

利用 R 语言构建 CHS-DRG 分组器*

徐州医科大学附属医院病案统计科(221000) 马赫[△] 张翔 孔甜甜 卞静 文娟

【摘要】 目的 探讨利用 R 语言构建 CHS-DRG 分组器的方法和思路。方法 利用实例论证如何使用 R 语言编写分组器,并给出数据结构及程序代码。结果 使用 R 语言自行构建分组器,相对商业软件可大幅节省成本,方便定制功能。在构建分组器时,应注意理解分组方案思想,同时根据地域差异灵活调整代码。结论 医疗机构可自行编写 DRG 分组器,以适应 DRG 浪潮,提升自身竞争力。

【关键词】 R 语言 CHS-DRG 分组器

【中图分类号】 R195.1 **【文献标识码】** A

DOI 10.11783/j.issn.1002-3674.2025.01.028

疾病诊断相关组(diagnosis related groups, DRG)是一种患者分组模式,常用于医保支付、医院管理和绩效评价。2021 年国家医保局发布《国家医疗保障局关于印发 DRG 支付方式改革三年行动计划的通知(医保发[2021]48 号)》,明确到 2024 年,全国所有统筹地区全面开展 DRG 付费方式改革工作。可以看出,推行 DRG 已是大势所趋。

DRG 系统运行依赖医院内部信息系统的支持,虽然国家医保局制定的国家医疗保障疾病诊断相关组(China healthcare security diagnosis related groups, CHS-DRG)分组方案已成为事实上的行业标准并对外公开,但绝大多数医院的信息部门无法自行建设基于 DRG 的医保结算及绩效管理考核系统。而商业 DRG 系统售价高昂,是一些较小型医疗机构难以承受的。因此,自行研发 DRG 管理系统,掌握核心技术,既能节省成本,也能跟随国家快速迭代 DRG 版本,益处颇多。

DRG 管理系统的核心是 DRG 分组器。目前少数已发表文献均是从构建 DRG 分组器的理论原理及规则入手^[1-3],尚未有文献对如何使用编程语言可实操的构造分组器的技术细节进行报道。

R 语言是一种统计专用语言,具有免费、易学、灵活、扩展性高等优点,特别适合数据清洗和计算,可部署于单机或服务器上,不要求使用者具有计算机专业背景。因此本研究拟利用 R 语言构建 DRG 分组器,并展示其技术细节及注意事项,为医疗机构自行构建 DRG 管理系统核心提供一种范例。

软件环境与数据结构

本研究代码适用于 R 语言 3.6 及更新版本。R 语言可接受 excel 文件、数据库直连等不同数据输入,建议数据在 R 中的保存类型为 data.frame,以利于后期的数据整合和计算。另外,由于分组器程序可能需要部署

在医院内网上,不与外部互联网连接,考虑到 R 语言离线安装外部扩展包并不方便,分组器本身代码全部使用基础 R 语言构建,不使用 tidyverse 等扩展包。

本研究分组标准根据《国家医疗保障疾病诊断相关分组(CHS-DRG)分组方案(1.1 版)》(以下简称“分组方案”)构建。CHS-DRG 分组器主要根据医院出院病案首页中的诊断信息及手术信息进行分组,患者基本信息作为辅助分组信息。本分组器采用的出院病案首页原始数据架构根据徐州市医保局制定的《徐州市 DRG 综合管理应用平台与医院信息系统接口规范 V1.6》,取出院前 16 个出院诊断编码、前 8 个手术诊断编码作为主要分组依据,辅助分组信息主要为患者年龄(以天计)和新生儿出生体重(克),用于主要诊断大类 P 分组(major diagnostic category P, MDGP)新生儿分组。这些字段将作为参数纳入分组器中,它们在程序中的参数名分别为:主要诊断编码(ZYZDBM_YB)、其他诊断编码 1(QTZDBM1_YB)、其他诊断编码 2(QTZDBM2_YB)……其他诊断编码 15(QTZDBM15_YB)、手术操作编码 1(SSJCZBM_YB1)、手术操作编码 2(SSJCZBM_YB2)……手术操作编码 8(SSJCZBM_YB8)、患者年龄(BZYZSNL)、新生儿出生体重(XSECSTZ)。其他地区的医疗机构可根据实际需要自行增减字段。对于待分组的原始数据,本分组器采用的数据输入架构为一位患者一次住院的所有信息在同一行,结构为表 1 所示。其中主要诊断编码(ZYZDBM_YB)要求非空,其他字段不做要求。

