# 合理进行多重Logistic回归分析——结合多水平模型分析

胡纯严[1]，胡良平[1,2*]

（1. 军事科学院研究生院，北京 100850；
2. 世界中医药学会联合会临床科研统计学专业委员会，北京 100029
*通信作者：胡良平，E-mail：lphu927@163.com）

【摘要】 本文目的是介绍如何结合多水平模型分析，合理地进行多重Logistic回归分析的方法。第一，介绍了与多水平模型分析有关的4个基本概念。第二，介绍了构建多水平模型的3个步骤。第三，通过一个多中心药物临床试验的实例，介绍了如何用SAS软件进行分析的全过程，其内容如下：①检验各中心优势比之间是否具有齐性；②对试验中心产生哑变量后构建多重Logistic回归模型；③将试验中心视为分层变量构建多重Logistic回归模型；④构建随机截距多水平多重Logistic回归模型；⑤构建随机截距和随机斜率多水平多重Logistic回归模型。得到的结论是，当具有二值结果变量的各层级资料间存在差异时，最合适的做法是构建多水平多重Logistic回归模型。

【关键词】 层级结构；分层变量；多水平模型；随机截距；随机斜率
中图分类号：R195.1 文献标识码：A doi：10.11886/scjsws20221113003

## Reasonably conduct the multiple Logistic regression analysis combined with the multilevel model analysis

*Hu Chunyan*[1], *Hu Liangping*[1,2*]

（1. *Graduate School, Academy of Military Sciences PLA China, Beijing* 100850, *China*;
2. *Specialty Committee of Clinical Scientific Research Statistics of World Federation of Chinese Medicine Societies, Beijing* 100029, *China*
*Corresponding author: *Hu Liangping*, *E-mail*: *lphu*927@163.*com*）

【Abstract】 The purpose of this paper was to introduce how to reasonably analyze the multiple Logistic regression models in combination with the multilevel model analysis. Firstly, four basic concepts related to the multilevel model analysis were introduced. Secondly, three steps for building a multilevel model were given. Thirdly, through an example of a multicenter drug clinical trial, the whole process of how to use SAS software for the analysis was presented. The contests were as follows: ① testing whether the odds ratios of each center were homogenous; ② building the multiple Logistic regression model after generating dummy variables for the trial center; ③ constructing a multiple Logistic regression model with the trial center as a stratified variable; ④ building a random intercept multilevel multiple Logistic regression model; ⑤ constructing a random intercept and random slope multilevel multiple Logistic regression model. The conclusion was that when there were differences among the data at different hierarchies with binary outcome variables, the most appropriate approach was to build a multilevel multiple Logistic regression model.

【Keywords】 Hierarchical structure; Hierarchical variable; Multilevel model; Random intercept; Random slope

统计资料可分为无层级结构和有层级结构两种类型。通常情况下，研究者收集的统计资料为无层级结构的资料。此时，可运用常规统计分析方法处理资料。然而，在调查研究中，受试对象常归属于不同层级，例如某省某市某县某区。又例如，学生常归属于某市某中学某年级某班级。当资料中的受试对象归属于多层级结构时，一般来说，不适合选用处理无层级结构资料的统计分析方法。本文将结合一个具有2个层级结构的常见临床试验资料，介绍二水平多重Logistic回归模型的构建方法和SAS实现方法。

## 1 基本概念

### 1.1 分层变量

在统计资料中，有一类变量被称为“分层变量”。例如，在人口普查资料中，会涉及省、市、县、区、街道或村委会，它们都可以被称为“分层变量”。从全国角度考量，可以把“省”作为最高级别的分层变量，把“市”作为第二级别的分层变量，以此类推，把“街道或村委会”作为最低级别的分层变量。

