%MVMODELS: a Macro for Survival and Logistic Analysis

Jeffrey Meyers, Mayo Clinic:

Statistical Programmer Analyst III within Mayo Clinic's Cancer Center Statistics team.

Research focus on gastrointestinal and respiratory cancers as well as meta-analysis.

10 years of SAS programming experience focusing on macros, graphics, SQL and reports.

%MVMODELS: a Macro for Survival and Logistic Analysis

What is %MVMODELS

- SAS macro to perform either survival analysis or logistic regression and output results to a forest plot or table
- Capable of performing univariate or multivariate modeling
- Multiple options and features for subgrouping models
- Flexible macro parameters to customize the output
- Outputs the same to RTF, HTML, PDF, EXCEL, and PowerPoint
- Contains error checking, documentation, and cleans up after itself

Sample Dataset EXAMPLE

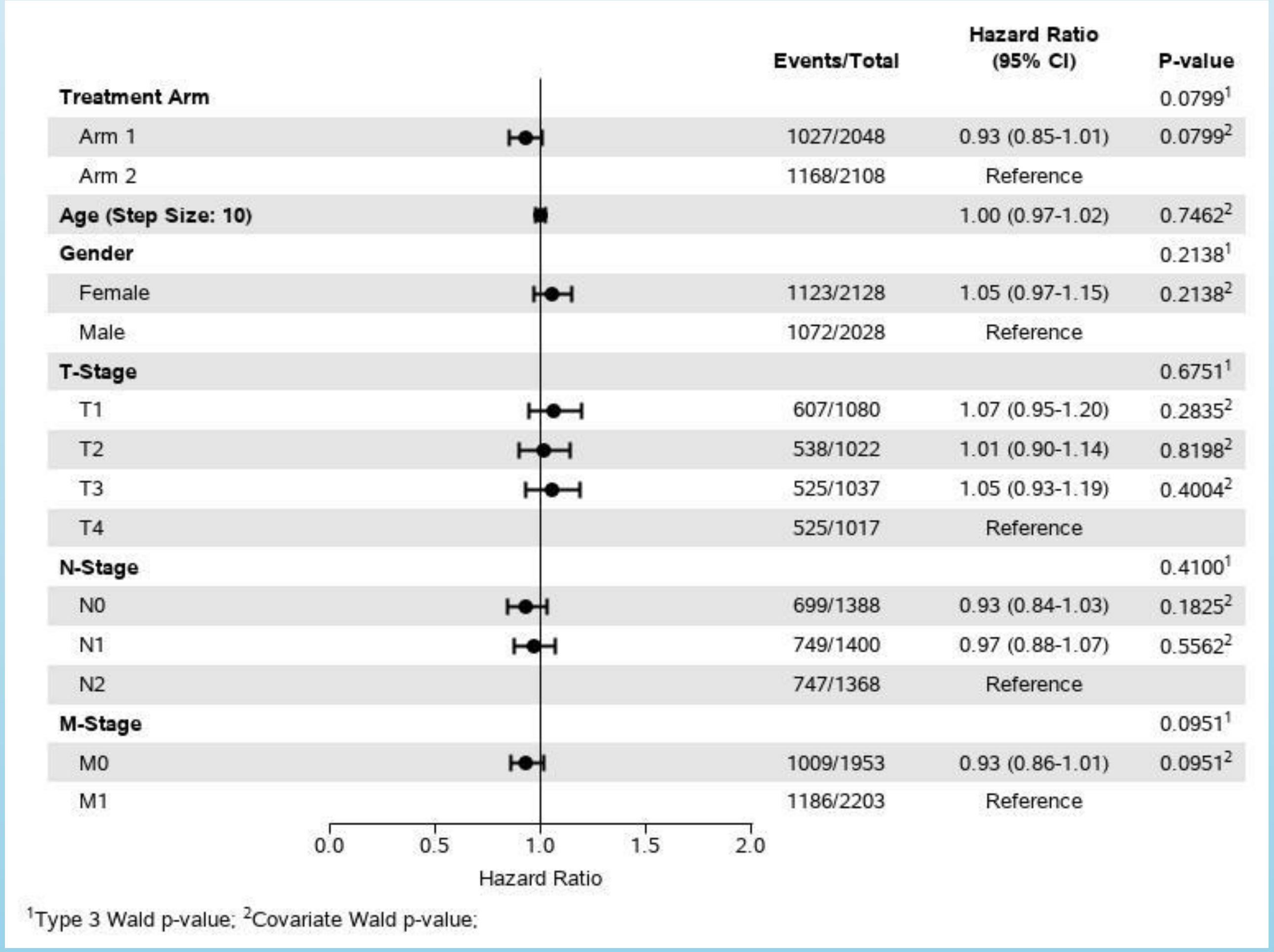
- Randomly created "Meta-analysis" dataset of 5 studies
- Includes variables for arm, baseline characteristics (age, gender, TNM staging), tumor response status, and overall survival time and status
- Code to produce available in paper
- Not a realistic dataset but will work for examples

Sample Dataset RANDOM

Study Number	Treatment Arm	Age	Gender	T-Stage	N-Stage	M-Stage	Response Status	Overall Survival Time (months)	Overall Survival Status
1	Arm 1	22	Female	T2	N1	M0	No Response	118	0
1	Arm 2	74	Male	T3	N0	M1	Response	12	0
1	Arm 1	78	Female	T2	N1	M0	Response	22	2
1	Arm 2	35	Female	T3	N0	M1	No Response	99	2
1	Arm 2	71	Female	T2	N0	M1	Response	72	1
1	Arm 2	34	Female	T1	N2	M1	No Response	120	3
1	Arm 1	48	Male	T2	N0	M1	Response	60	2
1	Arm 1	64	Female	T1	N1	M1	Response	82	2
1	Arm 1	34	Male	T4	N2	M1	Response	50	2
1	Arm 2	75	Female	T3	N0	M1	No Response	67	3
1	Arm 1	35	Male	T1	N1	M0	No Response	34	1
1	Arm 2	32	Male	T4	N2	M1	No Response	112	2
1	Arm 2	64	Male	T1	N1	M0	No Response	42	3
1	Arm 2	58	Male	T2	N0	M0	No Response	1	0
1	Arm 2	77	Male	T1	N2	M0	No Response	88	1
1	Arm 2	25	Female	T2	N1	M1	Response	25	1

Example Forest Plot

```
%MVMODELS(
    DATA=random,
    METHOD=survival,
    TIME=os_time,
     CENS=os_stat,
     COVARIATES=arm age gender tstage
                       nstage mstage,
    TYPE=2 1 2,
    CAT_DISPLAY=4,
    CONT_STEP=10,
    INCREMENT=0.5,
    REFLINE=1);
```



Example Table

```
%MVMODELS(
    DATA=random,
    METHOD=survival,
    TIME=os_time,
     CENS=os_stat,
     COVARIATES=arm age gender tstage
                   nstage mstage,
    TYPE=2 1 2,
    CAT_DISPLAY=4,
    CONT_STEP=10,
    INCREMENT=0.5,
    REFLINE=1);
```

	Events/Total	Hazard Ratio (95% CI)	P-value
Treatment Arm			0.0799[1]
Arm 1	1027/2048	0.93 (0.85-1.01)	0.0799[2]
Arm 2	1168/2108	Reference	
Age (Step Size: 10)		1.00 (0.97-1.02)	0.7462[2]
Gender			0.2138[1]
Female	1123/2128	1.05 (0.97-1.15)	0.2138[2]
Male	1072/2028	Reference	
T-Stage			0.6751[1]
T1	607/1080	1.07 (0.95-1.20)	0.2835[2]
T2	538/1022	1.01 (0.90-1.14)	0.8198[2]
T3	525/1037	1.05 (0.93-1.19)	0.4004[2]
T4	525/1017	Reference	
N-Stage			0.4100[1]
N0	699/1388	0.93 (0.84-1.03)	0.1825[2]
N1	749/1400	0.97 (0.88-1.07)	0.5562[2]
N2	747/1368	Reference	
M-Stage			0.0951[1]
M0	1009/1953	0.93 (0.86-1.01)	0.0951[2]
M1	1186/2203	Reference	

[1]Type 3 Wald p-value;[2]Covariate Wald p-value;

Required Parameters

- DATA: specifies user input dataset
- NMODELS: number of models being run. Default=1
- METHOD: determines if SURVIVAL or LOGISTIC models will run
- PLOT_DISPLAY: determines which output plots/statistics are shown and which order in the forest plot
- TABLE_DISPLAY: determines which output statistics are shown and which order in the table
- Default for DISPLAY parameters is STANDARD which changes depending on what method and options are chosen

DISPLAY Parameters

- PLOT_DISPLAY and TABLE_DISPLAY have keywords based on prefixes and suffixes
  - Prefixes: MED_, KM_, HR_, C_, BIN_, and OR_ stand for median, Kaplan-Meier, Hazard Ratio, C-index, Binomial success rate, and Odds Ratio
  - Suffixes: ESTIMATE, RANGE, EST_RANGE, PLOT stand for estimate, range (lower limit - upper limit), estimate (range), and a plot
- Other keywords include: Subtitle (row headers), TOTAL, EVENTS, EV_T (Events/Total), PCT (percentage Events/Total), EV_T_PCT (Events/Total (%)), and PVAL
- Additionally REF_ is a prefix that can be used on statistics to be shown specifically for reference groups (e.g. REF_EV_T)

Modeling Key Parameters

- COVARIATES: space delimited list of independent covariates for either survival analysis or logistic regression
  - Not required. Can still compute other statistic such as median time-to-event without covariates
- TYPE: indicates whether covariates are continuous (TYPE=1) or categorical (TYPE=2)
  - Space delimited list where first type will match with first COVARIATE, second type with second COVARIATE, etc.
  - Last TYPE listed will carry forward if more COVARIATES exist.
    - Example: TYPE=1 2 2 2 will be the same as TYPE=1 2 for four covariates

Continuous Covariate Parameters

- CONT_STEP: determines the numeric step size that a hazard ratio or odds ratio constitutes
  - E.g. CONT_STEP=10 for age will give the hazard ratio for a 10 step increase in age
- CONT_DISPLAY: determines how continuous covariates are displayed

	Hazard Ratio (95% CI)	P-value
CONT_DISPLAY: 1 and CONT_STEP=1		
Age (Step Size: 1)	1.00 (1.00-1.00)	0.7650[1]
CONT_DISPLAY: 2 and CONT_STEP=10		
Age	1.00 (0.97-1.02)	0.7650[1]
CONT_DISPLAY: 3 and CONT_STEP=100		
Age		0.7650[2]
Step Size: 100	0.97 (0.77-1.22)	0.7650[1]

[1]Covariate Wald p-value;[2]Type 3 Wald p-value;

Categorical Covariate Parameters

- CAT_ORDER: changes the order categorical values are displayed.
  - Default is unformatted values
  - Example: CAT_ORDER=2 3 1 will cause the default 2nd value to be first, default 3rd value to be second and the default 1st value to be 3rd
  - ` (lowercase ~) is used as a delimiter between multiple categorical covariates
- CAT_REF: determines the reference value for hazard and odds ratios
  - Formatted value should be listed without quotes
  - ` (lowercase ~) is used as a delimiter between multiple categorical covariates
- CAT_DISPLAY: determines how categorical covariates are displayed
  - Some statistics can only be shown on the covariate level if CAT_DISPLAY=4

