An Overview of Competing Risk Analysis in Time-to-Event Outcomes Using SAS

CYP-C Research Champion Webinar Series

JUNE 26, 2019

Presented by: Giancarlo Di Giuseppe, MPH
Research Associate & Analyst
Pediatric Oncology Group of Ontario
gdigiuseppe@pogo.ca

Outline

- Refresher and overview of time-to-event analysis
- What is a competing risk? Why do we need to consider them?
- Assumptions of competing risk analysis
- Data structure
- Cumulative Incidence Function
- Sub-distributional Hazard – Fine and Gray method
- Cause-specific Hazard – alternative method

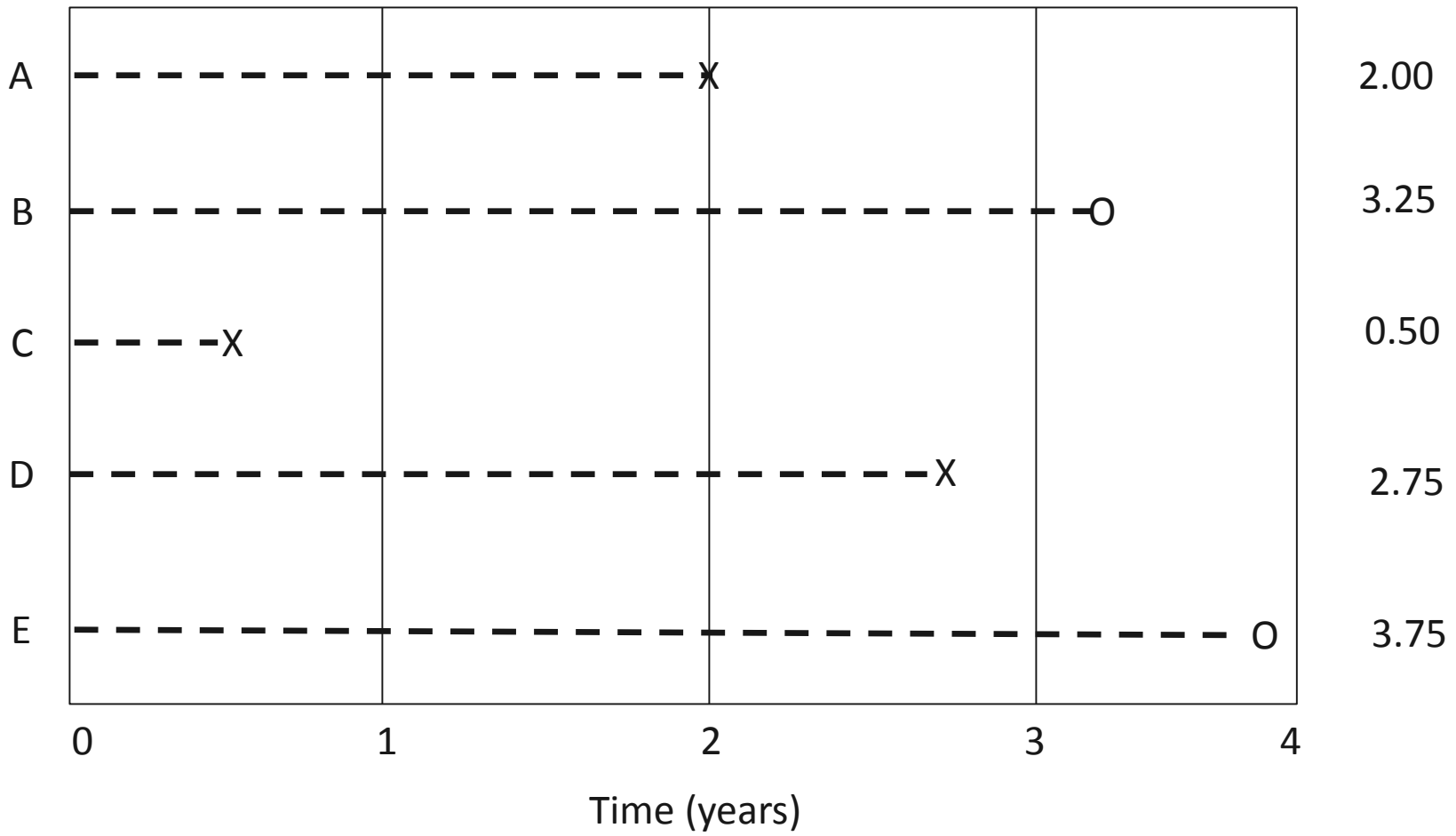
Time-to-event analysis refresher

- Synonymous with survival analysis

- Models the occurrence and **timing** of an outcome of interest
  - Origin of observation window (t_0) varies by research objective

- Censoring of individuals being followed describes periods of no observation
  - Left, right, and interval

- Reason for censoring may vary (critical for competing risk analysis) for individuals and depends on the research objectives
  - Examples include: lost-to-follow up, outcome of interest, end of study observation, death, etc.


