prop.table(ftable(with(h1n1, table(agegp, et))),2),3)
ftable(with(h1n1, table(mv, et)))
round(prop.table(ftable(with(h1n1, table(mv, et))),2),3)
ftable(with(h1n1, table(asthma, et)))
round(prop.table(ftable(with(h1n1, table(asthma, et))),2),3)
ftable(with(h1n1, table(copd, et)))
round(prop.table(ftable(with(h1n1, table(copd, et))),2),3)

ggplot(h1n1, aes(x=agegp, fill=factor(et))) + geom_bar() +
facet_grid(et ~ .)
```

```
ggplot(h1n1, aes(x=male, fill=factor(et))) + geom_bar() +
facet_grid(et ~ .)
ggplot(h1n1, aes(x=mv, fill=factor(et))) + geom_bar() + facet_grid(et
~ .)
ggplot(h1n1, aes(x=asthma, fill=factor(et))) + geom_bar() +
facet_grid(et ~ .)
ggplot(h1n1, aes(x=copd, fill=factor(et))) + geom_bar() +
facet_grid(et ~ .)
```

c) Estimate the propensity score of receiving early oseltamivir treatment by including all available predictors as main effects in a logistic regression model. Comment on the main characteristics of patients receiving early treatment. [4 mark]

Table: Estimated adjusted OR and corresponding 95% confidence intervals (CI) of receiving early treatment

Factors	aOR	95% CI	p-value
Male	0.96	0.76, 1.23	0.763
Age group			
Children	ref		
Adult	0.80	0.50, 1.26	0.342
Elderly	2.83	2.06, 3.95	<0.001
With mechanic ventilation	1.65	1.30, 2.11	<0.001
Asthma	1.33	0.96, 1.83	0.087
COPD	2.10	1.57, 2.78	<0.001

- Discuss both OR and p-value/95% CI for each factor.
- In general, patients receiving early treatment tend to be older, more severe with the use of mechanic ventilation and more likely to have COPD which were shown statistical significance at 5%.

```
h1n1$et<-as.factor(h1n1$et)
ps.model<-glm(et~factor(agegp)+male+mv+asthma+copd,
data=h1n1,family=binomial)
summary(ps.model)
round(exp(cbind(coef(ps.model),confint(ps.model))),2)
h1n1$ps<-predict(ps.model,type='response')
ggplot(h1n1,aes(x=ps,fill=et))+geom_histogram(binwidth=0.05)+facet_gri
d(et~.)
```

d) Propensity score stratification (by propensity score quintiles) will be used to analyze the effect of early treatment. Quote the propensity score quintiles. [2 marks]

0%	20%	40%	60%	80%	100%
0.077	0.121	0.227	0.234	0.335	0.584

We divided the propensity score (i.e. the predicted mean from the logistic model) into five equal groups by the following cutoff points: 0.121, 0.227, 0.234 and 0.345.

```
ps.boundary<-quantile(h1n1$ps,0:5/5)
ps.boundary
```

e) Assess the balance of the patient characteristics and sample size across treatment groups by propensity score strata. Is there any potential way to improve the balance? [4 marks]

The numbers of patients in each group were 307, 330, 205, 411 and 347 respectively. The percentages of patients with early treatment were 8.5%, 16.4%, 21.0%, 28.5% and 38.3% in each group.

```
h1n1$psq=cut(h1n1$ps,ps.boundary,right=F,include.lowest=T,label=1:5)
n.psq<-as.numeric(table(h1n1$psq))
n.psq
round(with(h1n1, prop.table(table(psq, et),1)),3)
```

The balance of patient characteristics across treatment groups are satisfactory for most variables, except there was some imbalance for asthma and COPD. The balance can be potentially improved by consider a more complicated propensity score model, such as including interaction terms.

Table: Tabulation between treatment and age group by groups of propensity scores

Group	Age group	Late treatment	Early treatment
1	Children	143 (50.9%)	14 (53.8%)
	Adult	138 (49.1%)	12 (46.2%)
	Elderly	0 (0.0%)	0 (0.0%)
2	Children	175 (63.4%)	34 (63.0%)
	Adult	101 (36.6%)	20 (37.%)
	Elderly	0 (35.2%)	0 (0.0%)
3	Children	2 (1.2%)	0 (0.0%)
	Adult	10 (6.2%)	2 (4.7%)
	Elderly	150 (92.6%)	41 (95.3%)
4	Children	16 (5.4%)	6 (5.1%)
	Adult	8 (2.7%)	1 (0.9%)
	Elderly	270 (91.8%)	110 (94.0%)
5	Children	0 (0.0%)	0 (0.0%)
	Adult	0 (0.0%)	0 (0.0%)
	Elderly	214 (100.0%)	133 (100.0%)

Table: Numbers and percentages of patients by treatment in each group of propensity scores

Factor	PS Strata	Late treatment	Early treatment
Male	1	133 (47.3%)	14 (53.8%)
	2	147 (53.3%)	69 (51.9%)
	3	150 (92.6%)	0 (95.3%)
	4	127 (43.2%)	58 (49.6%)
	5	57 (26.6%)	42 (31.6%)
Mechanic ventilation needed	1	0 (0.0%)	0 (0.0%)
	2	174 (63.0%)	29 (53.7%)
	3	10 (6.2%)	2 (4.7%)
	4	110 (37.4%)	57 (48.7%)
	5	147 (68.7%)	91 (68.4%)
Asthma	1	19 (6.8%)	4 (15.4%)
	2	75 (27.2%)	16 (29.6%)
	3	2 (1.2%)	0 (0%)
	4	55 (18.7%)	16 (13.7%)
	5	35 (16.4%)	29 (21.8%)
COPD	1	0 (0.0%)	0 (0.0%)
	2	71 (25.7%)	21 (38.9%)
	3	12 (7.4%)	2 (4.7%)
	4	24 (8.2%)	7 (6.0%)
	5	96 (44.9%)	72 (54.1%)

