# REPEATED EVENTS

# INTRODUCTION

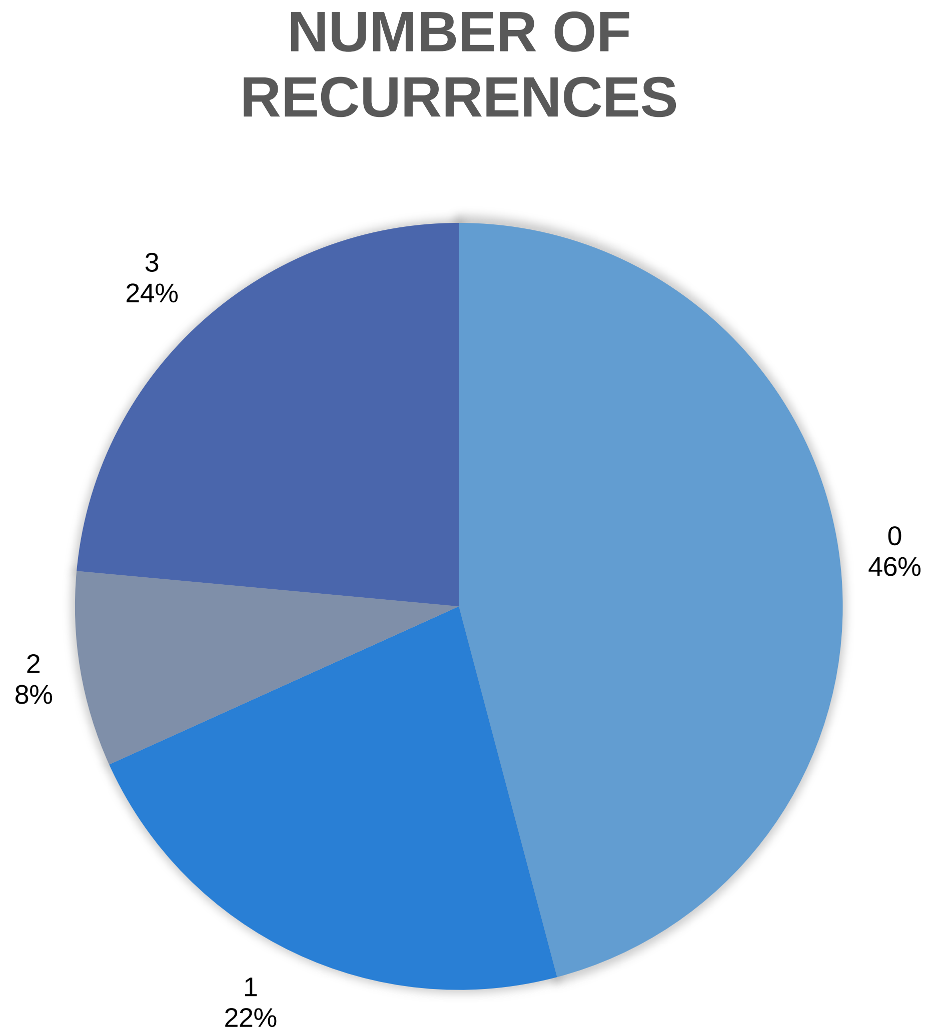
# Multiple Events

- Previously discussed how to analyze:
  - Time to **single** event
  - Time to **one** of many events

- What if we extended this again to the possibility of multiple occurrences of a single event?
- **Repeated events**, like competing risks, is a particular type of multi-state analysis that builds upon the previous things we have learned.
- We will be using the PH model for all of these examples

# Bladder Tumors Data Set

- Randomized trail of 85 patients.
- Count of recurrences of bladder tumors.
- Andrews DF, Hertzberg AM (1985)
- Subjects were followed for 64 months



# Bladder Tumors Data Set

- **Start**: Either a 0 or time of previous recurrence (in months)
- **Stop**: Current recurrence time (or time of censoring)
- **Event**: Tumor recurrence during the observed **start**, **stop** 1 if tumor, 0 if no tumor (at stop time)
- **ID**: Patient ID
- **rx**: placebo (1) or treatment (2) group (either placebo or thiotepa)
- **number**: number of tumors initially present (truncated at 8)
- **size**: diameter (cm) of largest initial tumor
- **enum**: # of previous times with tumors (up to max of 4)

