利用R软件进行COX回归calibration curve的做法

作者：王晓晓，赵一鸣 来源：临床流行病学和循证医学

本文原地址：https://www.iikx.com/news/statistics/1746.html

本文仅供学习交流之用，版权归原作者所有，请勿用于商业用途！

曾记否，咱们一起讨论过calibration curve，也就是校准曲线，校准曲线其实是实际发生率和预测发生率的散点图。还记得logistic回归的calibration curve怎么做吗?链接在此。预测模型好不好，看看校正曲线就知道

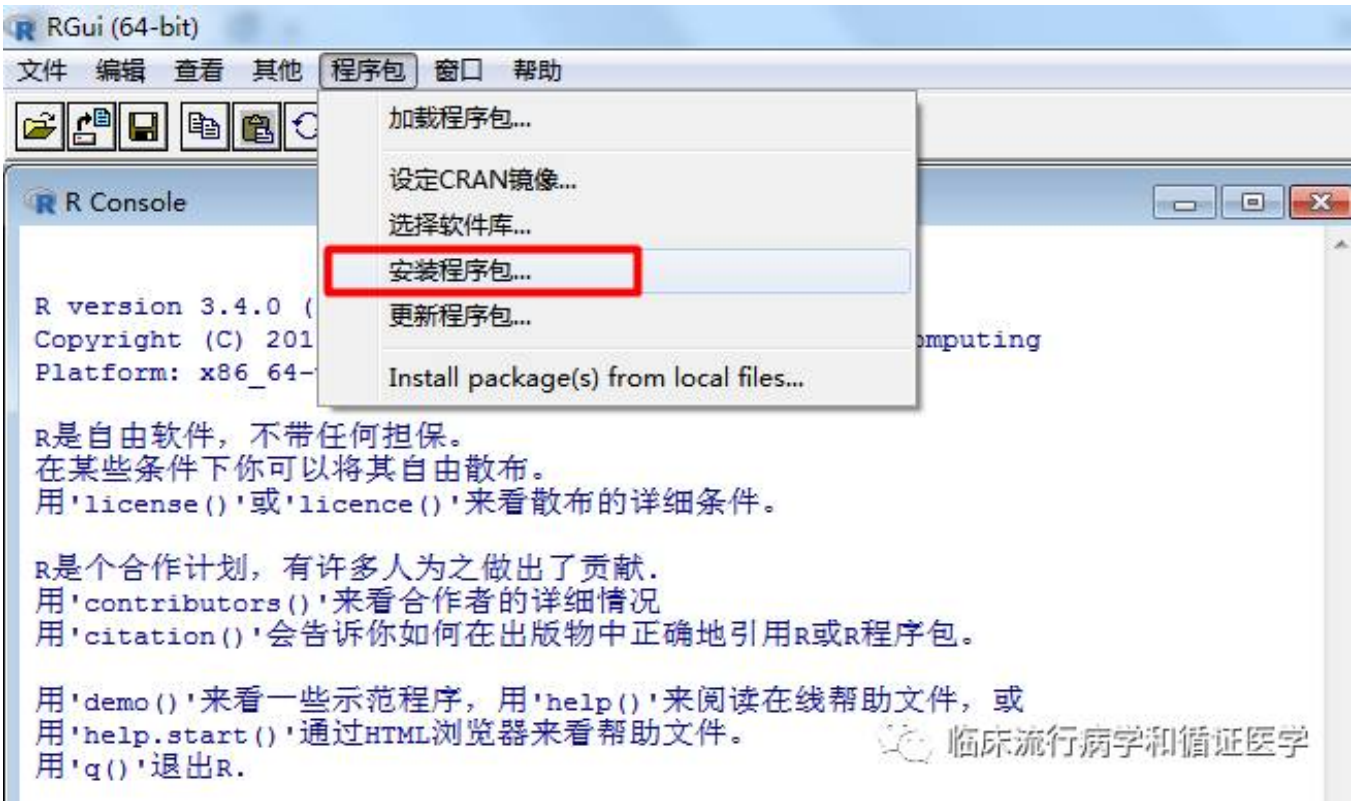
文章一出，有小伙伴发来贺电，同时催写COX回归calibration curve的做法，时隔半年，小编终于补上这一篇了。

