

R coding supplementary

1. A. Multivariate binomial regression

```
> t3<-read.csv("D:/Dr Ka Yu/sarcoma/manuscript/Database for building prediction model.csv")
> t3.complete<-t3[complete.cases(t3),]
```

```
t3.complete.binomial <-
glm(ThreeYearSurvival~Age+log_tum_dia+log_MF+Cervix+Parametrium+LymphNode+Distant_
_exceptLN+MF_VascularInvasion+MF_Circumscription, family=binomial, data=t3.complete)
> summary(t3.complete.binomial)
```

B. multivariate cox regression

```
t3.complete.survival<-
Surv(OS,Dead)~Age+log_tum_dia+log_MF+Cervix+Parametrium+LymphNode+Distant_except
LN+MF_VascularInvasion+MF_Circumscription
> t3.cox<-coxph(t3.complete.survival, data=t3.complete)
> summary(t3.cox)
```

2. Backward selection

```
Full<-
glm(ThreeYearSurvival~Age+log_tum_dia+log_MF+Cervix+Parametrium+LymphNode+Distant_
_exceptLN+MF_VascularInvasion+MF_Circumscription, family=binomial, data=t3.complete)
> Backward<-step(Full, direction="backward")
```

3. Final model

```
t3f.complete.binomial<-
glm(ThreeYearSurvival~Age+log_tum_dia+log_MF+Distant_exceptLN+MF_Circumscription,
family=binomial, data=t3.complete)
```

4. Model validation

```
install.packages("DAAG")
> require(DAAG)
cv.binary(t3f.complete.binomial,nfolds = 10)
```

5. 3 year survival prediction of the two examples

```
> newdata <- data.frame(Age=45,log_tum_dia=1,log_MF=1,Distant_exceptLN=0,
MF_Circumscription=1)
predict(t3f.complete.binomial, newdata, type="response")

newdata2 <- data.frame(Age=60,log_tum_dia=1,log_MF=1,Distant_exceptLN=1,
MF_Circumscription=0)
> predict(t3f.complete.binomial, newdata2, type="response")
```