coloncancer

2024-10-20

```
#only need to do this once, unhashtag these lines to run
#install.packages("biostat3")
#install.packages("gtsummary")
```

```
#bring in the data, you would bring in your own data here
library(biostat3)
```

```
## Loading required package: survival
```

```
## Loading required package: MASS
```

```
##
## Attaching package: 'biostat3'
```

```
## The following object is masked from 'package:survival':
##
##     colon
```

```
data(colon)
head(colon)
```

```
##      sex age     stage mmdx yydx surv_mm surv_yy       status
## 1 Female  77   Distant    3 1977    16.5     1.5 Dead: cancer
## 2 Female  78 Localised    7 1978    82.5     6.5  Dead: other
## 3   Male  78   Distant   10 1978     1.5     0.5 Dead: cancer
## 4   Male  76   Distant   10 1976     1.5     0.5 Dead: cancer
## 5   Male  80 Localised   12 1980     8.5     0.5 Dead: cancer
## 6 Female  75 Localised   11 1975    23.5     1.5 Dead: cancer
##                  subsite        year8594 agegrp         dx       exit id
## 1             Transverse Diagnosed 75-84    75+ 1977-03-04 1978-07-20  1
## 2   Coecum and ascending Diagnosed 75-84    75+ 1978-07-26 1985-06-11  2
## 3 Descending and sigmoid Diagnosed 75-84    75+ 1978-10-10 1978-11-25  3
## 4 Descending and sigmoid Diagnosed 75-84    75+ 1976-10-28 1976-12-13  4
## 5 Descending and sigmoid Diagnosed 75-84    75+ 1980-12-20 1981-09-05  5
## 6   Coecum and ascending Diagnosed 75-84    75+ 1975-11-16 1977-11-01  6
##        ydx    yexit
## 1 1977.170 1978.548
## 2 1978.564 1985.441
## 3 1978.772 1978.898
## 4 1976.824 1976.950
## 5 1980.969 1981.676
## 6 1975.873 1977.832
```

```
library(gtsummary)
```

```
##
## Attaching package: 'gtsummary'
```

```
## The following object is masked from 'package:MASS':
##
##     select
```

```
colon |> tbl_summary(include = c(sex,age, stage,subsite))
```

Characteristic	N = 15,564[1]
sex	
Male	6,340 (41%)
Female	9,224 (59%)
age	71 (62, 78)
stage	
Unknown	2,356 (15%)
Localised	6,274 (40%)
Regional	1,787 (11%)
Distant	5,147 (33%)
subsite	
Coecum and ascending	5,576 (36%)
Transverse	2,734 (18%)
Descending and sigmoid	6,270 (40%)
Other and NOS	984 (6.3%)

[1] n (%); Median (Q1, Q3)

```
colon |> tbl_cross(row= stage, col= subsite) |> add_p()
```

	subsite					
	Coecum and ascending	Transverse	Descending and sigmoid	Other and NOS	Total	p-value[1]
stage						<0.001
Unknown	756	369	879	352	2,356	
Localised	2,235	1,025	2,761	253	6,274	
Regional	798	291	621	77	1,787	
Distant	1,787	1,049	2,009	302	5,147	
Total	5,576	2,734	6,270	984	15,564	

[1] Pearson's Chi-squared test

```
colon |> tbl_summary(include = c(age,stage,subsite), by = sex) |> add_p()
```

Characteristic	Male N = 6,340[1]	Female N = 9,224[1]	p-value[2]
age	69 (60, 76)	73 (64, 79)	<0.001
stage			0.004
Unknown	885 (14%)	1,471 (16%)	
Localised	2,620 (41%)	3,654 (40%)	
Regional	715 (11%)	1,072 (12%)	
Distant	2,120 (33%)	3,027 (33%)	
subsite			<0.001
Coecum and ascending	2,132 (34%)	3,444 (37%)	
Transverse	1,139 (18%)	1,595 (17%)	
Descending and sigmoid	2,671 (42%)	3,599 (39%)	
Other and NOS	398 (6.3%)	586 (6.4%)	

[1] Median (Q1, Q3); n (%)

[2] Wilcoxon rank sum test; Pearson's Chi-squared test