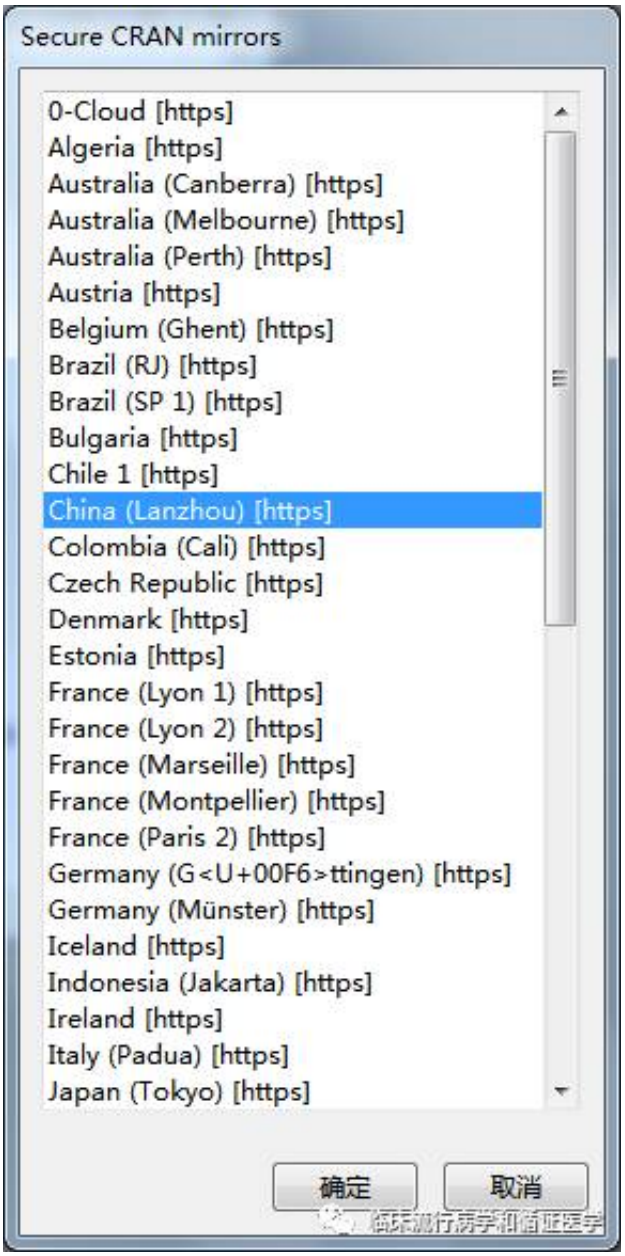
今天用的还是R软件，我们先来认识下所用的示例数据。数据命名为dtSurv，包含8个变量，我们所用到的主要是：生存时间—survTime、事件状态—status、以及自变量x1、x2和grp。

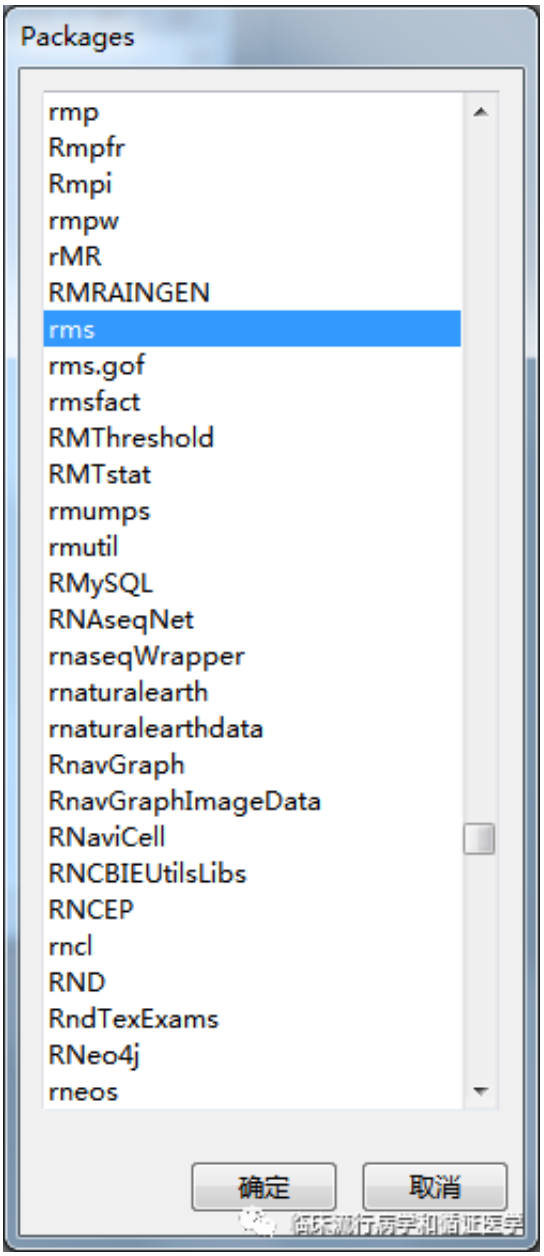
```
> head(dtSurv)
   id x1 x2 grp survTime censorTime obsTime status
1:  1  1  1   1       16         225      16      1
2:  2  1  1   0        6          16       6      1
3:  3  0  0   0        1          76       1      1
4:  4  0  1   0        7           3       3      0
5:  5  1  1   1      174          85      85      0
6:  6  0  0   0       15           2       2      0
>
```