A
B
C
D
E
X
O
X
X
O
2.00
3.25
0.50
2.75
3.75
0
1
2
3
4
Time (years)

What is a Competing Risk?

- Competing risks are said to be present when a patient is at risk of more than one mutually exclusive event, where the occurrence of one event will prevent any other from happening

- An individual can experience a failure event from one of several possible causes, with one failure cause precluding the others

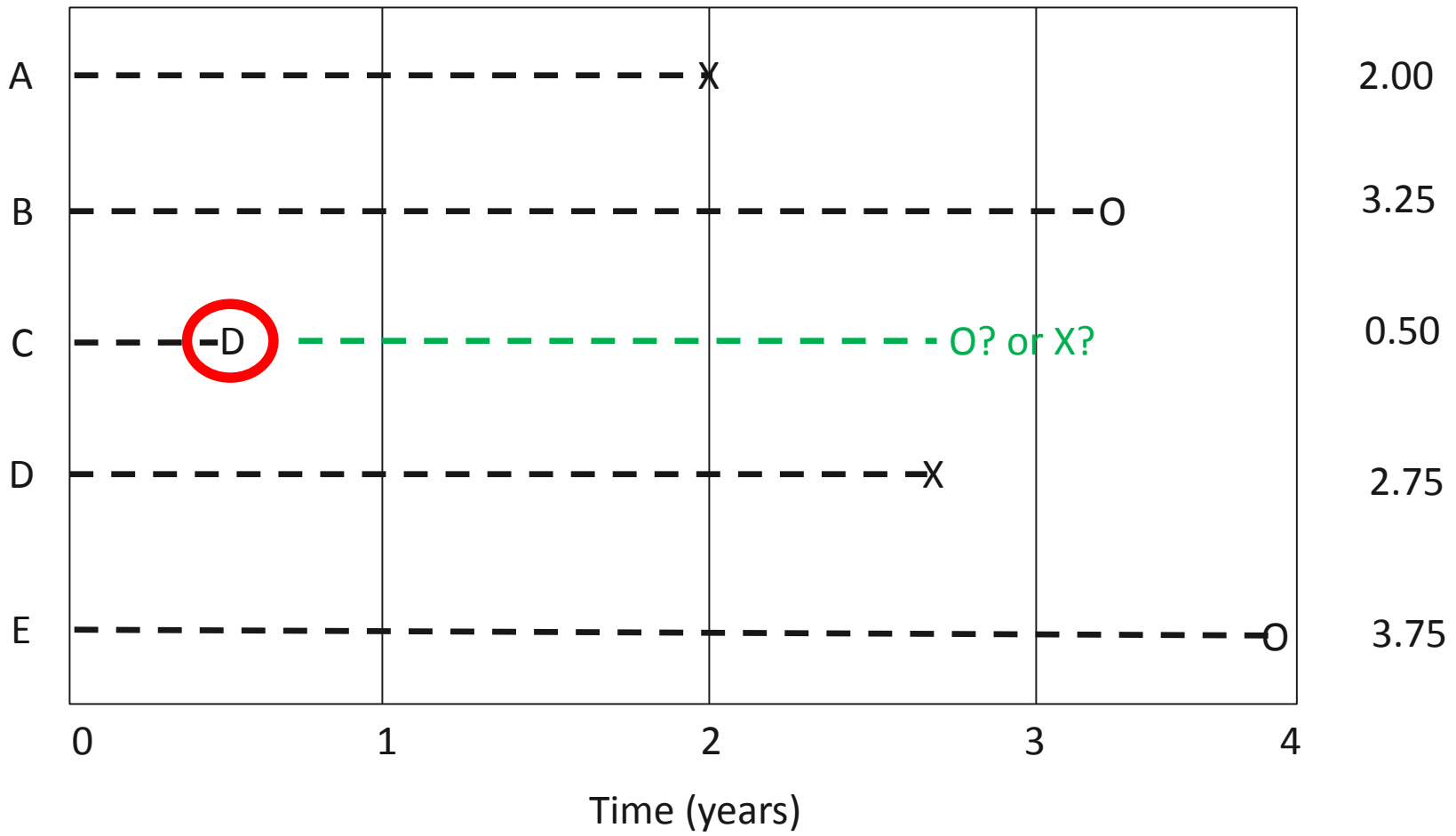
- Examples: All-cause mortality (can be a comp risk for anything, really); treatment-related mortality, progressive disease, or relapse in BMT studies

When & Why?