### 1.2 层级结构

在研究人的身高与体重关系的资料中，若受试

对象的来源涉及省、市、县3个行政级别，在统计学上就称此资料为具有3个层级结构的资料。也就是说，省下面嵌套着市、市下面又嵌套着县。同理，在一个多中心临床试验研究中，常把“试验中心”和“受试对象”视为2个不同层级上的变量，因为“试验中心”下面嵌套着“受试对象”。

### 1.3 k水平单位

在一个涉及省、市、县3个行政级别的关于居民健康情况的调查研究中，受调查者为1水平单位，县为2水平单位，市为3水平单位，省为4水平单位；另一种等价的表述为：受调查者为水平1单位，县为水平2单位，市为水平3单位，省为水平4单位。

### 1.4 个体水平与组水平

所谓个体水平，实际上就是指多层级资料中的最底层，通常为“个体”。若试验中需要对每个个体进行多次重复观测，则时间点就成为最底层了，此时，时间点就叫做“个体水平”，而每个个体就被称为“组水平”。在没有重复观测的试验研究资料中，“个体”就叫做“个体水平”，而它们的上一级就叫做“组水平”。在多层级结构中，必须给“组水平”附上相应的标记，以示区别。例如，在一个涉及省、市、县3个行政级别的关于居民健康情况的调查研究中，就有3种“组水平”，即“省级组水平”“市级组水平”和“县级组水平”。

## 2 多水平模型的构建步骤

多水平模型是基于层级结构数据形成的一种统计模型[1]，对于具有层级结构的二值结果变量的统计资料，需要采用多水平多重Logistic回归模型分析，它是一般Logistic回归模型的延伸，通过在模型中纳入随机效应项来处理层级结构数据的组内相关[2]。其模型表达见式(1)。

$$\ln\left(\frac{\mathrm{P}}{1-\mathrm{P}}\right)=\mathrm{X}\beta+\mathrm{Z}U \tag{1}$$

式(1)中，X是固定效应的解释变量设计矩阵，Z是随机效应的解释变量设计矩阵，$\beta$是水平1固定回归系数向量，$U$是随机回归系数向量，服从均值为0、协方差为矩阵G的正态分布。

下面介绍2水平多重Logistic回归模型的建模过程及分析步骤。假定在一个临床试验中，涉及试验药种类(drug)和临床试验中心(center)两个主要的原因变量，还涉及一个代表疗效的结果变量(y)。假设：y=0表示治疗成功，y=1表示治疗失败，以$p_{ij}$表示个体y=0发生的概率。建模过程如下：

第一步，建立空模型，计算组内相关系数(ICC)的值。空模型中仅有一个随机截距而不包含任何自变量，其模型见式(2)。

$$\ln\left(\frac{p_{ij}}{1-p_{ij}}\right)=\beta_{0j},$$
$$j=1,2,\cdots,C;\ i=1,2,\cdots,m_j \tag{2}$$

式(2)中，$j$代表2水平单位的编号，$C$代表2水平单位的数目；$i$代表第$j$个2水平单位组中个体的编号，$m_j$代表第$j$个2水平单位组中个体的数目，$\beta_{0j}$代表第$j$个2水平单位组中的截距，其表达式见式(3)。

$$\beta_{0j}=\beta_0+\mu_{0j} \tag{3}$$

式(2)和式(3)两个模型可合并成式(4)。

$$\ln\left(\frac{p_{ij}}{1-p_{ij}}\right)=\beta_0+\mu_{0j} \tag{4}$$

式(4)中，$\beta_0$为y=0的总平均logit值，$\mu_{0j}$为组水平(本资料为中心)的平均logit变异值，它表示第$j$个组的平均logit值与总平均logit值之间的差异，且$\mu_{0j}\sim N(0,\sigma^2_{\mu_0})$。

多水平多重Logistic回归模型的组间变异也可用ICC进行评估，因Logistic回归模型的残差方差为$\pi^2/3$，所以ICC的计算公式可由式(5)表示。

$$\mathrm{ICC}=\frac{\sigma^2_{\mu_0}}{\sigma^2_{\mu_0}+\pi^2/3}=\frac{\sigma^2_{\mu_0}}{\sigma^2_{\mu_0}+3.289868134} \tag{5}$$