DRG 分组器除了加载待分组首页数据外,在分组时还需加载分组方案数据库,而其数据量较大,不可直接写死在程序内,因此需要存储在数据集中,以供程序调用。同出院病案首页数据一样,分组方案数据读取至 R 中时也建议使用 data.frame 格式。

DRG 分组原理主要包括两阶段:核心疾病诊断相关组(adjacent diagnosis related groups, ADRG)分组阶段和考虑并发症的 DRG 分组阶段,因此需加载的分组方案数据包括 ADRG 分组阶段的诊断表和手术表,

*基金项目:2022 年徐州市卫生健康委医学科技创新项目(XWKYHT20220141)

[△]通信作者:马赫, E-mail: mahexz@163.com

及 DRG 分组阶段的严重并发症表、一般并发症表及 并发症排除表,共 5 张表。

表 1 病案首页原始数据输入架构示例

住院号	ZYZDBM_YB	QTZDBM1_YB	QTZDBM2_YB	...	QTZDBM15_YB	SSJCZBM_YB	SSJCZBM_YB2	...	SSJCZBM_YB8	BZYZSNL	XSECSTZ
10001	L08.900	G82.200				86.7503	86.2200	...	77.6802		
10002	P55.101	P29.800x902	G93.000x015							4	3640
10003	G04.904	A86.x00	I61.900x002	...	E11.900	96.7201	31.1x00x005	...	38.9800x001		

ADRG 诊断表的存储结构如表 2 所示,分为三列,第一列存储诊断编码,第二列存储所属的 MDC 名。第三列存储该 MDC 下内科组的组名,三列的列名分别为 bianma、mdcsscz、adrgnk。有些外科手术操作 ADRG 除了有对应的主手术编码要求外,还有自己特殊的主诊断编码要求,如 BC1 和 DA1,这些特殊主诊断编码表是所属 MDC 主诊断编码列表的子集,根据 R 语言的特性,长表的处理速度要快于宽表,因此为了减少整张表的列数,这类外科组要求的特殊主诊编码可再次在 bianma 列列出,其 mdcsscz 列展示对应 ADRG 名,adrgnk 列展示对应内科组名称,如表 2 的 I60.001 和 C00.101。MDCZ 主诊表按照身体部位分为九个部分,要求患者主诊断在一个部位的编码范围中,并且至少一个其他诊断在另一个部位的编码范围中,结构与一般 MDC 不同,因此存储架构也比较特殊,主要区别在 adrgnk 列展示对应编码的部位而不是内科组的组名,具体例子见表 2。

表 2 ADRG 诊断表示例

bianma	mdcsscz	adrgnk
I60.001	MDCB	BR1
G96.100x001	MDCB	BZ1
I60.001	BC1	BR1
H35.004	MDCC	CX1
C00.101	MDCD	DR1
C00.101	DA1	DR1
S01.800x011	MDCZ	头颈部
S21.100x002	MDCZ	胸部
S31.800x003	MDCZ	腹部
S35.400x001	MDCZ	泌尿
S35.500x005	MDCZ	生殖
S12.000	MDCZ	躯干
S14.300	MDCZ	上肢
S71.800x011	MDCZ	下肢
S31.800x012	MDCZ	骨盆