- Traditional survival analyses tend to focus on failure-time data that have a single type of failure

- Competing risks should be considered when the occurrence of one event hinders the occurrence of other types of events from ever happening (i.e. death)

- Competing risk analysis allows us to model separate survival probabilities for events in the presence of competing events


A
B
C
D
E
X
O
D
O? or X?
X
O
2.00
3.25
0.50
2.75
3.75
0
1
2
3
4
Time (years)

Assumptions of Survival Analysis

- All assumptions of traditional survival analysis apply to competing risks
- Have to assume that the reason for censoring are **independent** and reasonable
  - No way of testing independence assumption
- Censoring is assumed to be: random & non-informative
- Individuals have the same future risk of the event of interest as individuals who have not been censored and have not had the event of interest

Hypothetical Competing Risk Study – an example

Mock Study to Understand Comp Risk

- Interested in studying the effect of treatment received for a **primary** cancer has on the development of a *subsequent malignant neoplasm* (SMN)
- Intensity of Treatment Rating Scale (ITR-3) is a composite measure of the treatment received for paediatric cancer protocols
- Patients are followed from their initial diagnosis date of the primary cancer (t_0) to the development of an SMN or when the study ends (December 31, 2016)
- Death must be considered a competing risk

POGO
PEDIATRIC ONCOLOGY GROUP OF ONTARIO

Kazak AE, *et al.* A revision of the intensity of treatment rating scale: classifying the intensity of pediatric cancer treatment. Pediatr Blood Cancer. 2012;59(1):96–99. doi:10.1002/pbc.23320

Data Structure

```
DATA T2; SET T;
  /* DEFINING MY COHORT */
  IF 1985 <= DX1_YEAR <= 2012;
  IF 0 < DX1_AGE < 15;
  IF ITR IN (1:4);

  /* MAKING SURE MY EVENTS OF INTEREST HAPPEN WITHIN THE OBSERVATION WINDOW */
  *  DEATHS;
  IF . < DEATH_DATE <= '31DEC2016'D THEN DO; DEATH = 1; END;
    ELSE DO; DEATH = 0; DEATH_DATE = .; END;

  *  SMN'S;
  IF . < DX_DATE2 <= '31DEC2016'D THEN DO; SMN = 1; END;
    ELSE DO; SMN = 0; DX_DATE2 = .; END;

  LABEL TIME_DEATH = "NO. DAYS BETWEEN DX1 DATE AND DEATH";
  IF DEATH = 1 THEN DO;
    TIME_DEATH = DEATH_DATE - DX_DATE1;
    IF TIME_DEATH < 0 THEN TIME_DEATH = 0; /*POST-MORTEM DEATHS TO DAY ZERO */
  END;

  LABEL TIME_SMN = "NO. DAYS BETWEEN DX1 DATE AND DX2 DATE";
  TIME_SMN = DX_DATE2 - DX_DATE1;
```

Ensures no outcomes occur outside of the observation window

Calculating time between dates of interest

POGO
PEDIATRIC ONCOLOGY GROUP OF ONTARIO

Data Structure – contd.

Creates a censor date variable to be used to calculate FU time

```
 /* CENSORED ON THE EARLIEST OF: SMN DX, DEATH, OR DEC 31 2016 */
FORMAT CENSOR_DATE DATE9.;
CENSOR_DATE = MIN(DX_DATE2, DEATH_DATE, '31DEC2016'D);
```

Defines the competing risk and reason for censoring

```
/* DEFINES MY CENSOR VARIABLE WHERE EXITS ARE DUE TO:
  1 = SMN DIAGNOSIS
  2 = DEATH (FROM ANY CAUSE)
  0 = NO OUTCOME EVENT EXPERIENCED AND CENSORED AT STUDY END */
LABEL CENS_CMPRSK = "CENSOR VARIABLE STATUS FOR CMP RSK";
IF SMN = 1 AND (DX_DATE2 < DEATH_DATE OR DEATH_DATE = .) THEN CENS_CMPRSK = 1;
  ELSE IF DEATH = 1 AND SMN = 0 THEN CENS_CMPRSK = 2;
ELSE CENS_CMPRSK = 0;

/* FOLLOW-UP TIME VARIABLE USING THE DEFINED CENSOR DATE */
LABEL CENS_TIME = "CENSOR TIME (IN DAYS)";
CENS_TIME = CENSOR_DATE - DX_DATE1;

/* CREATING A VARIABLE WHICH DOES NOT CAPTURE DEATH AS A REASON FOR EXIT */
IF CENS_CMPRSK = 1 THEN STATUS = 1; ELSE STATUS = 0;


RUN; *N = 9,659;
```

POGO
PEDIATRIC ONCOLOGY GROUP OF ONTARIO

Data Structure

SUBJECT_ID	DEATH	SMN	ITR	TIME_DEATH	TIME_SMN	CENS_CMPRSK	CENS_TIME	STATUS
1	0. NO	0. NO	2. MODERATELY INTENSIVE			0	6815	0
2	0. NO	0. NO	2. MODERATELY INTENSIVE			0	7031	0
3	1. YES	0. NO	4. MOST INTENSIVE	107		2	107	0
26	1. YES	1. YES	4. MOST INTENSIVE	2453	2094	1	2094	1
35	1. YES	0. NO	3. VERY INTENSIVE	206		2	206	0

Cumulative Incidence

Cumulative Incidence Function (CIF)

- Cumulative incidence is the probability that an event of interest occurs before a given time *t*

- In competing risk analysis, the CIF is the cumulative probability of failure from a specific cause over time accounting for the fact that patients can fail from other causes (the competing risk)

- Recall: Cumulative incidence is equal to 1 – survival probability when only right censoring is present