第二步，建立包含自变量的随机截距模型，即在随机截距的基础上再考察变量drug的固定效应。构建的模型见式(6)。

$$\ln\left(\frac{p_{ij}}{1-p_{ij}}\right)=\beta_{0j}+\beta_1 drug_{ij} \tag{6}$$

式(6)中的$\beta_{0j}$由式(7)给出。

$$\beta_{0j}=\beta_0+\mu_{0j} \tag{7}$$

式(6)和式(7)两个模型可合并成式(8)的形式。

$$\ln\left(\frac{p_{ij}}{1-p_{ij}}\right)=\left(\beta_0+\beta_1 drug_{ij}\right)+\mu_{0j} \tag{8}$$

式(8)中，$\beta_0+\beta_1 drug_{ij}$为固定效应，$\mu_{0j}$为随机效应，且$\mu_{0j}\sim N(0,\sigma^2_{\mu_0})$。

第三步，建立包含自变量的随机截距-随机斜率模型，即截距项和自变量drug的回归系数中均含有随机效应项。构建的模型见式(9)。

$$\ln\left(\frac{p_{ij}}{1-p_{ij}}\right)=\beta_{0j}+\beta_{1j} drug_{ij} \tag{9}$$

式(9)中等号右边的$\beta_{0j}$和$\beta_{1j}$分别由式(10)和式

(11)给出。

$$\beta_{0j}=\beta_0+\mu_{0j} \quad (10)$$

$$\beta_{1j}=\beta_1+\mu_{1j} \quad (11)$$

式(9)、式(10)、式(11)的3个模型可合并成式(12)的形式。

$$\ln\left(\frac{p_{ij}}{1-p_{ij}}\right)=\left(\beta_0+\beta_1 drug_{ij}\right)+\left(\mu_{0j}+\mu_{1j}drug_{ij}\right) \quad (12)$$

式(12)中,$\beta_0+\beta_1 drug_{ij}$为固定效应,$\mu_{0j}+\mu_{1j}drug_{ij}$为随机效应,且$\mu_{0j}\sim N(0,\sigma^2_{\mu_0})$,$\mu_{1j}\sim N(0,\sigma^2_{\mu_1})$,$\mu_{0j}$与$\mu_{1j}$之间的协方差可能有统计学意义;若无统计学意义,则将它们之间的协方差设定为0。

## 3 实例与SAS实现

### 3.1 问题与数据结构

【例1】某临床研究中,研究者选择16所医院同时开展临床试验,每所医院均选取受试者120人,在医院内随机等分为两组,分别接受试验药物和对照药物的治疗。治疗结果见表1,试比较两种药物的疗效[3]。

**表1 多中心临床试验数据**

Table 1 Data from multicenter clinical trials

| 医院编号 | 药物种类 | 例数(人) 疗效:成功 | 失败 | 医院编号 | 药物种类 | 例数(人) 疗效:成功 | 失败 |
|---|---|---|---|---|---|---|---|
| 1 | 试验药 | 42 | 18 | 9 | 试验药 | 47 | 13 |
| | 对照药 | 28 | 32 | | 对照药 | 38 | 22 |
| 2 | 试验药 | 37 | 23 | 10 | 试验药 | 29 | 31 |
| | 对照药 | 29 | 31 | | 对照药 | 25 | 35 |
| 3 | 试验药 | 51 | 9 | 11 | 试验药 | 36 | 24 |
| | 对照药 | 22 | 38 | | 对照药 | 25 | 35 |
| 4 | 试验药 | 46 | 14 | 12 | 试验药 | 41 | 19 |
| | 对照药 | 37 | 23 | | 对照药 | 18 | 42 |
| 5 | 试验药 | 39 | 21 | 13 | 试验药 | 39 | 21 |
| | 对照药 | 38 | 22 | | 对照药 | 17 | 43 |
| 6 | 试验药 | 29 | 31 | 14 | 试验药 | 32 | 28 |
| | 对照药 | 25 | 35 | | 对照药 | 26 | 34 |
| 7 | 试验药 | 40 | 20 | 15 | 试验药 | 35 | 25 |
| | 对照药 | 29 | 31 | | 对照药 | 15 | 45 |
| 8 | 试验药 | 28 | 32 | 16 | 试验药 | 30 | 30 |
| | 对照药 | 12 | 48 | | 对照药 | 30 | 30 |