临床流行病学和循证医学

首先，我们需要安装rms包，这个包也可用来做列线图。小编在使用过程中发现install.packages("rms")在安装程序包时会报错，这个时候，大家可直接在R界面手动安装。通过程序包—安装程序包—选择中国镜像—选择所需程序包。


RGui (64-bit)
文件 编辑 查看 其他 程序包 窗口 帮助
加载程序包...
设定CRAN镜像...
选择软件库...
安装程序包...
更新程序包...
Install package(s) from local files...
R Console
R version 3.4.0 (
Copyright (C) 201
Platform: x86_64-
mputing
R是自由软件，不带任何担保。
在某些条件下你可以将其自由散布。
用'license()'或'licence()'来看散布的详细条件。
R是个合作计划，有许多人为之做出了贡献.
用'contributors()'来看合作者的详细情况
用'citation()'会告诉你如何在出版物中正确地引用R或R程序包。
用'demo()'来看一些示范程序，用'help()'来阅读在线帮助文件，或
用'help.start()'通过HTML浏览器来看帮助文件。
用'q()'退出R.
临床流行病学和循证医学


Secure CRAN mirrors
0-Cloud [https]
Algeria [https]
Australia (Canberra) [https]
Australia (Melbourne) [https]
Australia (Perth) [https]
Austria [https]
Belgium (Ghent) [https]
Brazil (RJ) [https]
Brazil (SP 1) [https]
Bulgaria [https]
Chile 1 [https]
China (Lanzhou) [https]
Colombia (Cali) [https]
Czech Republic [https]
Denmark [https]
Estonia [https]
France (Lyon 1) [https]
France (Lyon 2) [https]
France (Marseille) [https]
France (Montpellier) [https]
France (Paris 2) [https]
Germany (G<U+00F6>ttingen) [https]
Germany (Münster) [https]
Iceland [https]
Indonesia (Jakarta) [https]
Ireland [https]
Italy (Padua) [https]
Japan (Tokyo) [https]
确定
取消



在程序包安装之后：

```
> library(rms)
> s<-Surv(dtSurv$survTime,dtSurv$status,type="right")
> f <- cph(s~grp+x1+x2, x=TRUE, y=TRUE,surv = TRUE,time.inc=5,data=dtS
urv)
> cal<-calibrate(f,u=5,cmethod='KM',m=100)
Using Cox survival estimates at   5 Days
There were 50 or more warnings (use warnings() to see the first 50)
>
```

临床流行病学和循证医学

library(rms)##加载程序包

s<-Surv(dtSurv$survTime,dtSurv$status,type="right")##申明为生存分析的数据，涂黄部分为生存时间

，涂绿部分为事件发生与否。其中dtSurv是数据的名称，通过$美元符号连接相应的变量名。大家需要更改涂颜色的部分。

f <- cph(s~grp+x1+x2, x=TRUE, y=TRUE,surv = TRUE,time.inc=5,data=dtSurv)##拟合COX回归模型，大家需要更改涂颜色的部分，grp+x1+x2为相应的自变量，time.inc表示计算多长时间的生存率，data需要等于数据的名称。

cal<-calibrate(f,u=5,cmethod='KM',m=100)##这里的u应和time.inc保持一致。关于m，一般为50、100，和总样本量有关，calibration
curve中，数据点的数目=总样本量n/m。dtSurv共计500人，m设为100，calibration
curve应有数据点5个，待会可以验证下。