Cumulative Incidence Function (CIF)

- CIF can easily be calculated in SAS 9.4

```
/* CIF METHOD WITH GRAY'S TEST OF EQUALITY */
PROC LIFETEST DATA=T3 NOTABLE
        OUTCIF=CR_CIF_OUTPUT
        PLOTS=CIF(TEST);
        TIME CENS_TIME*CENS_CMPRSK(0) / EVENTCODE=1;
        STRATA ITR;
RUN;
```

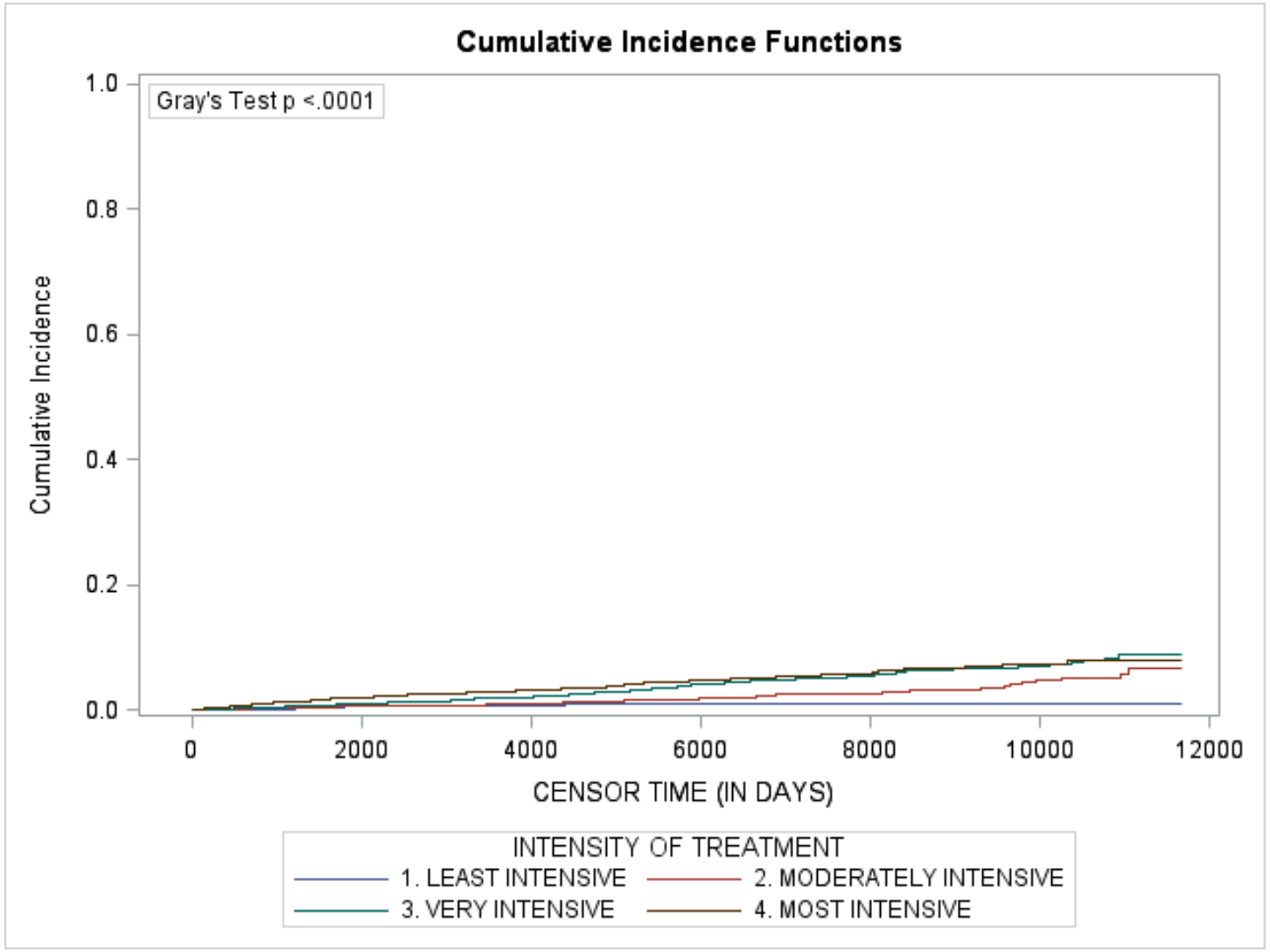
Grouping variable

Competing Risk code = 1 is our outcome (SMN) of interest

CIF in older versions of SAS

- Previous versions of SAS have a CIF macro built-in using the %CIF function

```
%CIF (
        DATA = T3,
        TIME = CENS_TIME,
        STATUS = CENS_CMPRSK,
        EVENT = 1,
        CENSORED = 0,
        GROUP = ITR
);
```



Gray's Test for Equality of Cumulative Incidence Functions		
Chi-Square	DF	Pr > Chi-Square
52.9431	3	<.0001

Kaplan-Meier – an overestimation in CR

- Primary assumption in Kaplan-Meier is that individuals who are censored have the same survival probability as those who continue to be followed – violated in competing risk analysis