### 3.2 检验各中心优势比之间是否具有齐性

为了检验各中心优势比之间是否具有齐性,可利用Breslow-Day检验[4]。设所需要的SAS程序如下:

```
data a1;
do center=1 to 16;
do drug=0 to 1;
do y=0 to 1;
input f @@;
do i=1 to f;
output;
end; end; end; end;
cards;
42 18 28 32 37 23 29 31
51  9 22 38 46 14 37 23
39 21 38 22 29 31 25 35
40 20 29 31 28 32 12 48
47 13 38 22 29 31 25 35
36 24 25 35 41 19 18 42
39 21 17 43 32 28 26 34
35 25 15 45 30 30 30 30
;
run;
/*考查各中心优势比之间是否具有齐性*/
proc freq data=a1;
tables center*drug*y/cmh;
run;
```

【SAS主要输出结果及解释】优势比齐性的Breslow-Day检验结果:$\chi^2$=37.994,$df$=15,$P$=0.0009。此结果表明:16个中心的优势比不满足齐性要求,说明不应将16个中心的资料简单合并后构建多重Logistic回归模型。

### 3.3 采用4种方法构建多重Logistic回归模型

#### 3.3.1 对试验中心产生哑变量后构建多重Logistic回归模型

设所需要的SAS程序如下:

```
proc logistic data=a1;
class center;
model y=center drug;
run;
```

【SAS主要输出结果及解释】对二重Logistic回归模型中两个参数进行的"3型效应分析"的结果见表2。

**表2 模型中两个参数的"3型效应分析"的结果**

Table 2 Results of "Type 3 effect analysis" of two parameters in the model

| 效应 | 自由度 | Wald $\chi^2$ | Pr>$\chi^2$ |
|---|---|---|---|
| center | 15 | 74.886 | <0.010 |
| drug | 1 | 74.844 | <0.010 |

由以上输出结果可知，16个中心疗效之间的差异有统计学意义，两种药物疗效之间的差异也有统计学意义。因模型中参数估计部分的输出结果很多，以下仅简单呈现与第16中心比较差异有统计学意义的结果，见表3。

**表3　与第16中心比较差异有统计学意义的结果**

Table 3　Statistically significant results compared with the 16th center

| 参　数 | 自由度 | 估　计 | 标准误差 | Wald $\chi^2$ | Pr>$\chi^2$ |
|---|---|---|---|---|---|
| Intercept | 1 | 0.539 | 0.068 | 62.174 | <0.010 |
| center4 | 1 | 0.716 | 0.194 | 13.566 | <0.010 |
| center5 | 1 | 0.482 | 0.188 | 6.582 | 0.010 |
| center8 | 1 | −0.847 | 0.191 | 19.723 | <0.010 |
| center9 | 1 | 0.798 | 0.197 | 16.381 | <0.010 |
| center15 | 1 | −0.476 | 0.183 | 6.759 | 0.009 |
| drug | 1 | −0.826 | 0.096 | 74.844 | <0.010 |

由以上输出结果可知，仅5个中心与第16中心的疗效之间的差异有统计学意义；试验药与对照药疗效之间的差异有统计学意义。由于程序中设定：拟合“y=0（治疗成功）”的概率模型，并且，drug=0代表试验药，drug=1代表对照药，drug的回归系数为−0.826，即对照药相对于试验药的优势比OR=exp（−0.826）=0.438。也就是说，试验药相对于对照药的优势比为1/OR=1/0.438=2.283倍。