	Events/Total	Hazard Ratio (95% CI)	P-value
CAT_DISPLAY: 1			
Gender (Female vs Male)		1.06 (0.97-1.15)	0.2050[1]
CAT_DISPLAY: 2			
Gender (Reference: Male)			0.2050[2]
Female		1.06 (0.97-1.15)	0.2050[1]
CAT_DISPLAY: 3			
Gender			0.2050[2]
Female		1.06 (0.97-1.15)	0.2050[1]
CAT_DISPLAY: 4			
Gender			0.2050[2]
Female	1123/2128	1.06 (0.97-1.15)	0.2050[1]
Male	1072/2028	Reference	
CAT_DISPLAY: 5			
Gender			0.2050[2]
Female vs Male		1.06 (0.97-1.15)	0.2050[1]

[1]Covariate Wald p-value;[2]Type 3 Wald p-value;

Survival Method Key Parameters

- TIME: numeric time-to-event variable (Required)
- CENS: numeric event indicator. (Required)
  - CEN_VL: determines the censor value (Default=0)
  - EV_VL: determines event of interest when SURV_METHOD=CIF (Default=missing)
- SURV_METHOD: determines if survival methods will be Kaplan-Meier (KM), 1-Kaplan-Meier (1-KM), or cumulative incidence (CIF). Default=KM
- TIMELIST: specifies time-points for event-free rates
  - TDIVISOR: transforms TIME variable into other units
  - e.g. TDIVISOR=12 will divide time by 12 transforming months to years

Survival Method Available Statistics

KM and 1-KM Methods	CIF Methods
Number patients and events	Number patients and events
Median time-to-event	Median time-to-event
Time-point event-free rates	Time-point event-free rates
Hazard ratios	Hazard ratios
P-values (type-3 tests, logrank, wilcoxon, Wald)	P-values (type-3 tests, Gray k-sample test, Wald)
Concordance index	
Fit statistics	

Survival Method Procedures Used

KM and 1-KM Methods

- LIFETEST Procedure
  - Number patients and events
  - Median time-to-event
  - Time-point event-free rates
  - Logrank/Wilcoxon p-values
- PHREG Procedure
  - Hazard ratios
  - P-values (type-3 tests, Wald)
  - Fit statistics
- Manual programming[1]
  - Concordance index

CIF Methods

- LIFETEST Procedure
  - Number patients and events
  - Median time-to-event
  - Time-point event-free rates
  - Gray k-sample test p-value
- PHREG Procedure
  - Hazard ratios
  - P-values (type-3 tests, Wald)

[1] Therneau T (2014). _A Package for Survival Analysis in S_. R package version 2.37-7, <URL: http://CRAN.R-project.org/package=survival>.

Survival Method Available Plots

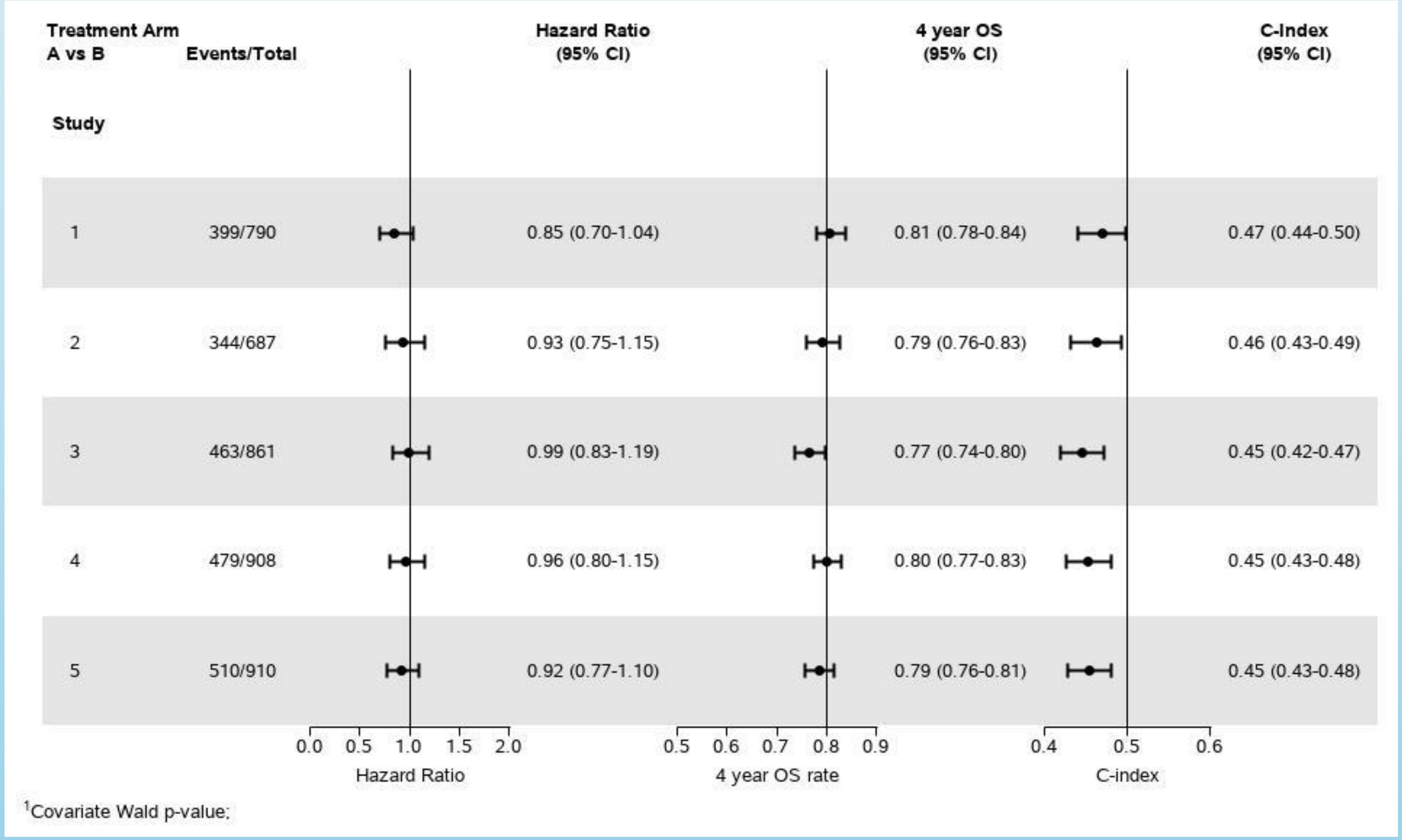
KM and 1-KM Methods

- Median time-to-event
- Time-point event-free rates
- Hazard ratios
- Concordance index

CIF Methods

- Median time-to-event
- Time-point event-free rates
- Hazard ratios

Survival Method Example Plot



Logistic Method Key Parameters

- EVENTCOV: determines the binomial event variable. This is the dependent covariate in logistic regression models

- EVENT: determines the event to be considered the “success” in binomial analyses and the event to be modeled in logistic regression models

Logistic Method Available Statistics

- Number patients and events
- Binomial success rates for designated event
- Odds ratios
- P-values (type-3 Wald test, Wald, Chi-square, Fisher’s exact)
- Concordance index (non-stratified models)
- Fit statistics

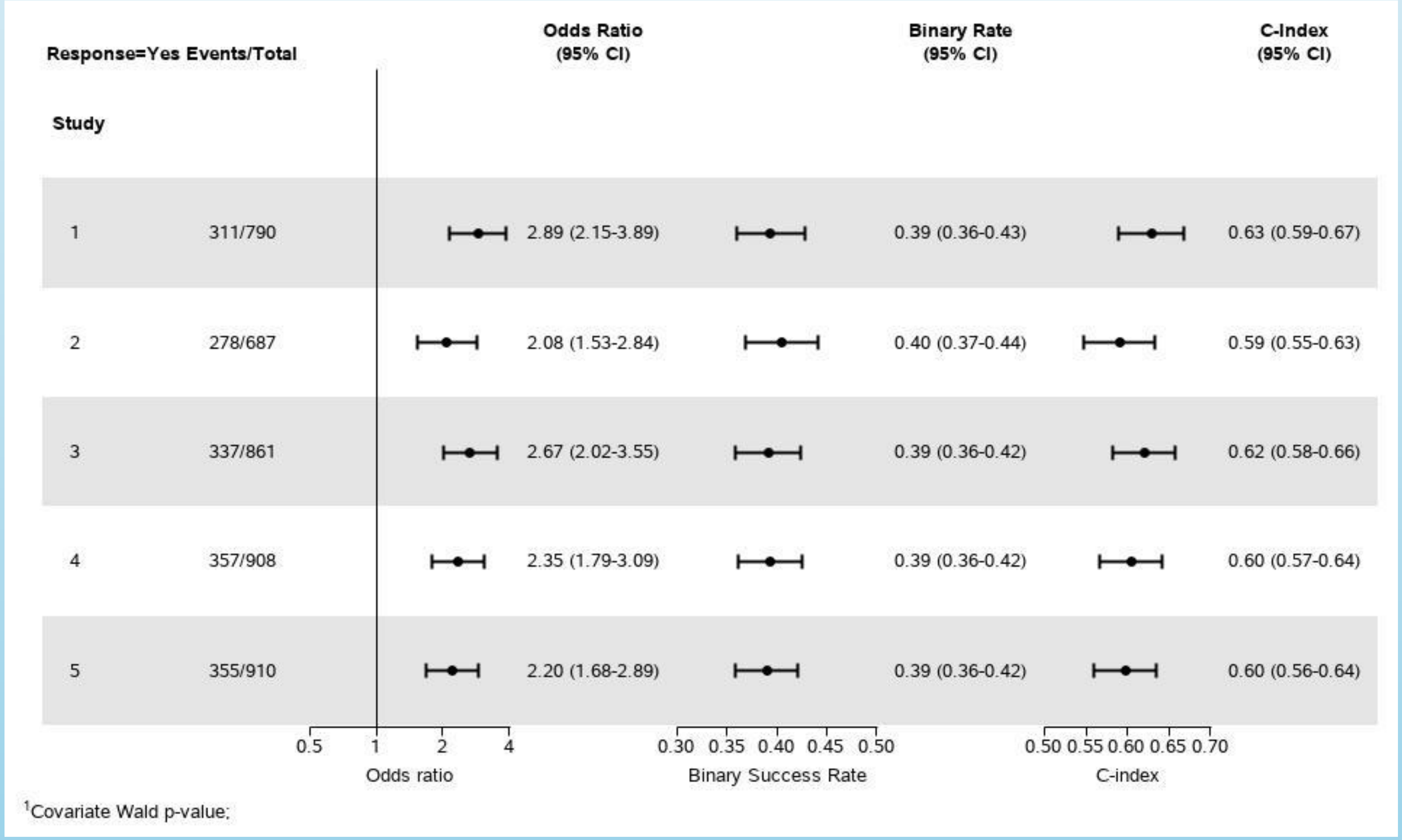
Logistic Method Procedures Used

- FREQ Procedure
  - Number patients and events
  - Binomial success rates for designated event
  - Chi-square and Fisher's exact p-value
- Logistic Procedure
  - Odds ratios
  - P-values (type-3 Wald test, Wald)
  - Fit statistics
- Manual Programming
  - Concordance index (non-stratified models)
  - cNeil BJ: The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology 143:29-36, 1982.

