

Demographic Table and Subgroup Summary Macro %TABLEN

Jeffrey Meyers, Mayo Clinic

ABSTRACT

Clinical trial publications frequently allocate at least one of the allotted tables to summarize the demographics, stratification factors, and other variables of interest of the patients involved with the study. These tables generally include basic distribution information such as frequencies, means, medians, and ranges while also comparing these distributions across a key factor, such as treatment arm, to determine whether there was any imbalance in patient populations when doing analysis. These distributions are not difficult to compute and combine into a table, but as treatments become more specific to patient characteristics such as genetic biomarkers and tumor stages there is a need to be able to display the distributions in subgroups. The macro %TABLEN is a tool developed to compute distribution statistics for continuous variables, date variables, discrete variables, univariate survival time-to-event variables, and univariate logistic regression variables and combine them into a table for publication. %TABLEN has multiple features for comparisons, subgrouping, and outputting to multiple destinations such as RTF and PDF. The macro %TABLEN is a valuable tool for any programmer summarizing patient level data.

INTRODUCTION

The most commonly displayed table in a clinical oncology manuscript is a descriptive summary of the patient demographic data. This includes basic background information such as age, gender, race, and performance score as well as stratification or grouping factors determined by the study protocol. The summary table is grouped by a variable such as treatment arm with comparison p-values to ensure that there is proper balance. Oncology treatments have become more specific to disease characteristics and patient genetics and as a result there is a need to summarize demographic data within key clinical subgroups as well as treatments. Calculating individual distributions of variables is simple to do within SAS, but organizing multiple variables within several subgroups into a singular table requires a large amount of coding. The macro %TABLEN was created to automate this process and output a flexible summary table to a multitude of Output Delivery System (ODS) locations.

EXAMPLE OF DEMOGRAPHIC SUMMARY TABLE

Clinical manuscripts often lead with a summary table of the demographics of patients included in the study to demonstrate a balance between arms and other subgroups. The format is rather standard, and the TABLEN macro focuses on creating tables in this format:

Table 1. Example of TABLEN output with different variable types

	Arm			P-value
	A (N=49)	B (N=55)	Total (N=104)	
Age				0.4116 ¹
N	49	55	104	
Mean (SD)	60.9 (11.34)	62.4 (8.73)	61.7 (10.02)	
Median	60.0	62.0	61.0	
Range	37.0, 81.0	37.0, 81.0	37.0, 81.0	
Randomization Date				0.2076 ¹
N	49	55	104	
Median	05/12/2010	04/19/2010	04/27/2010	
Range	01/14/2010, 08/11/2010	06/23/2009, 08/11/2010	06/23/2009, 08/11/2010	

	Arm			P-value
	A (N=49)	B (N=55)	Total (N=104)	
Gender, n (%)				0.6547 ²
Female	19 (38.8%)	19 (34.5%)	38 (36.5%)	
Male	30 (61.2%)	36 (65.5%)	66 (63.5%)	
Race, n (%)				0.0301 ²
White	46 (93.9%)	47 (85.5%)	93 (89.4%)	
Black or African American	0 (0.0%)	6 (10.9%)	6 (5.8%)	
Pacific Islander	0 (0.0%)	1 (1.8%)	1 (1.0%)	
Asian	0 (0.0%)	1 (1.8%)	1 (1.0%)	
Not reported	3 (6.1%)	0 (0.0%)	3 (2.9%)	
Smoking Status, n (%)				0.4379 ²
Current	16 (32.7%)	15 (27.3%)	31 (29.8%)	
Former	9 (18.4%)	16 (29.1%)	25 (24.0%)	
Never	24 (49.0%)	24 (43.6%)	48 (46.2%)	
Number of Metastatic Sites, n (%)				0.5648 ²
1	19 (38.8%)	29 (52.7%)	48 (46.2%)	
2	17 (34.7%)	15 (27.3%)	32 (30.8%)	
3	7 (14.3%)	6 (10.9%)	13 (12.5%)	
>=4	6 (12.2%)	5 (9.1%)	11 (10.6%)	

¹Kruskal-Wallis p-value; ²Chi-Square p-value;

Table 2. The TABLEN macro output summarizes continuous, discrete and date type variables in one easy to read format. The macro can also summarize survival and binomial variables.