ADRG 手术表的结构较之诊断表相对简单,分为两列:第一列 bianma 列展示手术编码,第二列 adrg 列展示所属外科手术组的组名,示例见表 3。

表 3 ADRG 手术表示例

bianma	adrg
34.0200x003	PB1
34.0301	PB1
80.5104	ZC1
80.5105	ZC1
02.0504	BB1
02.0505	BB1

DRG 严重并发症表分为两列:第一列 zbbianma 列展示严重并发症的诊断编码,第二列 pcb 列展示对应排除表的表名,见表 4。DRG 一般并发症表结构与严重并发症表结构相同,见表 5。

表 4 DRG 严重并发症表示例

zbbianma	pcb
A01.100	表 7-1
A01.200	表 7-1
A01.300	表 7-1
A17.800	表 7-2
A19.100	表 7-2
A20.000x001	表 7-3

注:pcb 列中的表名,指的是《国家医疗保障疾病诊断相关分组(CHS-DRG)分组方案(1.1 版)》中严重并发症或合并症排除列表的表序号,具体请参阅 CHS-DRG 国家试点技术指导组网站 <https://www.chsdr.org.cn/technicalplan.html>。

表 5 DRG 一般并发症表示例

zbbianma	pcb
A00.100	表 7-1
A00.100x001	表 7-1
A15.000x020	表 7-2
A15.000x022	表 7-2
A25.000x001	表 7-3
A25.100	表 7-3

注:pcb 列中的表名,指的是《国家医疗保障疾病诊断相关分组(CHS-DRG)分组方案(1.1 版)》中严重并发症或合并症排除列表的表序号,具体请参阅 CHS-DRG 国家试点技术指导组网站 <https://www.chsdr.org.cn/technicalplan.html>。

DRG 严重并发症表与一般并发症表共用一个主诊断排除表。主诊断排除表的结构同样为两列:第一列 pcbbianma 列为需要排除的编码,第二列 pcb 列为对应的排除表表名,见表 6。需要注意的是,因在程序中需要按照排除表表名将严重并发症表、一般并发症表分别与排除表合并成两张表,类似于数据库的完全连接(FULL JOIN),所以排除表第二列的列名一定要与严重并发症表和一般并发症表的第二列列名相同,如示例中的 pcb。

表 6 主诊断排除表示例

pcbbianma	pcb
A00.000x001	表 7-1
A00.100x001	表 7-1
A15.501	表 7-2
A15.502	表 7-2

注:pcb 列中的表名,均指的是《国家医疗保障疾病诊断相关分组(CHS-DRG)分组方案(1.1 版)》中严重并发症或合并症排除列表的表序号,具体请参阅 CHS-DRG 国家试点技术指导组网站 <https://www.chsdr.org.cn/technicalplan.html>。

程序示例

R 语言作为统计与数据分析专用语言,其中一个特点便是强调函数式编程,将程序写成函数的形式利用 apply 族函数执行,其效率和速度上都明显好于循环结构。因此本研究按照 DRG 分组原理将程序分为两个部分,ADRG 分组和 DRG 分组,并将其写成两个大型函数,以利于其他函数的调用,提升效率和灵活性。

1. ADRG 分组

首先应加载 ADRG 分组所需的分组方案数据:诊断表和手术表,变量名分别为 jibing 和 shoushu,二者在下述 ADRG 分组函数中将会被调用。

根据 R 语言语法,ADRG 分组函数的总体框架结构为:

```
ADRG<- function( BZYZSNL=99999999, XSECSTZ=999999, ZYZDBM_YB, QTZDBM1_YB='wu', QTZDBM2_YB='wu', ..., QTZDBM15_YB='wu', SSJCZBM_YB1='wu', SSJCZBM_YB2='wu', ..., SSJCZBM_YB8='wu') {
  quanbu.jibing<- c( ZYZDBM_YB, QTZDBM1_YB, QTZDBM2_YB, ..., QTZDBM15_YB)
  quanbu.shoushu<- c( SSJCZBM_YB1, SSJCZBM_YB2, ..., SSJCZBM_YB8)
  if( ... ) return( ... )
  else if( ... ) return( ... )
  else if( ... ) return( ... )
  ...
  else return( '未入组' )
}
```