# MODELS FOR REPEATED EVENTS

“Cluster” Model

# Treat Individuals as a “cluster”

- Looking at time since randomization of a “treatment” (drug in clinical trial, became a customer, etc)…known as “total time scale”
- Assume correlation between event times for a person can be explained by past events (time increments between events are conditionally uncorrelated)
- Only care about the overall effect, ignoring the order or type of recurrence.

# Example Patients

| ID | rx | number | size | start | stop | event | enum |
|---|---|---|---|---|---|---|---|
| 5 | 1 | 4 | 1 | 0 | 6 | 1 | 1 |
| 5 | 1 | 4 | 1 | 6 | 10 | 0 | 2 |
| 13 | 1 | 3 | 1 | 0 | 3 | 1 | 1 |
| 13 | 1 | 3 | 1 | 3 | 9 | 1 | 2 |
| 13 | 1 | 3 | 1 | 9 | 21 | 1 | 3 |
| 13 | 1 | 3 | 1 | 21 | 23 | 0 | 4 |
| 16 | 1 | 1 | 2 | 0 | 26 | 0 | 1 |
| 41 | 1 | 3 | 1 | 0 | 35 | 1 | 1 |
| 41 | 1 | 3 | 1 | 35 | 51 | 0 | 2 |

# Example Patients

| ID | rx | number | size | start | stop | event | enum |
|---|---|---|---|---|---|---|---|
| 5 | 1 | 4 | 1 | 0 | 6 | 1 | 1 |
| 5 | 1 | 4 | 1 | 6 | 10 | 0 | 2 |
| 13 | 1 | 3 | 1 | 0 | 3 | 1 | 1 |
| 13 | 1 | 3 | 1 | 3 | 9 | 1 | 2 |
| 13 | 1 | 3 | 1 | 9 | 21 | 1 | 3 |
| 13 | 1 | 3 | 1 | 21 | 23 | 0 | 4 |
| 16 | 1 | 1 | 2 | 0 | 26 | 0 | 1 |
| 41 | 1 | 3 | 1 | 0 | 35 | 1 | 1 |
| 41 | 1 | 3 | 1 | 35 | 51 | 0 | 2 |

# Example Patients

| ID | rx | number | size | start | stop | event | enum |
|---|---|---|---|---|---|---|---|
| 5 | 1 | 4 | 1 | 0 | 6 | 1 | 1 |
| 5 | 1 | 4 | 1 | 6 | 10 | 0 | 2 |
| 13 | 1 | 3 | 1 | 0 | 3 | 1 | 1 |
| 13 | 1 | 3 | 1 | 3 | 9 | 1 | 2 |
| 13 | 1 | 3 | 1 | 9 | 21 | 1 | 3 |
| 13 | 1 | 3 | 1 | 21 | 23 | 0 | 4 |
| 16 | 1 | 1 | 2 | 0 | 26 | 0 | 1 |
| 41 | 1 | 3 | 1 | 0 | 35 | 1 | 1 |
| 41 | 1 | 3 | 1 | 35 | 51 | 0 | 2 |

# Example Patients

| ID | rx | number | size | start | stop | event | enum |
|---|---|---|---|---|---|---|---|
| 5 | 1 | 4 | 1 | 0 | 6 | 1 | 1 |
| 5 | 1 | 4 | 1 | 6 | 10 | 0 | 2 |
| 13 | 1 | 3 | 1 | 0 | 3 | 1 | 1 |
| 13 | 1 | 3 | 1 | 3 | 9 | 1 | 2 |
| 13 | 1 | 3 | 1 | 9 | 21 | 1 | 3 |
| 13 | 1 | 3 | 1 | 21 | 23 | 0 | 4 |
| 16 | 1 | 1 | 2 | 0 | 26 | 0 | 1 |
| 41 | 1 | 3 | 1 | 0 | 35 | 1 | 1 |
| 41 | 1 | 3 | 1 | 35 | 51 | 0 | 2 |