Logistic Method Available Plots

- Binomial success rates for designated event
- Odds ratios
- Concordance index (non-stratified models)

Logistic Method Example Plot

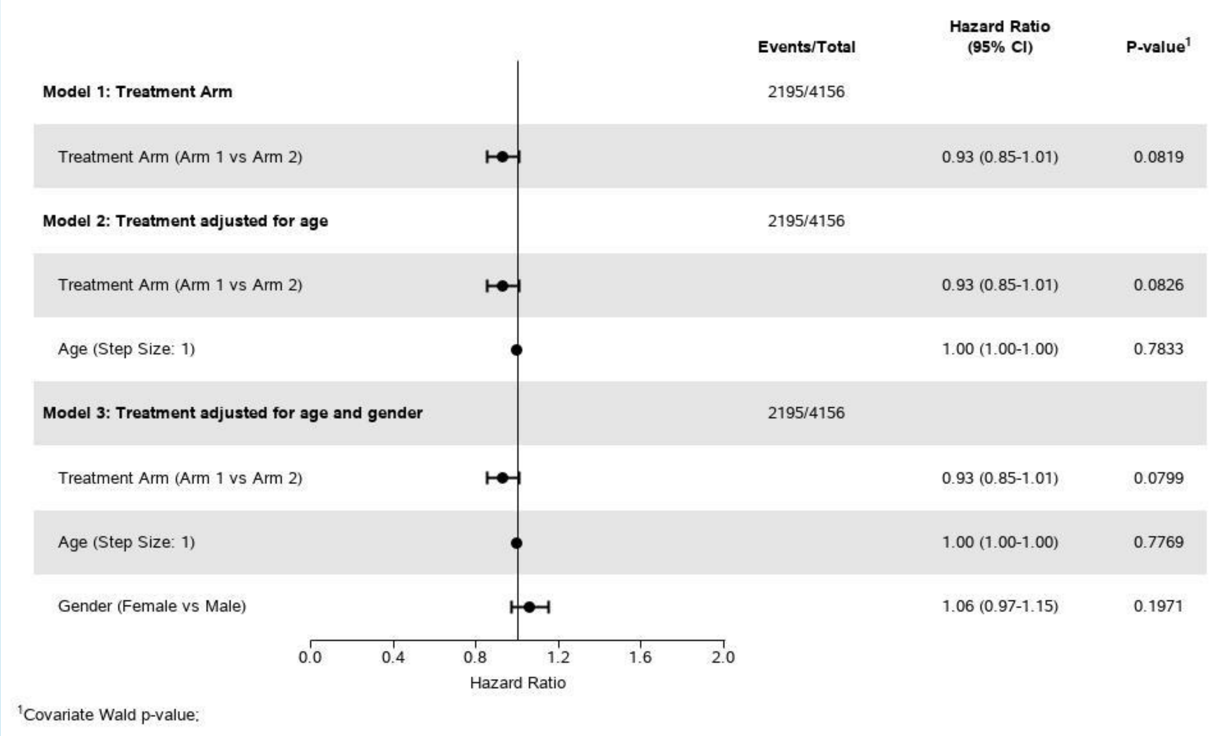


Running Multiple Models

- Two primary methods to run multiple models: manual or subgroup variables
- Manual:
  - NMODELS: determines how many models will be run
  - | (pipe) symbol changes options for each model
    - Example: COVARIATES=age gender sex|age gender will use age, gender and sex for model 1 and only age and gender for model 2
    - Any options without a | symbol will be used for all models
      - TIME=os_time without any | will use os_time for every model
  - Most control, but also more complicated macro call
  - WHERE: allows a where clause to be applied to an individual model

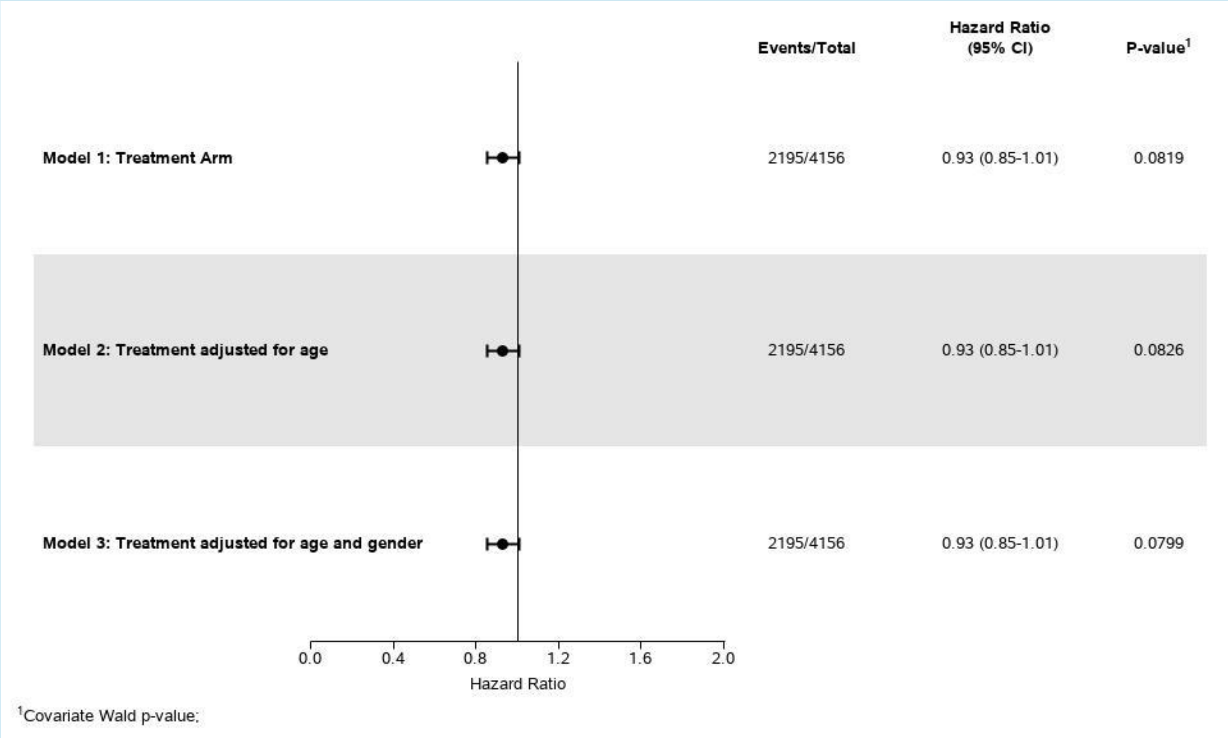
Running Multiple Models

```
%mvmodels(
    DATA=random, NMODELS=3,
    METHOD=survival,
    TIME=os_time, CENS=os_stat,
  COVARIATES=arm|arm age|arm age gender,
  TYPE=2|2 1|2 1 2,
  REFLINE=1,
  MODEL_TITLE=Model 1: Treatment Arm|
   Model 2: Treatment adjusted for age|
   Model 3: Treatment adjusted for age and gender,
  MODEL_TITLE_OWNROW=1,
  BOLD_COV_LABEL=0);
```



Running Multiple Models

```
%mvmodels(
    DATA=random, NMODELS=3,
    METHOD=survival,
    TIME=os_time, CENS=os_stat,
  COVARIATES=arm|arm age|arm age gender,
  TYPE=2|2 1|2 1 2,
   REFLINE=1,
  MODEL_TITLE=Model 1: Treatment Arm|
   Model 2: Treatment adjusted for age|
   Model 3: Treatment adjusted for age and gender,
  MODEL_TITLE_OWNROW=0,
  SHOW_ADJCOVARIATES=0);
```



Running Multiple Models

- MODEL_TITLE: gives each model its own title
  - If only one covariate is listed it will merge its output row with the title row unless MODEL_TITLE_OWNROW=1

- SHOW_ADJCOVARIATES: determines if any covariate beyond the first is shown in the plot/table.
  - Condenses table if only intending in showing how covariate of interest changes when adjusting for other factors

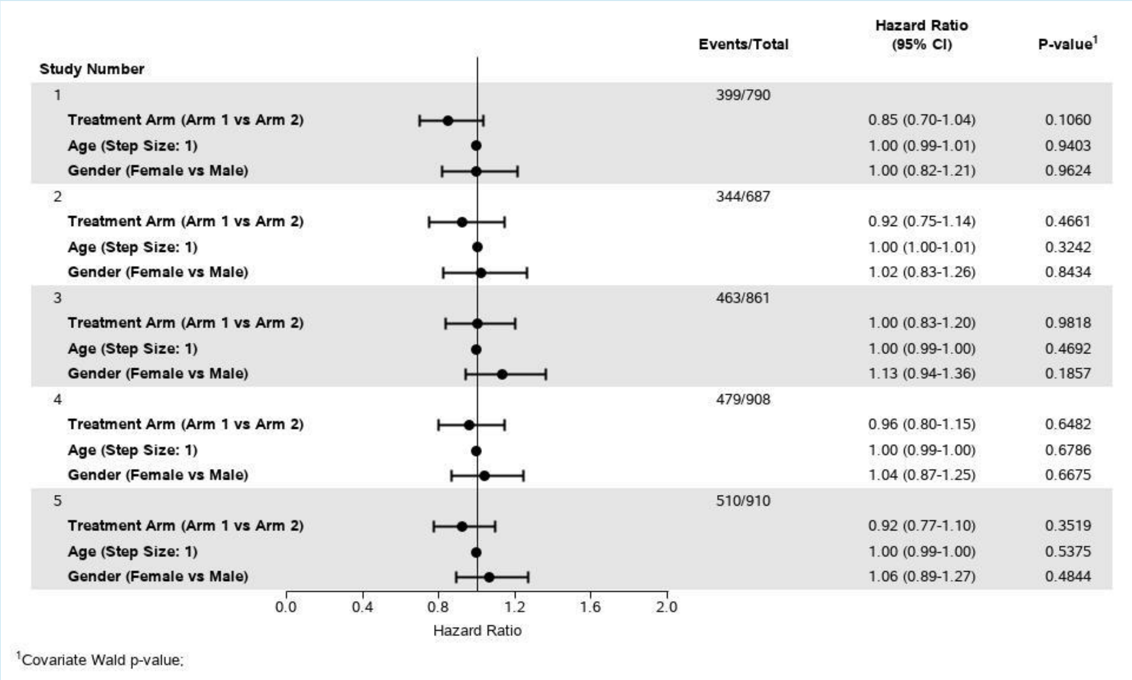
- BOLD_ and INDENT_ options: allows the user to control the boldface and indent of different output types
  - Example: BOLD_COV_LABEL controls the boldface of covariate label rows
  - Example: INDENT_MODEL_TITLE controls indent of model titles