### 3.3.2　将试验中心视为分层变量构建多重Logistic回归模型

相当于在每个试验中心内部进行配对设计，再进行条件Logistic回归分析。设所需要的SAS程序如下：

```
proc logistic data=a1;
class center;
model y=drug;
strata center/info check=all;
run;
```

【SAS主要输出结果及解释】基于center分层的一重Logistic回归模型的分析结果见表4。

**表4　条件最大似然估计的分析结果**

Table 4　Analysis results of conditional maximum likelihood estimation

| 参　数 | 自由度 | 估　计 | 标准误差 | Wald $\chi^2$ | Pr>$\chi^2$ |
|---|---|---|---|---|---|
| drug | 1 | −0.819 | 0.095 | 74.236 | <0.010 |

由以上输出结果可知，参数drug的回归系数为−0.819，即对照药相对于试验药的优势比OR=exp（−0.819）=0.441。也就是说，试验药相对于对照药的优势比为1/OR=1/0.441=2.268倍。

### 3.3.3　构建随机截距多水平多重Logistic回归模型

从16个“不同中心”抽取受试对象，同一个中心的个体之间的差异被称为1水平上的差异；而不同中心的个体之间的差异被称为2水平上的差异，需要构建随机截距多水平多重Logistic回归模型。设所需要的SAS程序如下：

```
/*拟合随机截距模型*/
proc glimmix data=a1 method=rspl;
class center;
model y(event='0')=drug/s dist=binary link=logit ddfm=bw;
random int /sub=center;
run;
```

【SAS主要输出结果及解释】与center对应的截距项的计算结果：截距项中随机部分V_u0=0.154（即随机部分的方差）。模型中固定效应的计算结果见表5。

**表5　模型中固定效应的计算结果**

Table 5　Calculation results of fixed effects in the model

| 效　应 | 估　计 | 标准误差 | 自由度 | $t$ | Pr>\|t\| |
|---|---|---|---|---|---|
| Intercept | 0.530 | 0.119 | 15 | 4.450 | 0.001 |
| drug | −0.814 | 0.095 | 1 903 | −8.600 | <0.010 |

由以上输出结果可知，$b_0$=0.530，$b_1$=−0.814，它们与0之间的差异均有统计学意义。drug的回归系数为−0.814，即对照药相对于试验药的优势比OR=exp（−0.814）=0.443。也就是说，试验药相对于对照药的优势比为1/OR=1/0.443=2.257倍。

为了检验截距项中随机部分V_u0=0.154与0之间的差异是否有统计学意义，设所需要的SAS程序如下：

```
/*将以上模型估计的3个参数值代入下面的程序中*/
proc nlmixed data=a1;
parms b0=0.530 b1=−0.814 V_u0=0.154;
z=b0+b1*drug+u0j;
if (y=0) then p=exp(z)/(1+exp(z));
else p=1−(exp(z)/(1+exp(z)));
ll=log(p);
model y~general(ll);
random u0j~normal(0,V_u0) sub=center;
run;
```

【SAS主要输出结果及解释】模型中各参数的估计结果见表6。

表 6 模型中各参数的估计结果

Table 6 Estimation results of parameters in the model

| 参数 | 估计 | 标准误差 | 自由度 | $t$ | Pr>\|t\| | 95% 置信限 | 梯度 |
|---|---|---|---|---|---|---|---|
| b0 | 0.533 | 0.117 | 15 | 4.580 | <0.010 | 0.285~0.782 | -9.5E-6 |
| b1 | -0.819 | 0.095 | 15 | -8.620 | <0.010 | -1.022~-0.617 | -1.44E-6 |
| V_u0 | 0.144 | 0.064 | 15 | 2.230 | 0.041 | 0.007~0.280 | 0.000 420 |