# Individuals as "cluster" model

```
bladder.td <- coxph(Surv(start, stop, event == 1) ~ rx + number +
                    size + cluster(id), data = bladder)

summary(bladder.td)
```

# “Cluster” model

```
           coef    exp(coef) se(coef) robust.se  z     Pr(>|z|)
rx      -0.46469   0.62833  0.19973   0.26556 -1.750  0.08015 .
number   0.17496   1.19120  0.04707   0.06304  2.775  0.00551 **
size    -0.04366   0.95728  0.06905   0.07762 -0.563  0.57376
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

       exp(coef) exp(-coef) lower .95 upper .95
rx        0.6283     1.5915    0.3734     1.057
number    1.1912     0.8395    1.0527     1.348
size      0.9573     1.0446    0.8222     1.115

Concordance= 0.634  (se = 0.032 )
Likelihood ratio test= 17.52  on 3 df,   p=6e-04
Wald test            = 11.54  on 3 df,   p=0.009
Score (logrank) test = 19.52  on 3 df,   p=2e-04,   Robust = 11.27
p=0.01
```

# MODELS FOR REPEATED EVENTS

Conditional Model

# Conditional Models

- Unlike the previous model, we can preserve the ordering of events if it's important.
- In the **conditional model**, we stratify on the number of events, so only those who have had a previous event are in the risk set for the next one.
  - Example: Not in the risk set for the 3rd event until you have had the 2nd event.

- Each recurrence is a separate stratum (imagine own model) with its **own baseline hazard** – no estimates/inferences on the number of recurrences.

# Conditional Model – Risk Set

- Risk set for 1$^{st}$ event:

| ID | start | stop | event | enum |
|---|---|---|---|---|
| 5 | 0 | 6 | 1 | 1 |
| 13 | 0 | 3 | 1 | 1 |
| 16 | 0 | 26 | 0 | 1 |
| 41 | 0 | 35 | 1 | 1 |

- Risk set for 2$^{nd}$ event:

| ID | start | stop | event | enum |
|---|---|---|---|---|
| 5 | 6 | 10 | 0 | 2 |
| 13 | 3 | 9 | 1 | 2 |
| 41 | 35 | 51 | 0 | 2 |

# Conditional Model

- Keeps information about “stratification”
- Strata is based on the number of times the event occurs (stratum 1 is the time til “first event”, stratum 2 is time til “second event”…we called this enum in the data set)
- Need to be aware if a risk set becomes too small (for example only one patient had the event happen 4 times…everyone else had the event 3 or fewer times)

# Conditional Model

```
bladder.con <- coxph(Surv(start, stop, event == 1) ~ rx + number +
                     size + strata(enum)+cluster(id), data =
bladder)

summary(bladder.con)
```

# Conditional Model

```
          coef      exp(coef) se(coef) robust.se    z     Pr(>|z|)
rx      -0.333489  0.716420  0.216168  0.204787 -1.628   0.1034
number   0.119617  1.127065  0.053338  0.051387  2.328   0.0199 *
size    -0.008495  0.991541  0.072762  0.061635 -0.138   0.8904
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

        exp(coef) exp(-coef) lower .95 upper .95
rx         0.7164     1.3958    0.4796     1.070
number     1.1271     0.8873    1.0191     1.246
size       0.9915     1.0085    0.8787     1.119
```

# CAN ALSO DO SAME ANALYSIS STRATIFIED

Assume effects change across “strata”

# Stratified

```
bladder.con2 <- coxph(Surv(start, stop, event == 1) ~
strata(enum)*rx + strata(enum)*number + strata(enum)*size +
cluster(id), data = bladder)

summary(bladder.con2)
```