此函数需要 BZYZSNL、XSECSTZ、ZYZDBM_YB 等 26 个字段作为输入参数,除主要诊断编码(ZYZDBM_YB)外,其余 25 个字段允许为空值,为防止空值导致程序出错,这些字段均给予默认值。患者年龄(BZYZSNL)、新生儿出生体重(XSECSTZ)为数值型参数,需要比较大小,因此一般非新生儿病例的这两个字段若为空值,为避免入 MDCP,默认值可给予一个很大的数字,如上文程序中的 99999999 或 999999。其余字段为字符型参数,默认值给予一个诊断或手术编码里不可能的值既可,本例中给予“wu”。函数内部,所有诊断和所有手术分别合并为向量 quanbu.jibing 和 quanbu.shoushu,以备使用。

本函数使用 if...else 结构根据参数判断分组,通常一行代码为一个 ADRG,返回值为 ADRG 名,如 BB1、BC1 等,具体见下文。若不符合所有入组条件,将返回“未入组”。R 语言程序一般按照由上到下的顺序执行,因篇幅所限,根据分组方案顺序,本文先对

MDCA、MDCP、MDCY、MDCZ 四个先期分组中入组条件比较特殊的 ADRG 进行描述,然后讲解常规 ADRG 代码的一般写法。

(1) 先期分组

①MDCA 中 AC1:

```
else if( any( quanbu.shoushu %in% c('52.8000', '52.8200', '52.8300'), na.rm = T) & any( quanbu.shoushu %in% c('55.6100', '55.6901'), na.rm=T) ) return('AC1')
```

AC1 组的入组要求是患者的所有手术编码中至少有一个编码为 52.8000、52.8200、52.8300 三者之一并且还应有至少一个手术编码为 55.6100、55.6901 两者之一,因此使用 quanbu.shoushu 向量进行判断,quanbu.shoushu %in% c('52.8000','52.8200','52.8300') 返回一个长度与 quanbu.shoushu 相同的布尔逻辑向量,向量中只要有一个值为 TRUE,则认为符合条件,因此使用 any() 函数。最后通过使用“&”将两个条件连接起来。

②MDCP 各组,以 PS1 为例:

```
else if( BZYZSNL < 29 & XSECSTZ < 1500 & any( quanbu.jibing %in% subset(jibing, adrgnk == 'PS1')[,1], na.rm=T) ) return('PS1')
```

MDCP 组的特点是每个 ADRG 组均以年龄作为入组条件,且 PS1~PS4 四个组对出生体重有不同的要求,除此之外与一般的 MDC 区别不大。主要诊断要求符合 MDCP 的主诊表,即 ZYZDBM_YB %in% subset(jibing, mdcscsz == 'MDCP')[,1],其中 subset() 函数是在 ADRG 诊断表中取 MDCP 的子集,如表 2 结构所示,“[,1]”为取子集的第一列,即编码列(bianma)。手术组还要求主手术符合 MDCP 下的 ADRG 相应要求,如 SSJCZBM_YB1 %in% subset(shoushu, adrg == 'PB1')[,1],表结构如表 3 所示。最后年龄、诊断和手术的“并且”关系使用 & 符号连接,即构成了手术操作组的入组条件。内科组的入组条件去掉了手术要求,主诊断的代码与手术组有细微的差别,即 subset 函数中由取 mdcscsz 列变为取 adrgnk 列。另外 PS1~PS4 组的入组条件为所有诊断中只要有一个编码在诊断表中既可,所以使用了 any() 函数。其余代码无特别之处。