由以上输出结果可知，对 3 个参数进行了重新估计，并且，对它们都进行了假设检验，$P$ 值均小于 0.05。$b_0$=0.533，$b_1$=-0.819，V_u0=0.144。

### 3.3.4 构建随机截距-随机斜率多水平多重 Logistic 回归模型

还可以尝试构建随机截距-随机斜率多水平多重 Logistic 回归模型。设所需要的 SAS 程序如下：

```
/*拟合随机截距和随机斜率模型*/
proc glimmix data=a1 method=rspl;
class center;
model y(event='0')=drug/s dist=binary link=logit ddfm=bw;
random int drug /sub=center type=chol;
run;
```

【SAS 主要输出结果及解释】模型中参数的协方差参数估计结果见表 7。

表 7 模型中参数的协方差参数估计结果

Table 7 Covariance parameter estimation results of parameters in the model

| 协方差参数 | 对象 | 估计 | 标准误差 |
|---|---|---|---|
| CHOL(1,1) | center | 0.432 | 0.112 |
| CHOL(2,1) | center | -0.199 | 0.182 |
| CHOL(2,2) | center | 0.427 | 0.119 |

由以上输出结果可知，截距中的随机部分 V_u0=0.432，斜率中的随机部分 V_u1=0.427，它们之间的协方差 Cov_u01=-0.199。模型中参数的固定效应的分析结果见表 8。

表 8 模型中参数的固定效应的分析结果

Table 8 Analysis results of fixed effects of parameters in the model

| 效应 | 估计 | 标准误差 | 自由度 | $t$ | Pr>\|t\| |
|---|---|---|---|---|---|
| Intercept | 0.532 | 0.128 | 15 | 4.170 | 0.001 |
| drug | -0.821 | 0.151 | 1 903 | -5.420 | <0.010 |

由以上输出结果可知，截距和斜率的固定部分的估计值分别为 $b_0$=0.532，$b_1$=-0.821，它们与 0 之间的差异均具有统计学意义。drug 的回归系数为 -0.821，即对照药相对于试验药的优势比 OR=exp(-0.821)=0.440。也就是说，试验药相对于对照药的优势比为 1/OR=1/0.440=2.273 倍。

为了检验 3 个随机部分的估计值与 0 之间的差异是否有统计学意义，设所需要的 SAS 程序如下：

```
/*将以上模型估计的5个参数值代入下面的程序中*/
proc nlmixed data=a1;
parms b0=0.5322 b1=-0.8208 V_u0=0.4323
Cov_u01=-0.1994 V_u1=0.4270;
z=b0+b1*drug+u0j+u1j;
if (y=0) then p=exp(z)/(1+exp(z));
else p=1-(exp(z)/(1+exp(z)));
ll=log(p);
model y~general(ll);
random u0j u1j~normal([0,0],[V_u0,Cov_u01,V_u1]) sub=center;
run;
```

【SAS 主要输出结果及解释】模型中参数的估计结果见表 9。

表 9 模型中参数的估计结果

Table 9 Analysis results of fixed effects of parameters in the model

| 参数 | 估计 | 标准误差 | 自由度 | $t$ | Pr>\|t\| | 95% 置信限 | 梯度 |
|---|---|---|---|---|---|---|---|
| b0 | 0.533 | 0.117 | 14 | 4.580 | <0.010 | 0.283~0.783 | -0.000 050 |
| b1 | -0.819 | 0.095 | 14 | -8.620 | <0.010 | -1.023~-0.615 | 0.000 014 |
| V_u0 | 0.380 | - | 14 | - | - | - | 0.000 326 |
| Cov_u01 | -0.305 | - | 14 | - | - | - | 0.000 652 |
| V_u1 | 0.374 | - | 14 | - | - | - | 0.000 326 |