Running Multiple Models

- Subgroup variables:
  - BY: specifies one or more variables to duplicate the current model across
  - COLBY (Column BY): specifies one variable to duplicate all models and subgroups in multiple columns
  - ROWBY(Row BY): specifies one variable to duplicate all models and subgroups in multiple rows
  - GROUPBY (Group BY): specifies one variable to duplicate all models such that all by levels are grouped into the same row with different colors

Running Multiple Models

```
%MVMODELS(
   DATA=random,
   METHOD=survival,
   TIME=os_time, CENS=os_stat,
   NMODELS=1,
   BY=study, SHADING=2,
   COVARIATES=arm age gender,
   TYPE=2 1 2,
   REFLINE=1);
```



Running Multiple Models

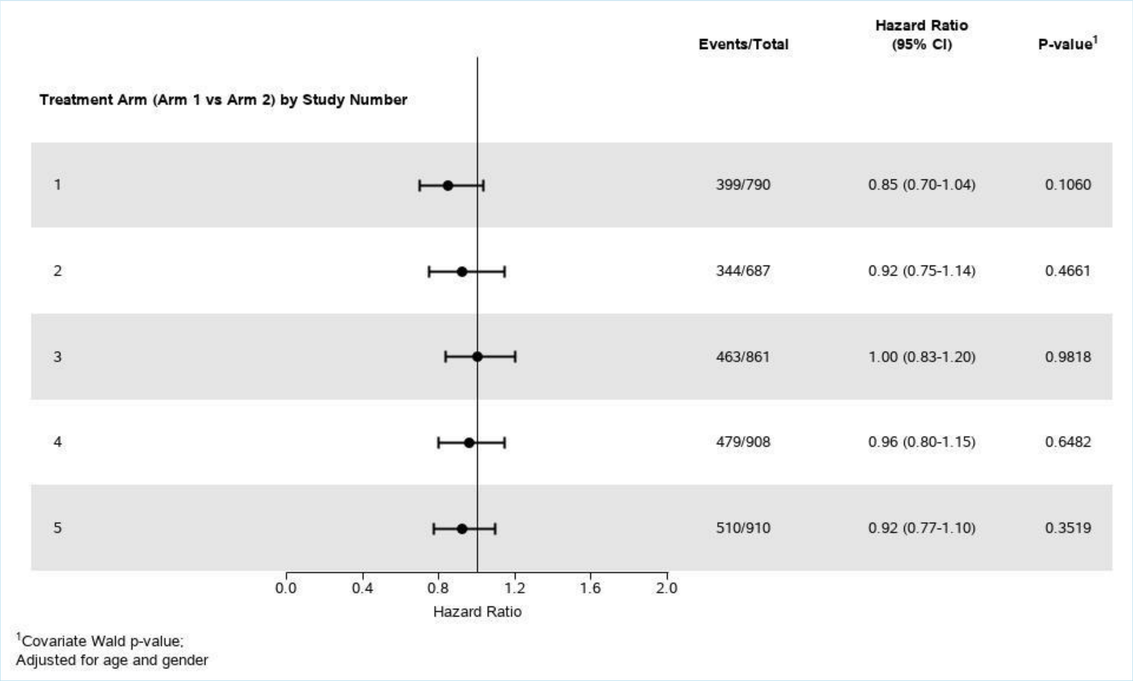
```
%MVMODELS(
  DATA=random,
  METHOD=survival,
  TIME=os_time, CENS=os_stat,
  NMODELS=1,
  BY=study, SHADING=2,
  COVARIATES=arm age gender,
  TYPE=2 1 2,
  REFLINE=1);
```

	Events/Total	Hazard Ratio (95% CI)	P-value[1]
Study Number			
1	399/790		
Treatment Arm (Arm 1 vs Arm 2)		0.85 (0.70-1.04)	0.1060
Age (Step Size: 1)		1.00 (0.99-1.01)	0.9403
Gender (Female vs Male)		1.00 (0.82-1.21)	0.9624
2	344/687		
Treatment Arm (Arm 1 vs Arm 2)		0.92 (0.75-1.14)	0.4661
Age (Step Size: 1)		1.00 (1.00-1.01)	0.3242
Gender (Female vs Male)		1.02 (0.83-1.26)	0.8434
3	463/861		
Treatment Arm (Arm 1 vs Arm 2)		1.00 (0.83-1.20)	0.9818
Age (Step Size: 1)		1.00 (0.99-1.00)	0.4692
Gender (Female vs Male)		1.13 (0.94-1.36)	0.1857
4	479/908		
Treatment Arm (Arm 1 vs Arm 2)		0.96 (0.80-1.15)	0.6482
Age (Step Size: 1)		1.00 (0.99-1.00)	0.6786
Gender (Female vs Male)		1.04 (0.87-1.25)	0.6675
5	510/910		
Treatment Arm (Arm 1 vs Arm 2)		0.92 (0.77-1.10)	0.3519
Age (Step Size: 1)		1.00 (0.99-1.00)	0.5375
Gender (Female vs Male)		1.06 (0.89-1.27)	0.4844

[1]Covariate Wald p-value;

Running Multiple Models

```
%MVMODELS(
   DATA=random,
   METHOD=survival,
   TIME=os_time, CENS=os_stat,
   NMODELS=1,
   BY=study, SHADING=2,
   COVARIATES=arm age gender,
   TYPE=2 1 2,
   REFLINE=1,
   SHOW_ADJCOVARIATES=0,
   FOOTNOTE=Adjusted for age
and gender);
```

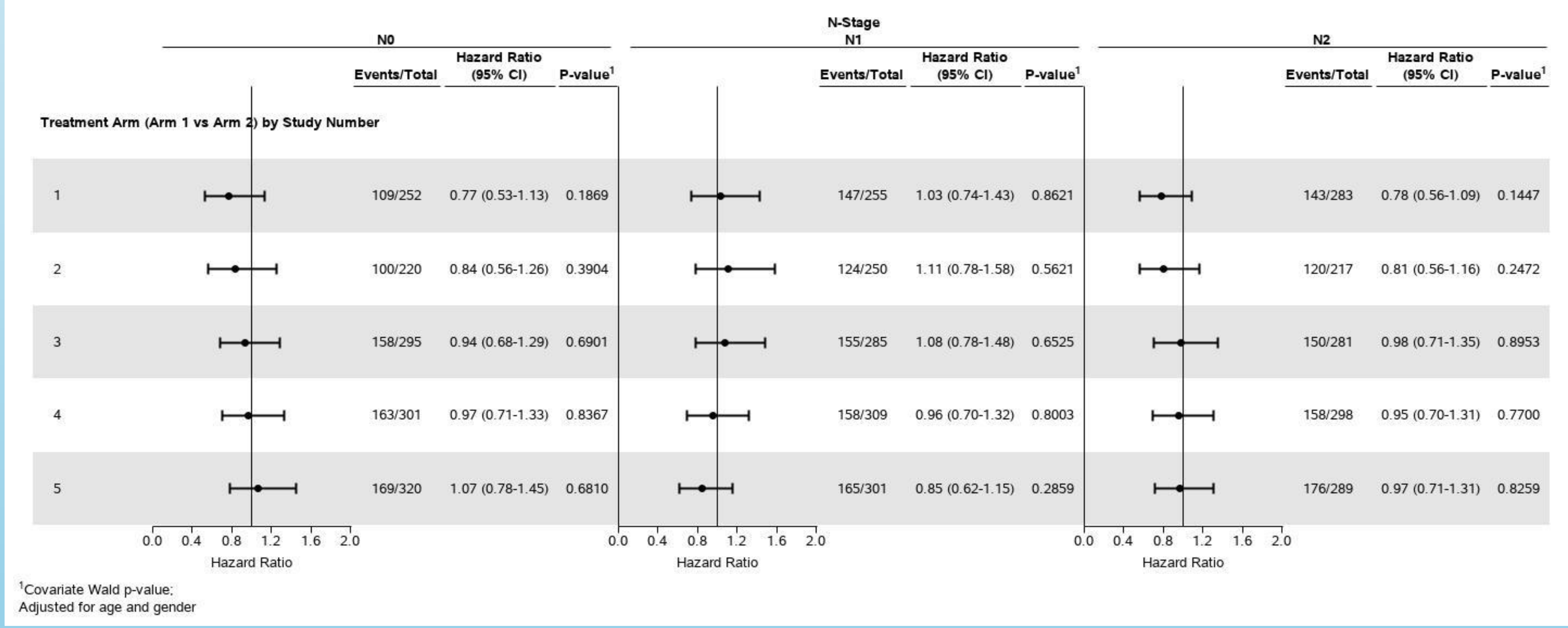


Running Multiple Models

%MVMODELS(

…,

COLBY=nstage, UNDERLINEHEADERS=1);



Treatment Arm (Arm 1 vs Arm 2) by Study Number	N0 Events/Total	N0 Hazard Ratio (95% CI)	N0 P-value[1]	N1 Events/Total	N1 Hazard Ratio (95% CI)	N1 P-value[1]	N2 Events/Total	N2 Hazard Ratio (95% CI)	N2 P-value[1]
1	109/252	0.77 (0.53-1.13)	0.1869	147/255	1.03 (0.74-1.43)	0.8621	143/283	0.78 (0.56-1.09)	0.1447
2	100/220	0.84 (0.56-1.26)	0.3904	124/250	1.11 (0.78-1.58)	0.5621	120/217	0.81 (0.56-1.16)	0.2472
3	158/295	0.94 (0.68-1.29)	0.6901	155/285	1.08 (0.78-1.48)	0.6525	150/281	0.98 (0.71-1.35)	0.8953
4	163/301	0.97 (0.71-1.33)	0.8367	158/309	0.96 (0.70-1.32)	0.8003	158/298	0.95 (0.70-1.31)	0.7700
5	169/320	1.07 (0.78-1.45)	0.6810	165/301	0.85 (0.62-1.15)	0.2859	176/289	0.97 (0.71-1.31)	0.8259

[1]Covariate Wald p-value;
Adjusted for age and gender

Running Multiple Models

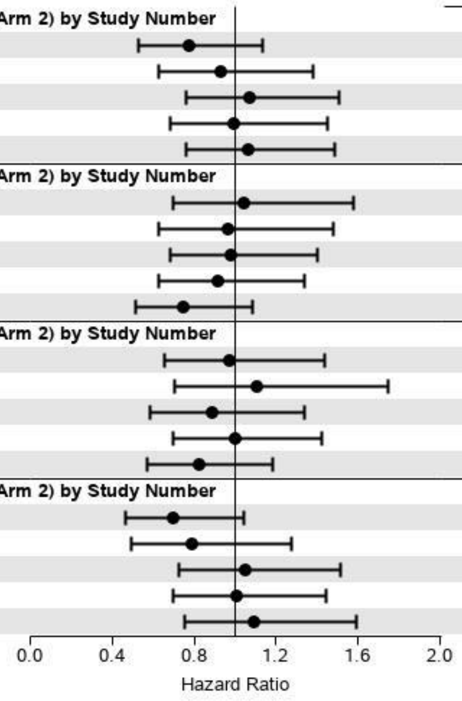
- COLBY works well with EXCEL destination due to its horizontal space