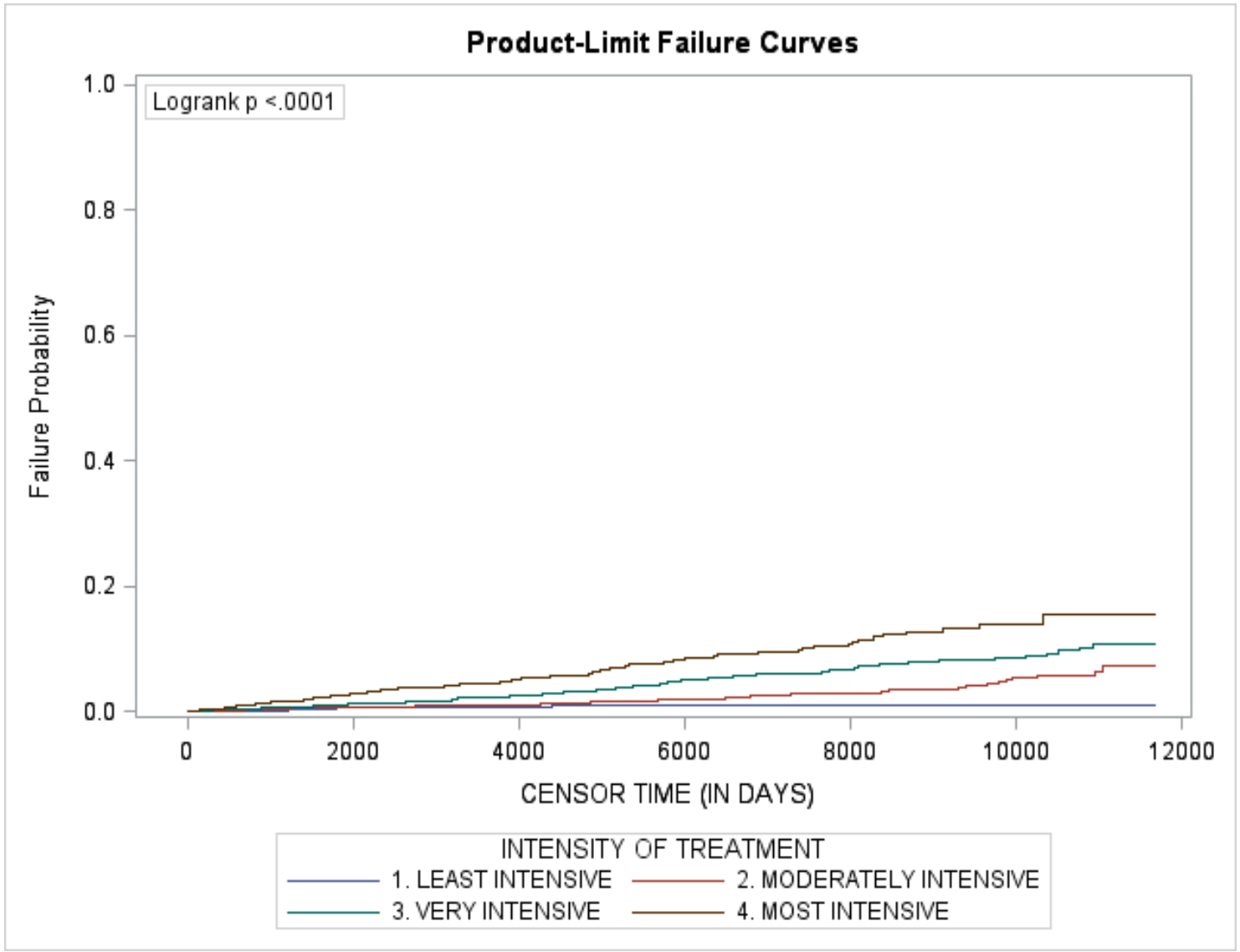
- Biased due to the fact that the probability of event occurrence is modified (aka conditional) by an antecedent competing event

- Traditional KM curves will result in biased and overestimated results in the presence of competing risks

```
/* STANDARD KAPLAN-MEIER METHOD */
PROC LIFETEST DATA=T3 NOTABLE
        OUTSURV=KM_OUTPUT
        PLOTS=SURVIVAL(FAILURE NOCENSOR TEST);
        TIME CENS_TIME*STATUS(0);
        STRATA ITR;
RUN;
```

POGO
PEDIATRIC ONCOLOGY GROUP OF ONTARIO

Kaplan-Meier – an overestimation in CR



Test of Equality over Strata			
Test	Chi-Square	DF	Pr > Chi-Square
Log-Rank	118.1122	3	<.0001
Wilcoxon	104.9924	3	<.0001
-2Log(LR)	114.4263	3	<.0001

KM vs. CIF

Summary of the Number of Censored and Uncensored Values					
Stratum	ITR	Total	Failed	Censored	Percent Censored
1	1. LEAST INTENSIVE	1022	9	1013	99.12
2	2. MODERATELY INTENSIVE	3269	72	3197	97.80
3	3. VERY INTENSIVE	2995	132	2863	95.59
4	4. MOST INTENSIVE	2370	110	2260	95.36
Total		9656	323	9333	96.65

