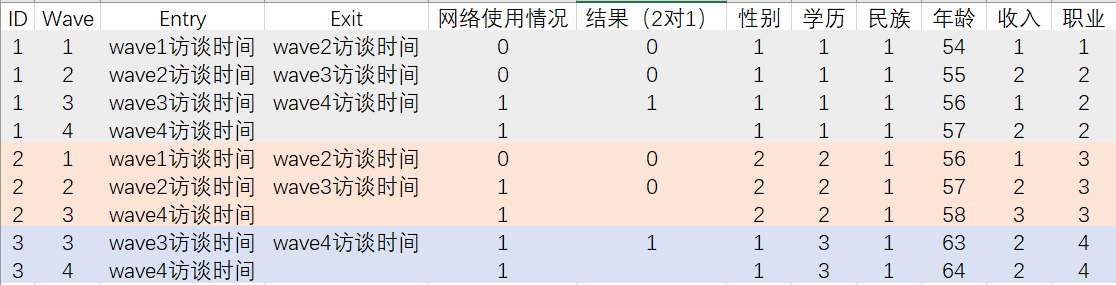
**变量编码情况**

**数据格式**

****

**一、协变量（共12个，表格中标黄部分）**

**1、Gender：** Male=1、Female=2

**2、Education level（Edu）：**

* Illiterate=1
* ≤ Elementary school=2
* Middle school=3
* ≥High school=4

**3、Nationality（Nat）：**Han=1、Ethnic minorities=2

**4、Work：：**Yes=1、No=2

**5、Age(years)：**连续数据，不分段，直接写入

**6、Area：**

* Central of City/Town=1
* Combination zone between urban and rural areas=2
* Village=3

**7、Marital status（Mari）：**

* Married/Cohabited=1
* Never-married=2
* Divorced/Widowed/separated=3

**8、Health insurance（Ins）：**

* 有社保也有商业保险=1
* 有社保没有商业保险=2
* 无保险=3

**9、smoking status（Smoking）：** Yes=1、No=2

**10、drinking status（Drinking）：** Yes=1、No=2

**11、Restless sleep times（Sleep）**：

* Rare=1
* Some or a little=2
* Occasionally=3
* Most or all=4

**12、ADL disability（ADL）：**

* health=1
* mild disability=2
* moderate disability=3
* partial disability=4
* severe disability=5

**二、自变量（共2个，表格中标蓝部分）**

**1、Internet Use（Internet\_Use）：** Yes=1、No=0

**2、Frequency（Freq）：**

* NA=0
* Not Regularly=1
* Almost Every Week=2
* Almost Daily=3

**三、因变量（共16个，表格中标绿部分）**

**1、慢性病发病情况（Chor\_1——14 共Chor\_14种）**： Yes=1、No=0

* **高血压（Chor\_1）**
* 血脂异常（高血脂或低血脂）
* 糖尿病或血糖升高（包括糖耐量异常和空腹血糖升高）
* 癌症等恶性肿瘤（不包括轻度皮肤癌）
* 慢性肺部疾患如慢性支气管炎或肺气肿、肺心病（不包括肿瘤或癌）
* 肝脏疾病（除外脂肪肝、肿瘤或癌）
* 心脏病（如心肌梗塞、冠心病、心绞痛、充血性心力衰竭和其他心脏疾病）
* 中风（包括脑梗和脑出血）
* 肾脏疾病（不包括肿瘤或癌）
* 胃部疾病或消化系统疾病（不包括肿瘤或癌）
* 情感及精神方面问题
* 与记忆相关的疾病（如老年痴呆症、脑萎缩、帕金森症）
* 关节炎或风湿病
* 哮喘

**2、视力损伤情况（Vis）**：

Yes=1、No=0

**3、抑郁情况（Dep）**：

Yes=1、No=0

以上是现有所以数据的变量编码情况。

因为因变量部分每个病患病率是不一样的，所以我们想的是每一个因变量都单独做一次cox，但是不麻烦师兄每一个都做。师兄以其中的一个疾病（比如高血压）为例进行基线筛选，然后做cox分析，之后师兄把代码发给我，告诉我们做下一次时哪些地方需要更改，然后剩下的因变量分析我们来完成就可以**。**

**以高血压为例，师兄限定在基线数据（即每一个参与人员的第一次调查）没有高血压（即Chro-1为0）的人中进行cox分析，结局有两种：y1=1：发生高血压，y1=0：未发生高血压，一旦发生随访时间终点（即最后一次随访时间）或者发生高血压即停止计算随访时间。**