	N-Stage								
	N0			N1			N2		
	Events/Total	Hazard Ratio (95% CI)	P-value[1]	Events/Total	Hazard Ratio (95% CI)	P-value[1]	Events/Total	Hazard Ratio (95% CI)	P-value[1]
Treatment Arm (Arm 1 vs Arm 2) by Study Number									
1	109/252	0.77 (0.53-1.13)	0.1869	147/255	1.03 (0.74-1.43)	0.8621	143/283	0.78 (0.56-1.09)	0.1447
2	100/220	0.84 (0.56-1.26)	0.3904	124/250	1.11 (0.78-1.58)	0.5621	120/217	0.81 (0.56-1.16)	0.2472
3	158/295	0.94 (0.68-1.29)	0.6901	155/285	1.08 (0.78-1.48)	0.6525	150/281	0.98 (0.71-1.35)	0.8953
4	163/301	0.97 (0.71-1.33)	0.8367	158/309	0.96 (0.70-1.32)	0.8003	158/298	0.95 (0.70-1.31)	0.77
5	169/320	1.07 (0.78-1.45)	0.681	165/301	0.85 (0.62-1.15)	0.2859	176/289	0.97 (0.71-1.31)	0.8259

[1]Covariate Wald p-value;
Adjusted for age and gender

Running Multiple Models

%MVMODELS(
…, UNDERLINEHEADERS=1,
ROWBY=tstage);



T-Stage	Treatment Arm (Arm 1 vs Arm 2) by Study Number	Events/Total	Hazard Ratio (95% CI)	P-value[1]
T1	1	108/203	0.78 (0.53-1.13)	0.1899
	2	103/181	0.93 (0.62-1.38)	0.7123
	3	134/234	1.07 (0.76-1.51)	0.7009
	4	115/225	1.00 (0.68-1.45)	0.9813
	5	147/237	1.06 (0.76-1.48)	0.7127
T2	1	94/196	1.05 (0.69-1.57)	0.8309
	2	88/182	0.96 (0.63-1.48)	0.8678
	3	120/212	0.98 (0.68-1.40)	0.8995
	4	117/217	0.91 (0.62-1.34)	0.6411
	5	119/215	0.74 (0.51-1.08)	0.1215
T3	1	100/202	0.97 (0.65-1.44)	0.8805
	2	80/158	1.11 (0.70-1.74)	0.6612
	3	93/198	0.88 (0.58-1.34)	0.5612
	4	127/245	1.00 (0.70-1.43)	0.9921
	5	125/234	0.82 (0.57-1.19)	0.2960
T4	1	97/189	0.70 (0.46-1.04)	0.0793
	2	73/166	0.79 (0.49-1.28)	0.3380
	3	116/217	1.05 (0.72-1.52)	0.8024
	4	120/221	1.00 (0.70-1.44)	0.9794
	5	119/224	1.09 (0.75-1.59)	0.6431

[1] Covariate Wald p-value;
Adjusted for age and gender

Running Multiple Models

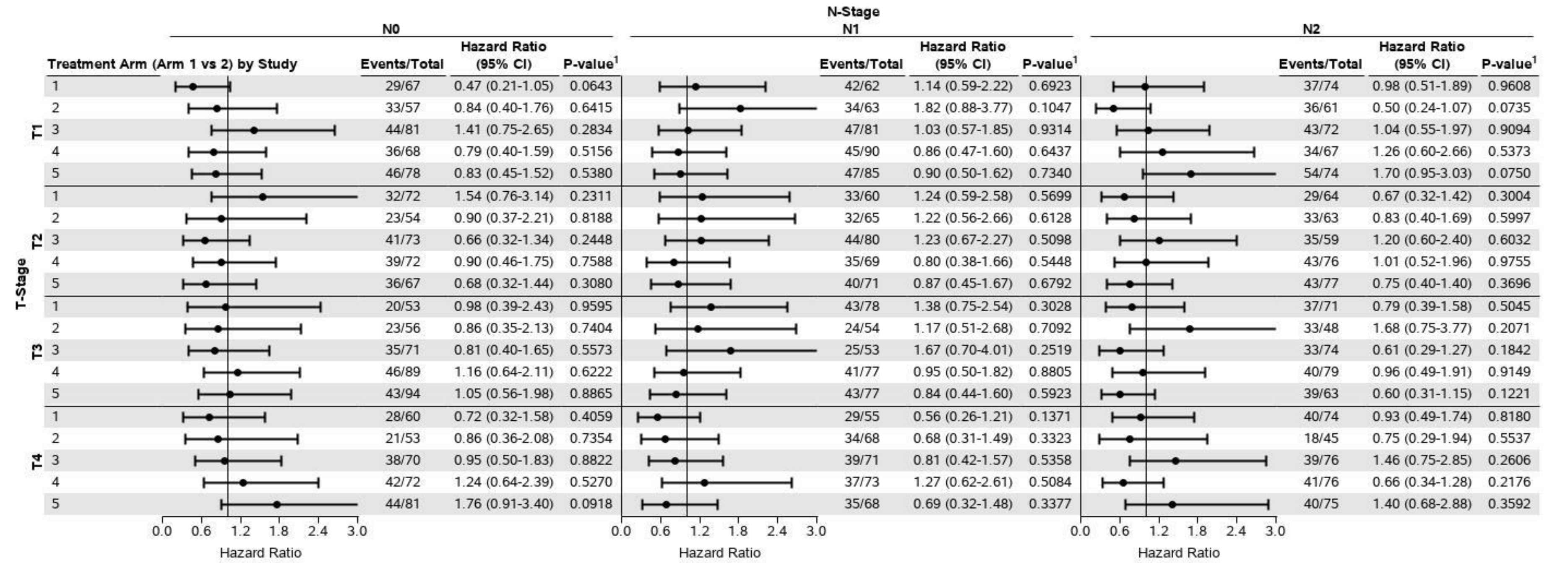
- ROWBY works well with skinny but long destinations such as PDF and RTF

T-Stage		Events/Total	Hazard Ratio (95% CI)	P-value[1]
T1	**Treatment Arm (Arm 1 vs Arm 2) by Study Number**			
	1	108/203	0.78 (0.53-1.13)	0.1899
	2	103/181	0.93 (0.62-1.38)	0.7123
	3	134/234	1.07 (0.76-1.51)	0.7009
	4	115/225	1.00 (0.68-1.45)	0.9813
	5	147/237	1.06 (0.76-1.48)	0.7127
T2	**Treatment Arm (Arm 1 vs Arm 2) by Study Number**			
	1	94/196	1.05 (0.69-1.57)	0.8309
	2	88/182	0.96 (0.63-1.48)	0.8678
	3	120/212	0.98 (0.68-1.40)	0.8995
	4	117/217	0.91 (0.62-1.34)	0.6411
	5	119/215	0.74 (0.51-1.08)	0.1215
T3	**Treatment Arm (Arm 1 vs Arm 2) by Study Number**			
	1	100/202	0.97 (0.65-1.44)	0.8805
	2	80/158	1.11 (0.70-1.74)	0.6612
	3	93/198	0.88 (0.58-1.34)	0.5612
	4	127/245	1.00 (0.70-1.43)	0.9921
	5	125/234	0.82 (0.57-1.19)	0.2960
T4	**Treatment Arm (Arm 1 vs Arm 2) by Study Number**			
	1	97/189	0.70 (0.46-1.04)	0.0793
	2	73/166	0.79 (0.49-1.28)	0.3380
	3	116/217	1.05 (0.72-1.52)	0.8024
	4	120/221	1.00 (0.70-1.44)	0.9794
	5	119/224	1.09 (0.75-1.59)	0.6431

[1]Covariate Wald p-value;
Adjusted for age and gender

Running Multiple Models

%MVMODELS(
…, COLBY=nstage, ROWBY=tstage, BYLABELON=0,
SUBTITLEHEADER=Treatment Arm (Arm 1 vs 2) by Study);



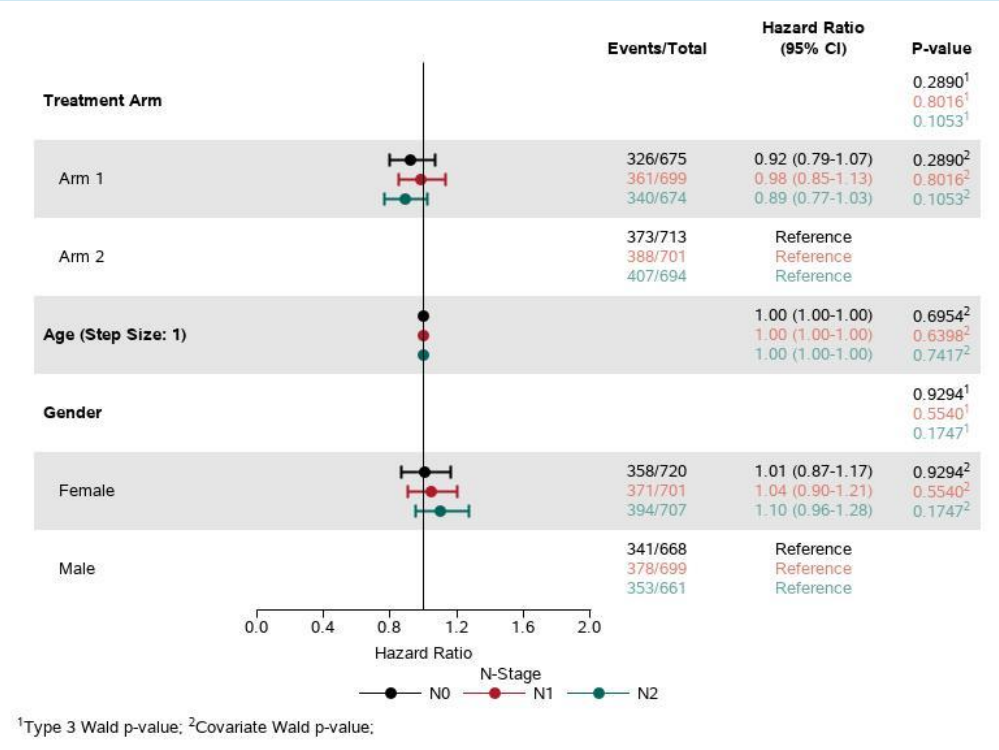
T-Stage	Treatment Arm (Arm 1 vs 2) by Study	N0 Events/Total	N0 Hazard Ratio (95% CI)	N0 P-value[1]	N1 Events/Total	N1 Hazard Ratio (95% CI)	N1 P-value[1]	N2 Events/Total	N2 Hazard Ratio (95% CI)	N2 P-value[1]
T1	1	29/67	0.47 (0.21-1.05)	0.0643	42/62	1.14 (0.59-2.22)	0.6923	37/74	0.98 (0.51-1.89)	0.9608
T1	2	33/57	0.84 (0.40-1.76)	0.6415	34/63	1.82 (0.88-3.77)	0.1047	36/61	0.50 (0.24-1.07)	0.0735
T1	3	44/81	1.41 (0.75-2.65)	0.2834	47/81	1.03 (0.57-1.85)	0.9314	43/72	1.04 (0.55-1.97)	0.9094
T1	4	36/68	0.79 (0.40-1.59)	0.5156	45/90	0.86 (0.47-1.60)	0.6437	34/67	1.26 (0.60-2.66)	0.5373
T1	5	46/78	0.83 (0.45-1.52)	0.5380	47/85	0.90 (0.50-1.62)	0.7340	54/74	1.70 (0.95-3.03)	0.0750
T2	1	32/72	1.54 (0.76-3.14)	0.2311	33/60	1.24 (0.59-2.58)	0.5699	29/64	0.67 (0.32-1.42)	0.3004
T2	2	23/54	0.90 (0.37-2.21)	0.8188	32/65	1.22 (0.56-2.66)	0.6128	33/63	0.83 (0.40-1.69)	0.5997
T2	3	41/73	0.66 (0.32-1.34)	0.2448	44/80	1.23 (0.67-2.27)	0.5098	35/59	1.20 (0.60-2.40)	0.6032
T2	4	39/72	0.90 (0.46-1.75)	0.7588	35/69	0.80 (0.38-1.66)	0.5448	43/76	1.01 (0.52-1.96)	0.9755
T2	5	36/67	0.68 (0.32-1.44)	0.3080	40/71	0.87 (0.45-1.67)	0.6792	43/77	0.75 (0.40-1.40)	0.3696
T3	1	20/53	0.98 (0.39-2.43)	0.9595	43/78	1.38 (0.75-2.54)	0.3028	37/71	0.79 (0.39-1.58)	0.5045
T3	2	23/56	0.86 (0.35-2.13)	0.7404	24/54	1.17 (0.51-2.68)	0.7092	33/48	1.68 (0.75-3.77)	0.2071
T3	3	35/71	0.81 (0.40-1.65)	0.5573	25/53	1.67 (0.70-4.01)	0.2519	33/74	0.61 (0.29-1.27)	0.1842
T3	4	46/89	1.16 (0.64-2.11)	0.6222	41/77	0.95 (0.50-1.82)	0.8805	40/79	0.96 (0.49-1.91)	0.9149
T3	5	43/94	1.05 (0.56-1.98)	0.8865	43/77	0.84 (0.44-1.60)	0.5923	39/63	0.60 (0.31-1.15)	0.1221
T4	1	28/60	0.72 (0.32-1.58)	0.4059	29/55	0.56 (0.26-1.21)	0.1371	40/74	0.93 (0.49-1.74)	0.8180
T4	2	21/53	0.86 (0.36-2.08)	0.7354	34/68	0.68 (0.31-1.49)	0.3323	18/45	0.75 (0.29-1.94)	0.5537
T4	3	38/70	0.95 (0.50-1.83)	0.8822	39/71	0.81 (0.42-1.57)	0.5358	39/76	1.46 (0.75-2.85)	0.2606
T4	4	42/72	1.24 (0.64-2.39)	0.5270	37/73	1.27 (0.62-2.61)	0.5084	41/76	0.66 (0.34-1.28)	0.2176
T4	5	44/81	1.76 (0.91-3.40)	0.0918	35/68	0.69 (0.32-1.48)	0.3377	40/75	1.40 (0.68-2.88)	0.3592