Summary of Failure Outcomes					
Stratum	ITR	Failed Events	Competing Events	Censored	Total
1	1. LEAST INTENSIVE	9	36	977	1022
2	2. MODERATELY INTENSIVE	72	312	2887	3271
3	3. VERY INTENSIVE	132	585	2279	2996
4	4. MOST INTENSIVE	110	1204	1056	2370
Total		323	2137	7199	9659

Hazard Function: Sub-distribution

Sub-distribution Hazards – Fine and Gray

- Fine and Gray (1999) proposed a proportional hazards model aimed at examining the effects of covariates in the context of competing risks
- Uses the cumulative incidence function to model **sub-distribution** hazards
- Risk set contains subjects who are currently event free, as well as those who have previously experienced a competing event
- Sub-distribution hazard subjects who are censored from the competing risk remain in the risk set and are given a weight which reduces with censoring time

Fine JP, Gray RJ. A Proportional Hazards Model for the Subdistribution of a Competing Risk. J Am Stat Assoc. 1999; 94(446): 496-509

Available in SAS 9.4

```
PROC PHREG DATA=T3;
  CLASS    ITR (REF='1. LEAST INTENSIVE')
           DX_GROUP1 (REF="4. SOLID TUMOR + OTHER")
           RAD (REF="0. NO")
           /PARAM=REFERENCE;

  MODEL CENS_TIME*CENS_CMPRSK(0) = ITR DX1_AGE DX_GROUP1 RAD / RL EVENTCODE=1;

RUN;
```

If using older version of SAS, use %PSHREG

- %PSHREG macro for older versions of SAS and will perform the Fine and Gray modelling regression

More information can be found here:

https://cemsiis.meduniwien.ac.at/kb/wf/software/statistische-software/pshreg/

Kohl M, et al. PSHREG: a SAS macro for proportional and nonproportional subdistribution hazards regression. Comput Methods Programs Biomed. 2015;118(2):218-33. doi: 10.1016/j.cmpb.2014.11.009.

The PHREG Procedure

Model Information

Data Set	WORK.T3	
Dependent Variable	CENS_TIME	CENSOR TIME (IN DAYS)
Status Variable	CENS_CMPRSK	CENSOR VARIABLE STATUS FOR CMP RSK
Event of Interest	1	CENSOR VARIABLE STATUS FOR CMP RSK
Competing Event	2	CENSOR VARIABLE STATUS FOR CMP RSK
Censored Value	0	CENSOR VARIABLE STATUS FOR CMP RSK

Summary of Failure Outcomes

Total	Event of Interest	Competing Event	Censored
9656	323	2134	7199

Type 3 Tests

Effect	DF	Wald Chi-Square	Pr > ChiSq
ITR	3	35.8569	<.0001
DX1_AGE	1	0.1630	0.6864
DX_GROUP1	3	38.4740	<.0001
RAD	1	16.4573	<.0001

POGO
PEDIATRIC ONCOLOGY GROUP OF ONTARIO

Analysis of Maximum Likelihood Estimates

Parameter		DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio	95% Hazard Ratio Confidence Limits	
ITR	2. MODERATELY INTENSIVE	1	0.76416	0.36525	4.3772	0.0364	2.147	1.049	4.393
ITR	3. VERY INTENSIVE	1	1.44259	0.36168	15.9089	<.0001	4.232	2.083	8.597
ITR	4. MOST INTENSIVE	1	1.58799	0.36911	18.5087	<.0001	4.894	2.374	10.089
DX1_AGE		1	0.00542	0.01342	0.1630	0.6864	1.005	0.979	1.032
DX_GROUP1	1. LEUKEMIA	1	-0.31836	0.14976	4.5193	0.0335	0.727	0.542	0.975
DX_GROUP1	2. LYMPHOMA	1	0.74805	0.17719	17.8223	<.0001	2.113	1.493	2.990
DX_GROUP1	3. CNS	1	0.00366	0.16430	0.0005	0.9822	1.004	0.727	1.385
RAD	1. YES	1	0.49547	0.12213	16.4573	<.0001	1.641	1.292	2.085

What about the competing risk?