③MDCZ 各组,以 ZB1 为例:

```
else if( ZYZDBM_YB %in% subset(jibing, mdcscsz == 'MDCZ')[,1] & sum( any( quanbu.jibing %in% subset(jibing, adrgnk == '腹部')[,1], na.rm = T), any( quanbu.jibing %in% subset(jibing, adrgnk == '骨盆')[,1], na.rm = T), ..., any( quanbu.jibing %in% subset(jibing, adrgnk == '胸部')[,1], na.rm = T), na.rm = T) > 1 & SSJCZBM_YB1 %in% subset
```

```
(shoushu, adrg == 'ZB1')[,1])return('ZB1')
```

ZB1 入组条件由三部分构成。第一,主诊断编码在 MDCZ 的主诊表中,即 `ZYZDBM_YB %in% subset(jibing, mdcscsz == 'MDCZ')[,1]`。第二, MDCZ 的主诊表由腹部、骨盆、泌尿等 9 个分表组成, MDCZ 要求患者所有诊断中至少有两个以上的诊断编码属于两个以上不同的分表,故我们首先利用 `any(quantbu.jibing %in% subset(jibing, adrgnk == 分表名)[,1], na.rm=T)` 判断所有诊断编码中是否有至少一个编码属于该分表,然后我们利用 `sum()` 函数计算 `any()` 函数返回的 True 的数量,若大于 1 则符合条件。第三,对于 ZB1 等手术组而言,主手术还应符合相应的手术编码表,最后使用 `&` 符号连接即构成 ZB1 的入组条件。其余手术组的入组条件与 ZB1 前两点相同,将主手术编码表的名称更换为相应 ADRG 组名既可。内科组 ZZ1 入组条件与手术组前两点相同,除此之外还应并列一个条件: `!SSJCZBM_YB1 %in% shoushu[,1]`,即主手术不在表 3 的所有手术编码中或不存在手术,才可入 ZZ1 组。

(2) 其余 ADRG 分组

其余 ADRG 分组按分组方案中的顺序编写,共 379 个。普通 ADRG 分组的代码分为三种情况:第一,对主诊断的要求除了符合相应 MDC 主诊断表外还需符合本 ADRG 的主诊断子表的手术操作组,如 BC1,该类型 ADRG 的代码一般为: `else if(ZYZDBM_YB %in% subset(jibing, mdcscsz == 'BC1')[,1] & SSJCZBM_YB1 %in% subset(shoushu, adrg == 'BC1')[,1])return('BC1')`;第二,本 ADRG 无特定诊断子表要求,只需符合所属 MDC 主诊表即可的手术操作组,如 BC2,该类型 ADRG 的代码一般为: `else if(ZYZDBM_YB %in% subset(jibing, mdcscsz == 'MDCB')[,1] & SSJCZBM_YB1 %in% subset(shoushu, adrg == 'BC2')[,1])return('BC2')`;第三,内科组,以 BR1,为例,该类型 ADRG 的代码一般为: `else if(ZYZDBM_YB %in% subset(jibing, adrgnk == 'BR1')[,1])return('BR1')`。编写完所有 ADRG 分组,代码的最后加上 `else return('不入组')`,以标记不满足上述任何 ADRG 入组条件的病例。

2. DRG 分组

使用 `ADRG()` 函数完成 ADRG 分组后,还需完成第二阶段考虑并发症的 DRG 分组。首先加载 DRG 分组所需的分组方案数据: DRG 严重并发症表、DRG 一般并发症表及主诊断排除表,分别赋给变量 `mcc`、`cc` 及 `paichu`。然后使用 `merge` 函数将严重并发症表、一般并发症表分别与排除表合并为 `quan.mcc` 及 `quan.cc` 两个变量,需注意的是, `merge` 函数的 `all` 参数要设置为 True。