N-Stage

[1]Covariate Wald p-value;
Adjusted for age and gender

Running Multiple Models

```
%MVMODELS(
    ...,
    GROUPBY=nstage,
    SHOW_ADJCOVARIATES=1,
    CAT_DISPLAY=4);
```



Running Multiple Models

- Using grouping variables creates many subgrouped models quickly but is not as customizable as manually running models
- BY is useful for quickly replicating the same model across multiple variable levels
- COLBY is useful for comparing models side-by-side
- ROWBY is useful for comparing models vertically
- GROUPBY is useful for comparing components of different models more easily

Creating the Forest Plot

- Graph Template Language (GTL) is used to create the forest plot
- Macro creates dataset conducive for generating a forest plot
- Macro uses annotation functions such as DRAWTEXT over data-driven graph statements
- SGRENDER is used to render the final graph

Creating the Forest Plot

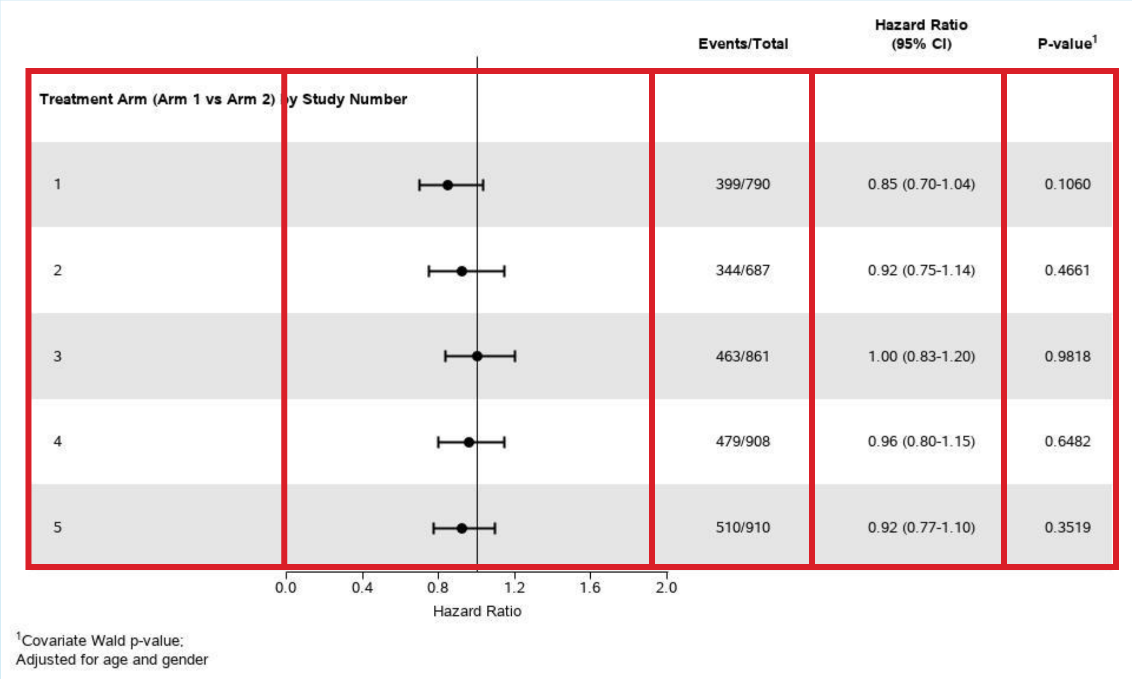
- Dataset is created to be one row for each row of the forest plot
- Each graph has variables for the estimate and confidence limits
- Each statistic has a character column of pre-formatted text
  - Example: HR_EST_RANGE is formatted to be x.xx (x.xx-x.xx)
- The subtitle (or row header) column has variables for indicating boldface, number of indentations, and whether the row should be shaded.
- When COLBY is used each variable has a suffix of _n (_1, _2, _3) to indicate which COLBY level the variable's belong to
- P-values are kept as two components: the p-value and the footnote indicator number. These are combined in the graph template
- The lower limit and upper limit numbers have "cap" variables to indicate whether the error bar should be capped or not
  - If error bars exceed the width of the graph space they are not capped by default to indicate they do not stop

Creating the Forest Plot

subind	boldind	subtitle	ev_t_1	estimate1_1	lcl1_1	ucl1_1	hr_est_range_1	pval_1	pfoot_index_1	y	shadeind	lcl1_cap1	ucl1_cap1
0	1	This is a really long title that goes across multiple panels								.5			
1	1	Study Number								1.5			
2	0	1	399/790							2.5			
3	1	Treatment Arm						0.1063	1	3.5			
4	0	Arm 1	185/397	0.85	0.70	1.04	0.85 (0.70-1.04)	0.1063	2	4.5		Serif	Serif
4	0	Arm 2	214/393				Reference			5.5			
2	0	2	344/687							6.5			
3	1	Treatment Arm						0.507	1	7.5			
4	0	Arm 1	163/346	0.93	0.75	1.15	0.93 (0.75-1.15)	0.507	2	8.5		Serif	Serif
4	0	Arm 2	181/341				Reference			9.5			
2	0	3	463/861							10.5			
3	1	Treatment Arm						0.939	1	11.5			
4	0	Arm 1	225/418	0.99	0.83	1.19	0.99 (0.83-1.19)	0.939	2	12.5		Serif	Serif
4	0	Arm 2	238/443				Reference			13.5			
2	0	4	479/908							14.5			
3	1	Treatment Arm						0.6708	1	15.5			
4	0	Arm 1	235/462	0.96	0.80	1.15	0.96 (0.80-1.15)	0.6708	2	16.5		Serif	Serif
4	0	Arm 2	244/446				Reference			17.5			
2	0	5	510/910							18.5			
3	1	Treatment Arm						0.3492	1	19.5			
4	0	Arm 1	219/425	0.92	0.77	1.10	0.92 (0.77-1.10)	0.3492	2	20.5		Serif	Serif
4	0	Arm 2	291/485				Reference			21.5			

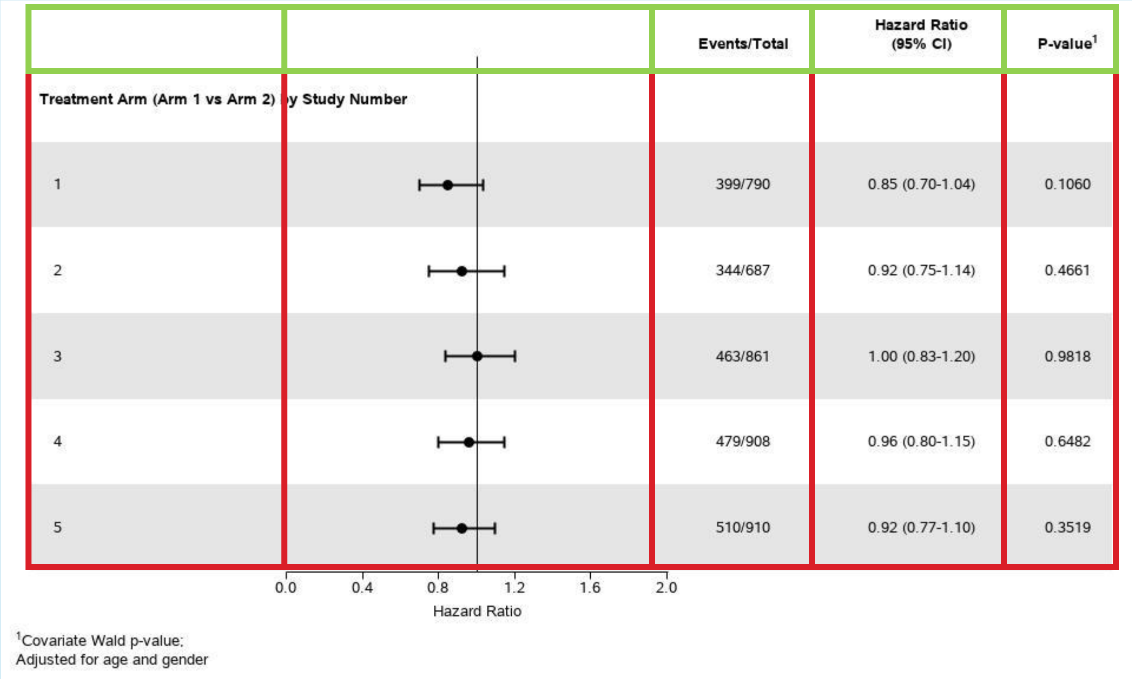
Creating the Forest Plot

- The plot is created using a lattice layout
  - One column for each plot or statistic
  - One row for each level of ROWBY
- Statistical columns have text drawn by DRAWTEXT
- Plot columns have estimates drawn by SCATTERPLOT and confidence bands drawn by HIGHLOWPLOT
- Axis labels are drawn in the COLUMNHEADERS space with ENTRY statements within GRIDDED layouts to allow multiple rows