- Can quantify the instantaneous hazard for the competing risk in our cohort by changing the **event code** of interest

- Same interpretation as the previous output – “the hazard of death in the presence of a SMN diagnosis as a competing risk”

```
PROC PHREG DATA=T3;
  CLASS    ITR (REF='1. LEAST INTENSIVE')
           DX_GROUP1 (REF="4. SOLID TUMOR + OTHER")
           RAD (REF="0. NO")
           /PARAM=REFERENCE;

  MODEL CENS_TIME*CENS_CMPRSK(0) = ITR DX1_AGE DX_GROUP1 RAD / RL EVENTCODE=2;

RUN;
```

POGO
PEDIATRIC ONCOLOGY GROUP OF ONTARIO

Summary of Failure Outcomes

Total	Event of Interest	Competing Event	Censored
9656	2134	323	7199

Type 3 Tests

Effect	DF	Wald Chi-Square	Pr > ChiSq
ITR	3	1099.9723	<.0001
DX1_AGE	1	4.1000	0.0429
DX_GROUP1	3	191.9531	<.0001
RAD	1	7.0828	0.0078

Analysis of Maximum Likelihood Estimates

Parameter		DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio	95% Hazard Ratio Confidence Limits	
ITR	2. MODERATELY INTENSIVE	1	1.08001	0.18202	35.2047	<.0001	2.945	2.061	4.207
ITR	3. VERY INTENSIVE	1	1.73951	0.17817	95.3230	<.0001	5.695	4.016	8.074
ITR	4. MOST INTENSIVE	1	2.89518	0.17602	270.5382	<.0001	18.087	12.810	25.538
DX1_AGE		1	0.01035	0.00511	4.1000	0.0429	1.010	1.000	1.021
DX_GROUP1	1. LEUKEMIA	1	-0.40460	0.05247	59.4659	<.0001	0.667	0.602	0.740
DX_GROUP1	2. LYMPHOMA	1	-0.33679	0.08890	14.3525	0.0002	0.714	0.600	0.850
DX_GROUP1	3. CNS	1	0.33409	0.05536	36.4231	<.0001	1.397	1.253	1.557
RAD	1. YES	1	0.12538	0.04711	7.0828	0.0078	1.134	1.034	1.243

Not distinguishing the event type

```
PROC PHREG DATA=T3;
  CLASS    ITR (REF='1. LEAST INTENSIVE')
           DX_GROUP1 (REF="4. SOLID TUMOR + OTHER")
           RAD (REF="0. NO")
             /PARAM=REFERENCE;
  MODEL CENS_TIME*CENS_CMPRSK(0) = ITR DX1_AGE DX_GROUP1 RAD / RL;
RUN;
```

Summary of the Number of Event and Censored Values

Total	Event	Censored	Percent Censored
9656	2457	7199	74.55

Type 3 Tests

Effect	DF	Wald Chi-Square	Pr > ChiSq
ITR	3	1217.8407	<.0001
DX1_AGE	1	5.3649	0.0205
DX_GROUP1	3	180.9814	<.0001
RAD	1	16.0137	<.0001

Analysis of Maximum Likelihood Estimates

Parameter		DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio	95% Hazard Ratio Confidence Limits	
ITR	2. MODERATELY INTENSIVE	1	1.05563	0.16060	43.2049	<.0001	2.874	2.098	3.937
ITR	3. VERY INTENSIVE	1	1.72749	0.15705	120.9900	<.0001	5.627	4.136	7.655
ITR	4. MOST INTENSIVE	1	2.82598	0.15610	327.7487	<.0001	16.877	12.429	22.918
DX1_AGE		1	0.01086	0.00469	5.3649	0.0205	1.011	1.002	1.020
DX_GROUP1	1. LEUKEMIA	1	-0.42269	0.05190	66.3297	<.0001	0.655	0.592	0.725
DX_GROUP1	2. LYMPHOMA	1	-0.12759	0.07868	2.6296	0.1049	0.880	0.754	1.027
DX_GROUP1	3. CNS	1	0.29993	0.05311	31.8885	<.0001	1.350	1.216	1.498
RAD	1. YES	1	0.17387	0.04345	16.0137	<.0001	1.190	1.093	1.296

Answers a different question!

Hazard Function: Cause-specific

Cause-specific Hazard – alternative method

- Standard Cox regression modelling strategy with competing events treated as censored observations
- Instantaneous risk from a specific event after censoring for the competing risk and conditional on survival until time *t* or later
- Risk set decreases with time after individuals are censored for the competing event
- Note: Cause-specific hazards do not allow us to examine the effects of covariates on the CIF → this is what led Fine and Gray to develop their regression method

Cause-specific Hazard – alternative method

```
PROC PHREG DATA=T3;
  CLASS ITR (REF='1. LEAST INTENSIVE')
       DX_GROUP1 (REF="4. SOLID TUMOR + OTHER")
       RAD (REF="0. NO")
       /PARAM=REFERENCE;

  MODEL CENS_TIME*CENS_CMPRSK(0,2) = ITR DX1_AGE DX_GROUP1 RAD / RL;
RUN;
```

Cause-specific Hazard – alternative method

The PHREG Procedure

Model Information

Data Set	WORK.T3	
Dependent Variable	CENS_TIME	CENSOR TIME (IN I
Censoring Variable	CENS_CMPRSK	CENSOR VARIABLE S
Censoring Value(s)	0 2	
Ties Handling	BRESLOW	

Summary of the Number of Event and Censored Values

Total	Event	Censored	Percent Censored
9656	323	9333	96.65

Type 3 Tests

Effect	DF	Wald Chi-Square	Pr > ChiSq
ITR	3	87.1208	<.0001
DX1_AGE	1	0.6325	0.4265
DX_GROUP1	3	40.2795	<.0001
RAD	1	17.5673	<.0001