DRG 分组函数的总体框架结构为:

```
DRG<- function ( ZYZDBM_YB = 'wu', QTZDBM1_YB = 'wu', ..., QTZDBM15_YB = 'wu' ) {
  if( ( QTZDBM1_YB %in% quan.mcc $ zbbianma & ! ZYZDBM_YB %in% subset( quan.mcc, zbbianma == QTZDBM1_YB ) [,3] ) |
  .....
  | ( QTZDBM15_YB %in% quan.mcc $ zbbianma & ! ZYZDBM_YB %in% subset( quan.mcc, zbbianma == QTZDBM15_YB ) [,3] ) ) return( '1' )
  else if( ( QTZDBM1_YB %in% quan.cc $ zbbianma & ! ZYZDBM_YB %in% subset( quan.cc, zbbianma == QTZDBM1_YB ) [,3] ) |
  .....
  | ( QTZDBM15_YB %in% quan.cc $ zbbianma & ! ZYZDBM_YB %in% subset( quan.cc, zbbianma == QTZDBM15_YB ) [,3] ) ) return( '3' )
  else return( '5' )
}
```

根据分组方案,满足条件下,优先划入严重并发症组(标记为 1),如不满足,再与一般并发症表比较,满足条件下划入一般并发症组(标记为 3),若也不满足则划入不伴并发症组(标记为 5),由此,程序代码也同样分为三部分,最后以 `else return('5')` 结尾。在判断严重并发症时,需要判断每个其他诊断编码是否在严重并发症表中,若在,还需主要诊断编码不能在其他诊断 1 对应的排除表中,即 `! ZYZDBM_YB %in% subset(quantbu.jibing, mdcscsz == 'MDCB')[,1] & SSJCZBM_YB1 %in% subset(shoushu, adrg == 'BC1')[,1]`,这两个条件使用“&”连接。每个其他诊断都需进行判断,只要有一个其他诊断和主诊断满足条件即进入严重并发症组,故条件之间使用“|”符号相连。一般并发症组的入组条件与之类似,不再赘述。

3. 函数调用与并行计算

`ADRG()` 与 `DRG()` 函数构建完成后即可对具体病例进行分组,可通过 `apply` 函数族中的 `mapply()` 函数进行调用,以 `ADRG()` 为例,即 `mapply(FUN = ADRG, SIMPLIFY = T, BZYZSNL = x $ BZYZSNL, XSECSTZ = x $ XSECSTZ, ZYZDBM_YB = x $ ZYZDBM_YB, QTZDBM1_YB = x $ QTZDBM1_YB, QTZDBM2_YB = x $ QTZDBM2_YB, QTZDBM3_YB = x $ QTZDBM3_YB, ...)`,其中 `x` 为待分组的病例数据集,运行后将得到一个包含每个病例的 ADRG 名称的向量,即可将该向量合并回原数据集。`DRG()` 函数与此类似,调用运行后可得到一个包含每个病例 1、3、5 并发症标记的向量,与 ADRG 名称向量合并后即得到如 BB11、FT13 的 DRG 组名。

`mapply()` 函数运行时仅使用 CPU 中的单核进行

运算,对 CPU 计算能力利用欠佳。若医疗机构数据量大,推荐使用 R 自带 parallel 包中的 clusterMap() 函数,它是 mapply() 函数的并行版本,通过调用所有 CPU 内核参与运算,可明显缩短程序运行时间,具体 clusterMap() 的使用请参见 parallel 包的说明文档,此处不再赘述。

讨 论

本研究通过实例探讨了如何使用 R 语言构建 DRG 分组器。由上文可以看出:①编写代码时应先分析需求,从总体程序架构入手,由上而下,逐步深化具象问题;②要深刻理解 R 语言的语法特性,做到灵活使用。如 MDCZ 中的 SUM() 和 ANY() 函数结合用来计算满足 ADRG 诊断表的个数,SUM() 函数一般用于计算数值型变量之和,在本例中 ANY() 函数返回逻辑值 TRUE 或 FALSE,TRUE 在 SUM() 中等价于 1,FALSE 等价于 0,故 SUM() 可用于计算 TRUE 的个数。