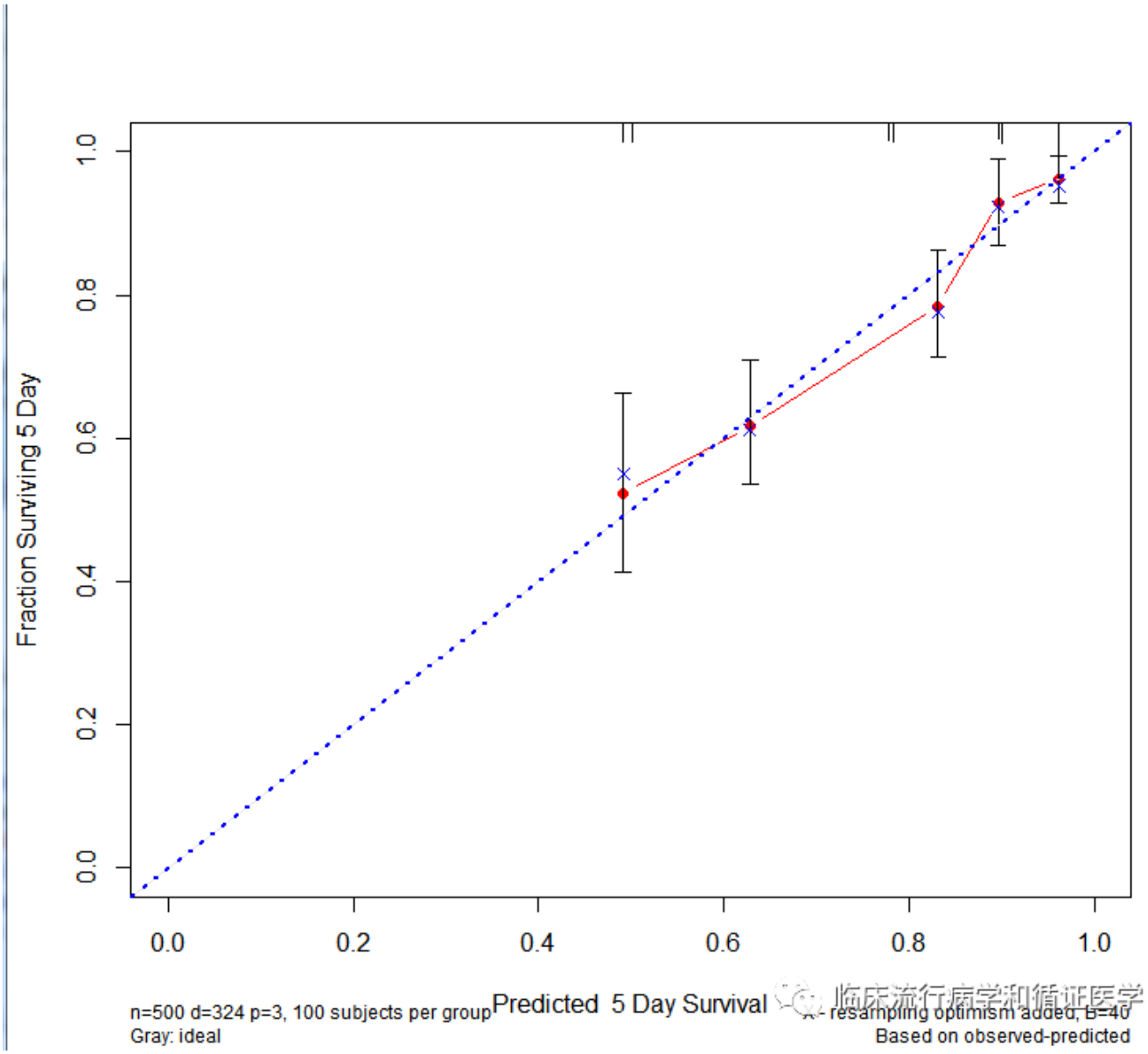
```
> plot(cal,xlim = c(0,1),ylim= c(0,1),
+      errbar.col=c(rgb(0,0,0,maxColorValue=255)),col=c(rgb(255,0,0,ma
xColorValue= 255)))
>
>
> abline(0,1,lty =3,lwd=2,col=c(rgb(0,0,255,maxColorValue= 255)))
>
```

临床流行病学和循证医学

plot(cal,xlim = c(0,1),ylim= c(0,1),

errbar.col=c(rgb(0,0,0,maxColorValue=255)),col=c(rgb(255,0,0,maxColorValue=255)))##绘图的参数，xlim和ylim限定x和y轴的区间。errbar.col定义误差线的颜色，col定义校准曲线的颜色，颜色可参考rgb调色进行调整。

abline(0,1,lty=3,lwd=2,col=c(rgb(0,0,255,maxColorValue=
255)))##添加y=x并设置线型、粗细、颜色。最后绘制的calibration curve如下。




```
library(rms)

s<-Surv(dtSurv$survTime,dtSurv$status,type="right")

f <- cph(s~grp+x1+x2, x=TRUE, y=TRUE,surv = TRUE,time.inc=5,data=dtSurv)

cal<-calibrate(f,u=5,cmethod='KM',m=100)

plot(cal,xlim = c(0,1),ylim= c(0,1),

errbar.col=c(rgb(0,0,0,maxColorValue=255)),col=c(rgb(255,0,0,maxColorValue= 255)))

abline(0,1,lty =3,lwd=2,col=c(rgb(0,0,255,maxColorValue= 255)))
```

更多 统计方法 请访问 https://www.iikx.com/news/statistics/

本文版权归原作者所有，请勿用于商业用途，爱科学iikx.com转发