```
ftable(with(h1n1, table(et, agegp, psq)))[1:3,]
round(prop.table(ftable(with(h1n1, table(et, agegp, psq)))[1:3,],2),3)
ftable(with(h1n1, table(et, agegp, psq)))[4:6,]
round(prop.table(ftable(with(h1n1, table(et, agegp, psq)))[4:6,],2),3)

ftable(with(h1n1, table(et, male, psq)))[c(2,4),]
round(prop.table(ftable(with(h1n1, table(et, male,
psq)))[1:2,],2)[2,],3)
round(prop.table(ftable(with(h1n1, table(et, male,
psq)))[3:4,],2)[2,],3)

ftable(with(h1n1, table(et, mv, psq)))[c(2,4),]
round(prop.table(ftable(with(h1n1, table(et, mv, psq)))[1:2,],2)[2,],3)
round(prop.table(ftable(with(h1n1, table(et, mv, psq)))[3:4,],2)[2,],3)

ftable(with(h1n1, table(et, asthma, psq)))[c(2,4),]
round(prop.table(ftable(with(h1n1, table(et, asthma,
psq)))[1:2,],2)[2,],3)
round(prop.table(ftable(with(h1n1, table(et, asthma,
psq)))[3:4,],2)[2,],3)

ftable(with(h1n1, table(et, copd, psq)))[c(2,4),]
round(prop.table(ftable(with(h1n1, table(et, copd,
psq)))[1:2,],2)[2,],3)
round(prop.table(ftable(with(h1n1, table(et, copd,
psq)))[3:4,],2)[2,],3)
```

f) Calculate the stratum-specific and overall effect of early treatment (simple mean of the 5 strata). Please also provide the 95% CI for the overall treatment effect. [4 marks]

The stratum-specific effectiveness are -0.0249, -0.0024, -0.0261, -0.0116 and -0.0183. The overall effectiveness is -0.0156 (early vs. late), the weighted mean of the stratum-specific effectiveness. The 95% CI for the overall effectiveness is (-0.0402, 0.0091).

```
mean.tp<-aggregate(death~et+psq,data=h1n1,FUN=mean)
count.tp<-aggregate(death~et+psq,data=h1n1,FUN=length)
cbind(mean.tp,count.tp$death)
# strata specific mean
str.mean <- mean.tp$death[mean.tp$et==1] - mean.tp$death[mean.tp$et==0]
str.mean
# overall weighted mean
overall <- sum((mean.tp$death[mean.tp$et==1] -
mean.tp$death[mean.tp$et==0])*n.psq)/sum(n.psq)
overall
# 95% CI
var.tp <- aggregate(death~et+psq, data=h1n1, FUN=var)
var.over <- sum((var.tp$death/count.tp$death)*rep(n.psq,
each=2)^2)/sum(n.psq)^2
lowb <- overall - 1.96*sqrt(var.over)
uppb <- overall + 1.96*sqrt(var.over)
print(c(lowb, uppb))
```

g) Estimate the overall treatment effect using conditional logistic regression. [3 marks]

The odds ratio of the overall treatment effect (early vs. late) is 0.77 (0.46, 1.28).

```
library(survival)
clog.h1n1<-clogit(death~et+strata(psq),data=h1n1)
summary(clog.h1n1)
```

h) Based on the above results, comment on the effectiveness of early treatment of oseltamivir in reducing mortality and validity of the results in terms of control for confounders. [3 marks]

- The overall effectiveness of 1.6% indicates that the death rate in late treatment is higher than in early treatment.
- However, the corresponding 95% CI indicates that the difference is not statistically significant
- The patient characteristics have been satisfactorily balanced across treatment groups for most variables and should have controlled for most of the confounding effects from observed factors.
- Residual confounding may not be fully controlled by the propensity score method.