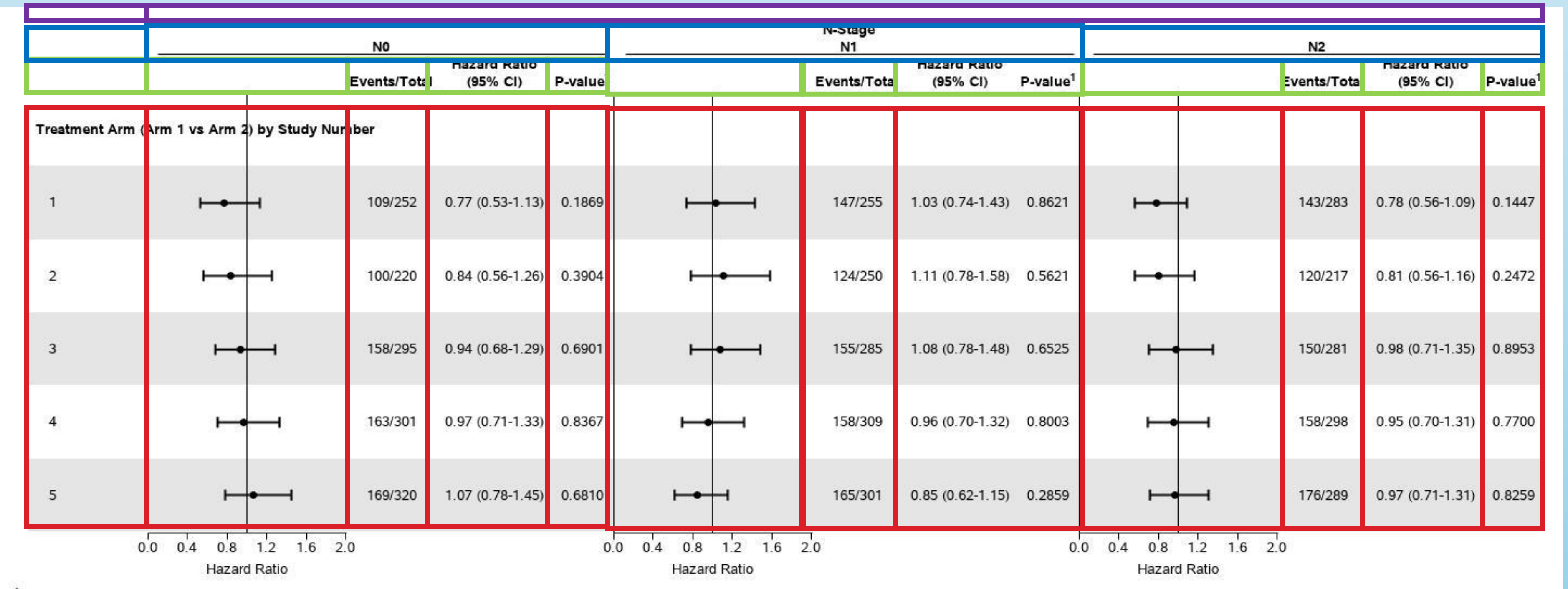
Creating the Forest Plot

- Column headers are drawn in a SIDEBAR block that holds another LATTICE layout with the same weights as the graph (Green)
  - Normally this could be done in a COLUMN2HEADERS block, but this interferes with putting spanning headers in for COLBY labels
  - DRAWTEXT is used here instead of ENTRY statements
- Titles and footnotes are drawn with the ENTRYTITLE and ENTRYFOOTNOTE statements



Creating the Forest Plot

- Spanning headers are drawn with another SIDEBAR block with a LATTICE layout (BLUE)
- COLBY labels are drawn with a final SIDEBAR block with a LATTICE layout (Purple)

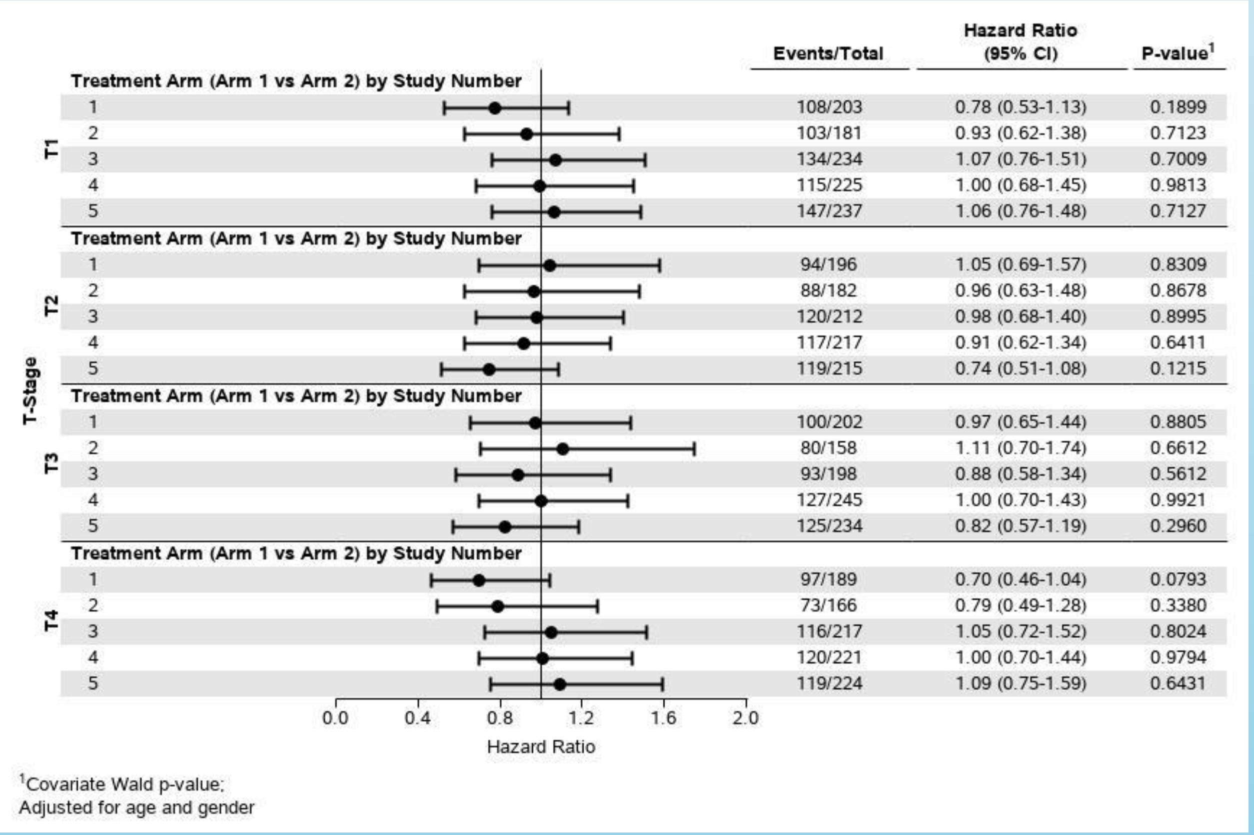


Treatment Arm (Arm 1 vs Arm 2) by Study Number	N0 Events/Total	N0 Hazard Ratio (95% CI)	N0 P-value[1]	N1 Events/Total	N1 Hazard Ratio (95% CI)	N1 P-value[1]	N2 Events/Total	N2 Hazard Ratio (95% CI)	N2 P-value[1]
1	109/252	0.77 (0.53-1.13)	0.1869	147/255	1.03 (0.74-1.43)	0.8621	143/283	0.78 (0.56-1.09)	0.1447
2	100/220	0.84 (0.56-1.26)	0.3904	124/250	1.11 (0.78-1.58)	0.5621	120/217	0.81 (0.56-1.16)	0.2472
3	158/295	0.94 (0.68-1.29)	0.6901	155/285	1.08 (0.78-1.48)	0.6525	150/281	0.98 (0.71-1.35)	0.8953
4	163/301	0.97 (0.71-1.33)	0.8367	158/309	0.96 (0.70-1.32)	0.8003	158/298	0.95 (0.70-1.31)	0.7700
5	169/320	1.07 (0.78-1.45)	0.6810	165/301	0.85 (0.62-1.15)	0.2859	176/289	0.97 (0.71-1.31)	0.8259

[1]Covariate Wald p-value;
Adjusted for age and gender

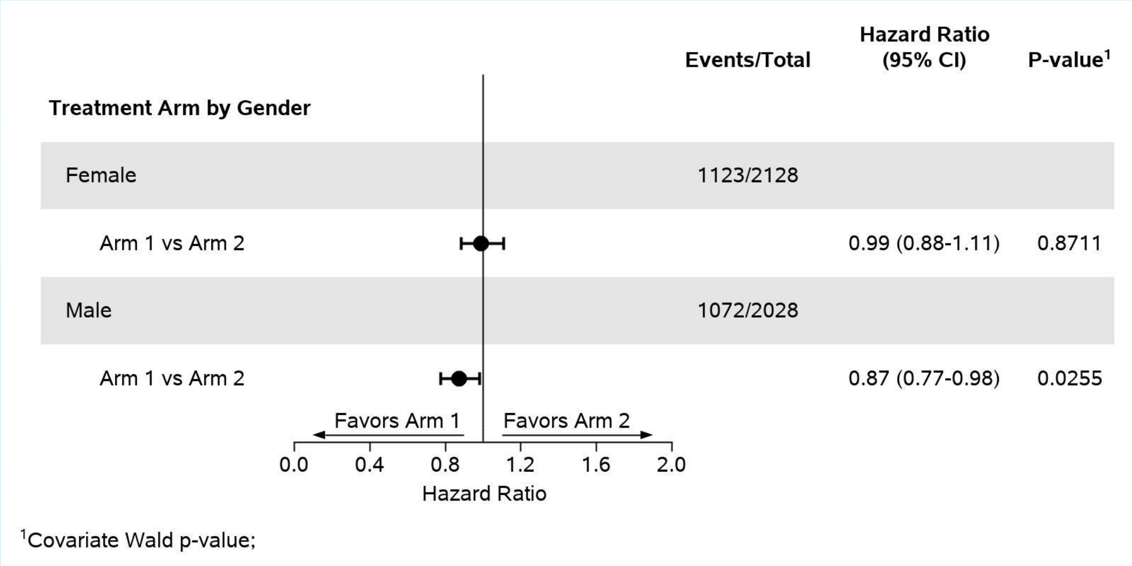
Creating the Forest Plot

- ROWBY headers are drawn within the ROWHEADERS block with DRAWTEXT statements and a GRIDDED layout
- ROWBY labels are drawn in a SIDEBAR block with DRAWTEXT statements and a GRIDDED layout
- Underlined column headers are created using DRAWLINE statements