# Stratified

| | coef | exp(coef) | se(coef) | robust.se | z | Pr(>\|z\|) | |
|---|---|---|---|---|---|---|---|
| rx | -0.52598 | 0.59097 | 0.31583 | 0.31524 | -1.669 | 0.09521 | . |
| number | 0.23818 | 1.26894 | 0.07588 | 0.07459 | 3.193 | 0.00141 | ** |
| size | 0.06961 | 1.07209 | 0.10156 | 0.08863 | 0.785 | 0.43220 | |
| strata(enum)enum=2:rx | 0.02215 | 1.02239 | 0.51451 | 0.60852 | 0.036 | 0.97097 | |
| strata(enum)enum=3:rx | 0.66664 | 1.94768 | 0.74348 | 0.57671 | 1.156 | 0.24771 | |
| strata(enum)enum=4:rx | 0.57632 | 1.77947 | 0.85238 | 0.62678 | 0.919 | 0.35784 | |
| strata(enum)enum=2:number | -0.26282 | 0.76888 | 0.11763 | 0.16532 | -1.590 | 0.11189 | |
| strata(enum)enum=3:number | -0.18852 | 0.82819 | 0.20026 | 0.14196 | -1.328 | 0.18420 | |
| strata(enum)enum=4:number | -0.03390 | 0.96667 | 0.25366 | 0.19351 | -0.175 | 0.86092 | |
| strata(enum)enum=2:size | -0.23033 | 0.79427 | 0.15910 | 0.17506 | -1.316 | 0.18827 | |
| strata(enum)enum=3:size | 0.09849 | 1.10350 | 0.28757 | 0.18033 | 0.546 | 0.58497 | |
| strata(enum)enum=4:size | -0.06052 | 0.94128 | 0.35382 | 0.37643 | -0.161 | 0.87228 | |

# GAP TIME

# Gap Time

- Notice that in the conditional model, each event's **start** time is determined by the previous event's **stop** time!
- An alternative time scale is the **gap time**, where we instead choose to model the time *since last event*.
- In gap-time models, time is reset to 0 after each event, so the time until the prior event has no bearing on the current event's risk set.

# Conditional time versus Gap Time?

- "The conditional probability approach uses the time since the beginning of the study to define the time intervals, and is appropriate when the interest is in the full course of the recurrent event process. The gap time approach essentially "resets the clock" for each recurrence by using the time since the previous event to define time intervals, and is more appropriate when event (or recurrence)-specific effect estimates are of interest."

Source: https://www.publichealth.columbia.edu/research/population-health-methods/time-event-data-analysis

# Gap Time – Risk Set

- Risk set for 1st event:

| ID | start | stop | event | enum |
|---|---|---|---|---|
| 5 | 0 | 6 | 1 | 1 |
| 13 | 0 | 3 | 1 | 1 |
| 16 | 0 | 26 | 0 | 1 |
| 41 | 0 | 35 | 1 | 1 |

- Risk set for 2nd event:

| ID | start | stop | event | enum |
|---|---|---|---|---|
| 5 | 0 | 4 | 0 | 2 |
| 13 | 0 | 6 | 1 | 2 |
| 41 | 0 | 16 | 0 | 2 |

# Gap Time – R

```
bladder.gap <- coxph(Surv(time = (stop - start), event == 1) ~ rx +
                     number + size + strata(enum)+cluster(id), data
= bladder)

summary(bladder.gap)
```

# Gap Time – R

```
          coef     exp(coef)  se(coef) robust.se  z     Pr(>|z|)
rx      -0.279005  0.756536  0.207348  0.215624 -1.294  0.19569
number   0.158046  1.171220  0.051942  0.050940  3.103  0.00192
size     0.007415  1.007443  0.070023  0.064333  0.115  0.90824
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

       exp(coef) exp(-coef) lower .95 upper .95
rx        0.7565     1.3218    0.4958     1.154
number    1.1712     0.8538    1.0599     1.294
size      1.0074     0.9926    0.8881     1.143

Concordance= 0.596  (se = 0.032 )
Likelihood ratio test= 9.33  on 3 df,   p=0.03
Wald test            = 11.84  on 3 df,   p=0.008
Score (logrank) test = 10.27  on 3 df,   p=0.02,   Robust = 9.92
p=0.02
```