另外需要指出的是,本研究重点在于从易于理解的角度为编写分组器提供一个范例和思路,代码的结

构和性能尚有提升优化的空间。其实编写分组器的关键不在于编程技术细节,而是应明确分组方案精神,深刻体会 CHS-DRG 的设计思想,以使用场景为导向,做到既能匹配上游输入数据格式,又能方便下游数据利用。另外不同地区对于分组方案的解释存在差异,各医疗机构可根据当地 DRG 结算平台返回的入组信息自行对代码进行调整。

本程序既可用于个人计算机进行小范围数据分析,也可部署于服务器上进行病例预分组和医院管理,相对于商业软件,节省了大量成本,尤其适用于中小型医疗机构,以助其适应 DRG 浪潮,提升竞争力。

参 考 文 献

- [1] 魏倩,陈尔真,孙木.S-DRG 分组器的构建与改进方法[J].中国卫生质量管理,2020,27(1):61-65.
- [2] 许岩,孙木,何萍,等.上海市医院疾病诊断分组模型及分组器的建立[J].中国卫生政策研究,2015,8(9):15-18.
- [3] 史森,王晖.基于 DRGs 分组器的北京农村学生儿童重大疾病医疗保障信息系统的设计与研究[J].中国卫生信息管理杂志,2011,8(5):60-62.

(责任编辑:郭海强)

(上接第 139 页)

人员有限的研究可能不具备应用 EDC 系统相应功能的能力。对于数据质量监测工作,纠正疑问数据是最终目的,发现疑问数据是首要和重要的手段。本文基于医学研究工作实际,提出了利用已有的工具(SAS 和计算机操作系统)、尽可能自动化完成此项工作的一种思路和方法。

医学研究领域宽泛、研究设计类型众多。然而,最终获得的数据集形式、变量类别是相似的。针对数据集进行核查的内容也大体相似,包括对单一数值或字符变量缺失或异常赋值、多个变量之间交叉逻辑关系(重复、一致/不一致、逻辑矛盾等)、日期时间变量的逻辑关系(日期范围、逻辑矛盾等)等^[4]。对于研究团队而言,无论是否采用 EDC 系统,制定详细的数据核查计划是数据核查的必要前置工作^[5]。在确定了核查类型后,撰写简单的 SAS 宏程序,即可较为便捷地完成具体的核查工作,实现与 EDC 系统发现疑问数据相似的功能。

SAS 是一个集数据访问和管理、数据分析、报告编制等功能于一体的模块化、集成化的大型应用软件,最常被使用的是其统计分析模块。既往也有研究利用该软件对公共卫生监测数据进行核查^[6]。本文在此

基础上,进一步利用了 SAS/ACCESS 模块,直接访问服务器上的数据,保证程序运行时使用的是最新的数据集;在完成数据核查后,可利用 SAS 直接发邮件报告,节省了人工成本。更重要的是,通过与计算机操作系统计划任务功能的配合,即可实现定时、自动的运行制定的 SAS 核查程序,帮助数据管理员在没 EDC 系统协助的情况时,也能完成初步的数据核查工作。

参 考 文 献

- [1] 刘晋,梁宁霞,李天萍,等.电子数据采集系统在临床科研管理中的实践与思考[J].中华医院管理杂志,2019,35(6):503-506.
- [2] 王娟,胡晓雯,周梅红,等.药物临床试验网络信息服务系统设计与实现[J].科技信息,2010,(34):5.
- [3] 谢高强,姚晨.数据管理在临床研究中的地位和作用[J].北京大学学报(医学版),2010,42(6):641-643.
- [4] 周蓓,于浩.临床试验逻辑核查的分类及应用[J].中国临床药理学与治疗学,2019,24(6):670-674.
- [5] 中国国家食品药品监督管理总局.总局关于发布临床试验数据管理工作技术指南的通告(2016 年第 112 号)[EB/OL].2016. <http://samr.cfda.gov.cn/WS01/CL0087/160961.html>
- [6] 李洪兴,董国庆,王丽,等.SAS 软件在公共卫生监测数据核查中的应用[J].中国卫生统计,2012,29(5):749-751.

(责任编辑:郭海强)