Creating the Forest Plot

- Reference guides are available to describe what each side of the reference line indicates for estimates
- Space is allocated within the graph by adding on one row for each row of text plus one row for the arrow
- Text is drawn with DRAWTEXT and the arrows are drawn with DRAWARROW




```
%mvmodels(..., REFLINE=1,
   REFGUIDELOWER=Favors Arm 1,
   REFGUIDEUPPER=Favors Arm 2);
```

Annotation vs Data Driven or ENTRY Statements

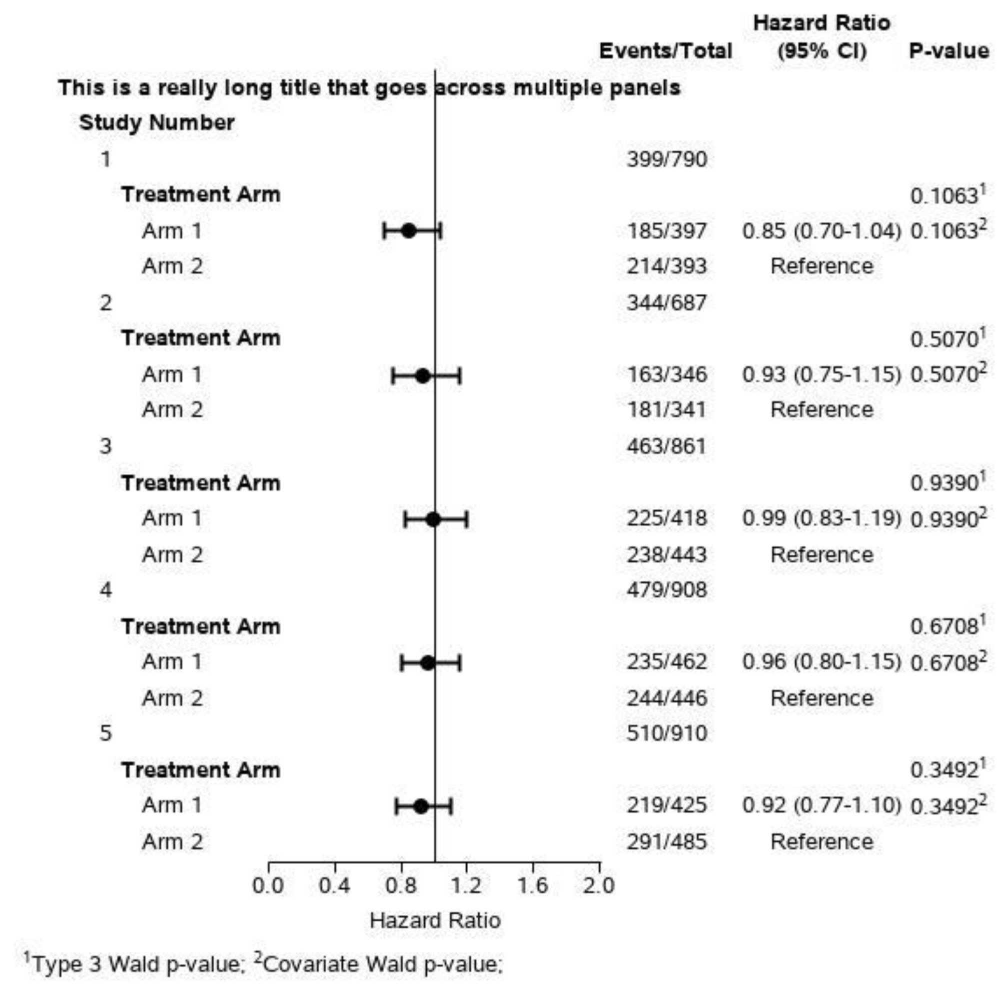
Annotation	Data Driven and ENTRY
- Allow use of Unicode, superscripts and subscripts - Can flow across multiple panels - More cumbersome to program manually - Requires plot space to be "initialized"	- Limited use of Unicode - Superscripts and subscripts only in ENTRY statements - Limited to current Layout block (Overlay, Gridded, etc.) and the potential to get cut off - Easier to program

Annotation vs Data Driven or ENTRY Statements

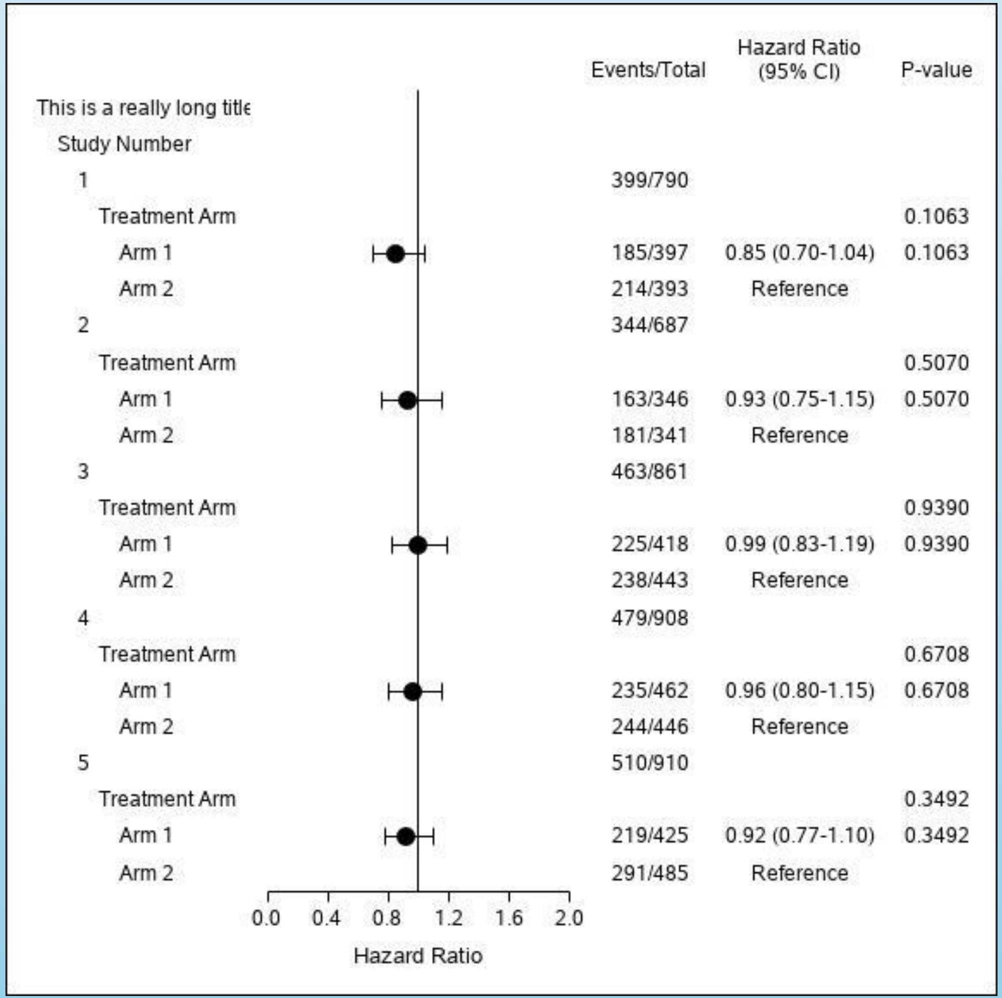
- Macro programming avoids the tediousness of annotation
- Forest plots have the potential to have a large number of text collisions across panels that annotation avoids
- Flexibility for Unicode, superscript and subscript is needed for variable levels and footnotes

Annotation vs Data Driven or ENTRY Statements

%MVMODELS Graph



Data Driven Same Dataset



Outputting Table to Different Destinations

- Macro is designed to output the table to various destinations while maintaining the same appearance
  - RTF, PDF, HTML, EXCEL, POWERPOINT, LISTING, and WORD

- Excel destination is good for wide and/or long tables
- RTF, PDf, WORD and HTML are good for long tables
- LISTING is great for outputting results to .LST files

Outputting Table to Different Destinations

ODS PDF output

	Events/Total	Hazard Ratio (95% CI)	P-value
Model 1	2195/4156		
Treatment Arm			0.0799 [1]
Arm 1	1027/2048	0.93 (0.85-1.01)	0.0799 [2]
Arm 2	1168/2108	Reference	
Age (Step size: 10.0)		1.00 (0.97-1.02)	0.7769 [2]
Gender			0.1971 [1]
Female	1123/2128	1.06 (0.97-1.15)	0.1971 [2]
Male	1072/2028	Reference	

[1]Type 3 Wald p-value;[2]Covariate Wald p-value;

ODS POWERPOINT output

	Events/Total	Hazard Ratio (95% CI)	P-value
Model 1	2195/4156		
Treatment Arm			0.0799[1]
Arm 1	1027/2048	0.93 (0.85-1.01)	0.0799[2]
Arm 2	1168/2108	Reference	
Age (Step size: 10.0)		1.00 (0.97-1.02)	0.7769[2]
Gender			0.1971[1]
Female	1123/2128	1.06 (0.97-1.15)	0.1971[2]
Male	1072/2028	Reference	

[1]Type 3 Wald p-value;[2]Covariate Wald p-value;

ODS EXCEL output

		Events/Total	Hazard Ratio (95% CI)	P-value
37				
38				
39		Events/Total	Hazard Ratio (95% CI)	P-value
40	**Model 1**	2195/4156		
41	**Treatment Arm**			0.0799[1]
42	Arm 1	1027/2048	0.93 (0.85-1.01)	0.0799[2]
43	Arm 2	1168/2108	Reference	
44	**Age (Step size: 10.0)**		1.00 (0.97-1.02)	0.7769[2]
45	**Gender**			0.1971[1]
46	Female	1123/2128	1.06 (0.97-1.15)	0.1971[2]
47	Male	1072/2028	Reference	
48	[1]Type 3 Wald p-value;[2]Covariate Wald p-value;			

ODS LISTING output

	Events/Total	Hazard Ratio (95% CI)	P-value
Model 1	2195/4156		
Treatment Arm			0.0799*
Arm 1	1027/2048	0.93 (0.85-1.01)	0.0799**
Arm 2	1168/2108	Reference	
Age (Step size: 10.0)		1.00 (0.97-1.02)	0.7769**
Gender			0.1971*
Female	1123/2128	1.06 (0.97-1.15)	0.1971**
Male	1072/2028	Reference	

*Type 3 Wald p-value;**Covariate Wald p-value;

Conclusion

- The MVMODELS macro is a powerful tool for performing and outputting survival or logistic regression analysis
- Many options for modifying the forest plot or table to meet the user's needs
- Uses annotation to create a graph that is otherwise unobtainable
- The macro is available for download on the SAS Communities page

Name: Jeffrey Meyers
Organization: Mayo Clinic
E-mail: meyers.jeffrey@mayo.edu / jpmeyers.spa@gmail.com