POGO
PEDIATRIC ONCOLOGY GROUP OF ONTARIO

Cause-specific Hazard – alternative method

Analysis of Maximum Likelihood Estimates

Parameter		DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio	95% Hazard Ratio Confidence Limits	
ITR	2. MODERATELY INTENSIVE	1	0.86775	0.36352	5.6981	0.0170	2.382	1.168	4.856
ITR	3. VERY INTENSIVE	1	1.59356	0.35736	19.8845	<.0001	4.921	2.443	9.914
ITR	4. MOST INTENSIVE	1	2.22457	0.36217	37.7282	<.0001	9.250	4.548	18.810
DX1_AGE		1	0.01045	0.01314	0.6325	0.4265	1.011	0.985	1.037
DX_GROUP1	1. LEUKEMIA	1	-0.49215	0.15011	10.7497	0.0010	0.611	0.456	0.820
DX_GROUP1	2. LYMPHOMA	1	0.61476	0.17464	12.3912	0.0004	1.849	1.313	2.604
DX_GROUP1	3. CNS	1	0.08382	0.16157	0.2691	0.6039	1.087	0.792	1.493
RAD	1. YES	1	0.50711	0.12099	17.5673	<.0001	1.660	1.310	2.105

POGO
PEDIATRIC ONCOLOGY GROUP OF ONTARIO

Do we unintentionally model competing risk?!

- Recall the dichotomized variable STATUS: where 1=SMN diagnosis & 0=censored

- Death's were captured, but contained within the composite censor value of '0'

```
*      COMPOSITE EVENT CAPTURED IN THE STATUS VARIABLE;
PROC PHREG DATA=T3;
  CLASS ITR (REF='1. LEAST INTENSIVE')
        DX_GROUP1 (REF="4. SOLID TUMOR + OTHER")
        RAD (REF="0. NO")/PARAM=REFERENCE;

  MODEL CENS_TIME*STATUS(0) = ITR DX1_AGE DX_GROUP1 RAD / RL;
RUN;
```

POGO
PEDIATRIC ONCOLOGY GROUP OF ONTARIO

Summary of the Number of Event and Censored Values

Total	Event	Censored	Percent Censored
9656	323	9333	96.65

Type 3 Tests

Effect	DF	Wald Chi-Square	Pr > ChiSq
ITR	3	87.1208	<.0001
DX1_AGE	1	0.6325	0.4265
DX_GROUP1	3	40.2795	<.0001
RAD	1	17.5673	<.0001

Analysis of Maximum Likelihood Estimates

Parameter		DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio	95% Hazard Ratio Confidence Limits	
ITR	2. MODERATELY INTENSIVE	1	0.86775	0.36352	5.6981	0.0170	2.382	1.168	4.856
ITR	3. VERY INTENSIVE	1	1.59356	0.35736	19.8845	<.0001	4.921	2.443	9.914
ITR	4. MOST INTENSIVE	1	2.22457	0.36217	37.7282	<.0001	9.250	4.548	18.810
DX1_AGE		1	0.01045	0.01314	0.6325	0.4265	1.011	0.985	1.037
DX_GROUP1	1. LEUKEMIA	1	-0.49215	0.15011	10.7497	0.0010	0.611	0.456	0.820
DX_GROUP1	2. LYMPHOMA	1	0.61476	0.17464	12.3912	0.0004	1.849	1.313	2.604
DX_GROUP1	3. CNS	1	0.08382	0.16157	0.2691	0.6039	1.087	0.792	1.493
RAD	1. YES	1	0.50711	0.12099	17.5673	<.0001	1.660	1.310	2.105

Sub-distribution vs. Cause-specific hazard

- Differences in the hazards are due to the underlying **risk set** used
- When the competing risk is common, cause-specific hazards will overestimate the hazard
- Degree of overestimation depends on the frequency and distribution of competing events

Andersen PK, et al. Competing risks in epidemiology: possibilities and pitfalls. Int J Epidemiol. 2012;41(3):861-70.

Summary

- Competing risk analysis is considered when subject is at risk of more than one mutually exclusive outcome event
- Models separate survival probabilities for outcome of interest in the presence of competing events
- Analysis is easily performed in SAS with slight modifications to the PROC LIFETEST and PROC PHREG procedures
- There are two methods to perform competing risk analysis in SAS: sub-distributional hazards or cause-specific hazards

Additional Readings

Andersen PK, et al. Competing risks in epidemiology: possibilities and pitfalls. Int J Epidemiol. 2012;41(3):861-70.

Dignam JJ, et al. The use and interpretation of competing risks regression models. Clin Cancer Res. 2012;18(8):2301-8.

Fine JP, Gray RJ. A Proportional Hazards Model for the Subdistribution of a Competing Risk. J Am Stat Assoc. 1999; 94(446): 496-509

Pintilie M. (2006), Competing Risks: A Practical Perspective, Statistics in Practice, Chichester, UK: John Wiley & Sons

Questions?

Email: gdigiuseppe@pogo.ca