由以上输出结果可知，后三行为随机部分，它们的标准误差无法计算出来，因此，无法检验它们与 0 之间的差异是否有统计学意义。从 SAS 日志可知，无法计算的原因是最终的海森矩阵不是正定矩阵，因此，估计的协方差矩阵不是满秩的[4]。

## 3.4 结论

比较“第 3.3.1 节”到“第 3.3.4 节”的计算结果，

可得出以下结论：第一，采用 4 种方法构建多重 Logistic 回归模型，所获得的截距和斜率的固定部分的估计值比较接近；第二，采用多水平 Logistic 回归模型分析，可以把不同中心资料优势比之间的不齐性以随机效应定量地呈现出来，即各中心试验药与对照药的成功率之间存在明显的跨中心效应。也就是说，采用无层级结构的多重 Logistic 回归模型[5-7]分析层级之间存在明显差异的统计资料是不合适的。

## 4 讨论与小结

### 4.1 讨论

由于本例资料中有一个多值名义变量“center（试验中心）”，使统计分析者有多种可能的分析策略。除了本文已经采用的 4 种建模方法之外，还有两种误用的做法：一是忽视变量“center”的存在；二是将变量“center”视为有序变量或定量变量（赋值为 1，2，…，16）。

### 4.2 小结

本文介绍了与多水平模型分析有关的 4 个基本概念；介绍了构建二水平 Logistic 回归模型的三个步骤；针对一个多中心药物临床试验的实例，介绍了用 SAS 实现分析的全过程，包括如何检验各试验中心优势比的齐性，如何将试验中心变换为哑变量后构建多重 Logistic 回归模型，如何将试验中心视为分层变量构建多重 Logistic 回归模型，如何构建随机截距多重 Logistic 回归模型，以及如何构建随机截距和随机斜率多重 Logistic 回归模型。对 4 种可行的建模方法所获得的估计结果进行了比较，作出了结论。最后，在讨论中还指出了两种可能被误用的做法。

## 参考文献


[1] 石磊. 多水平模型及其统计诊断[M]. 北京：科学出版社，2008：27-50.
Shi L. Multilevel model and its statistical diagnosis[M]. Beijing: Science Press, 2008: 27-50.

[2] 王济川，谢海义，姜宝法. 多层统计分析模型：方法与应用[M]. 北京：高等教育出版社，2008：15-190.
Wang JC, Xie HY, Jiang BF. Multilevel models: methods and applications [M]. Beijing: Higher Education Press, 2008: 15-190.

[3] 胡良平. 面向问题的统计学：(2)多因素设计与线性模型分析[M]. 北京：人民卫生出版社，2012：482-492.
Hu LP. Problem oriented statistics: (2) multifactor design and linear model analysis[M]. Beijing: People's Medical Publishing House, 2012: 482-492.

[4] SAS Institute Inc. SAS/STAT®15.1 user's guide[M]. Cary, NC: SAS Institute Inc, 2018: 2997-3216, 5749-6006.

[5] 李长平，胡良平. 非配对设计二值资料一水平多重 Logistic 回归分析[J]. 四川精神卫生，2019，32(4)：297-303.
Li CP, Hu LP. One-level multiple Logistic regression analysis with the dichotomous choice data collected from the unpaired design[J]. Sichuan Mental Health, 2019, 32(4): 297-303.

[6] 朱光，秉岩，刘丽娟. 上海市某区老年高血压患者焦虑状况及其影响因素[J]. 四川精神卫生，2022，35(1)：26-30.
Zhu G, Bing Y, Liu LJ. Prevalence and influencing factors of anxiety status among elderly hypertensive patients in a district of Shanghai[J]. Sichuan Mental Health, 2022, 35(1): 26-30.

[7] 万崇华，罗家洪. 高级医学统计学[M]. 北京：科学出版社，2014：235-264.
Wan CH, Luo JH. Advanced medical statistics [M]. Beijing: Science Press, 2014: 235-264.


（收稿日期：2022-11-13）
（本文编辑：戴浩然）