MACRO KEY PARAMETERS

REQUIRED PARAMETERS

The TABLEN macro has a large number of parameters available to ensure the flexibility to adjust the table as needed, but there are only three parameters that are required in order to run the macro:

1. DATA: identifies the user input dataset used to build a table.
2. VAR: space delimited list of variables to summarize in the table. The variables are listed in the table in the same order that they are listed within this parameter.
3. TYPE: space delimited list of values to classify the variables in the VAR parameter. The following are valid values that may be used:
 - 1: continuous variable
 - 2: discrete variable
 - 3: date variable
 - 4: survival time variable
 - 5: logistic regression event variable

An example macro call would look like:

```
%TABLEN(DATA=sashelp.class, var=age sex weight height, type=1 2 1 1)
```

A special note about the TYPE parameter is that the last listed type will carry forward to the remaining variables. Meaning the user does not have to repeat the same value if all remaining variables are the same type. The previous example can be replaced with the following:

```
%TABLEN(DATA=sashelp.class, var=age sex weight height, type=1 2 1)
```

KEY OPTIONAL PARAMETERS

BY Variable Options

The BY variable creates one set of statistics for each level of the BY variable and allows the comparison of these measures with metrics such as p-values and hazard ratios (see table 1). The following are key parameters associated with using a BY variable:

- BYORDER: Allows the user to change the order of the BY variable levels without modifying the original data. The order by default is the unformatted values.
- BY_INCMISS: Determines if missing values are included as a valid value for comparisons
- PVALS: Determines which type of p-value is used for comparison. Different p-values are available for each variable type and are set for each variable in the VAR list individually.
- REFERENCE: Specifies the reference level when computing hazard ratios or odds ratios

Continuous and Date Type Variable Options

Continuous and date type variables are analyzed the same way, but a date format is applied date type variable results. The following options are available:

- CONTDISPLAY/DATEDISPLAY: Controls which distribution statistics are shown with a space delimited list. The statistics are shown in the order they are listed. The options are:
 - N: Number of non-missing records
 - NMISS: Number of missing records
 - N_NMISS: Combines N and NMISS as N (NMISS)
 - MEAN: Mean of the distribution
 - SD: Standard Deviation of the distribution
 - MEAN_SD: Combines Mean and SD as MEAN (SD)
 - MEDIAN: Median of the distribution
 - RANGE: Range of the distribution as MIN, MAX
 - IQR: Interquartile range of the distribution as Q1, Q3
 - MEDIAN_RANGE: Combines Median and Range as MEDIAN (MIN, MAX)
 - MEDIAN_IQR: Combines Median and IQR as MEDIAN (Q1, Q3)
- DATEFMT: Specifies the date format applied to date variable type statistics. Default is MMDDYY10.
- Parameters to control the significant digits of statistics (MEANDIGITS, SDDIGITS, MEDIANDIGITS, RANGEDIGITS and IQRDIGITS). These can be set differently for each variable to account for variables such as age and laboratory values that are rounded differently.

Table 2. Example of continuous and date type variables

	Arm			P-value
	A (N=49)	B (N=55)	Total (N=104)	
Age				0.4116 ¹
N (Missing)	49 (0)	55 (0)	104 (0)	
Mean (SD)	60.9 (11.34)	62.4 (8.73)	61.7 (10.02)	
Median (Range)	60.0 (37.0, 81.0)	62.0 (37.0, 81.0)	61.0 (37.0, 81.0)	
Randomization Date				0.0409 ²
N	49	55	104	
Median	05/12/2010	04/19/2010	04/27/2010	
Range	01/14/2010, 08/11/2010	06/23/2009, 08/11/2010	06/23/2009, 08/11/2010	

¹Kruskal-Wallis p-value; ²ANOVA F-test p-value;

Table 2. Example of continuous and date type variables

The code to run the above example is shown below:

```
%TABLEN(data=example, by=arm, var=age date_on, type=1 3, pvals=1 5,
        contdisplay=n_nmiss mean_sd median_range, datedisplay=n median range)
```

Discrete Type Variable Options

Options for discrete type variables revolve giving the user flexibility in displaying the results without the need to modify their source data.

- **DIS_DISPLAY:** Determines if the frequency counts, percentages or both are shown.
- **DIS_ORDER:** Allows the user to change the order of each discrete type variable's values individually. This parameter also allows the user to only print certain levels of the variable (e.g. only male percentages of a gender variable)
- **DIS_INCMISS:** Determines if missing values are counted as valid values and thus included in percentages.
- **PCTDIGITS:** Determines the number of significant digits for percentages. This can be changed for each variable.

Table 3. Example of discrete type variables

	Arm			P-value
	A (N=49)	B (N=55)	Total (N=104)	
Smoking Status, n (%)				0.4379 ¹
Never	24 (49.0%)	24 (43.6%)	48 (46.2%)	
Current	16 (32.7%)	15 (27.3%)	31 (29.8%)	
Former	9 (18.4%)	16 (29.1%)	25 (24.0%)	
Gender, n (%)				0.6547 ²
Female	19 (38.8%)	19 (34.5%)	38 (36.5%)	
Measurable#Disease, n (%)				0.3406 ³
Yes	46 (93.9%)	54 (98.2%)	100 (96.2%)	
No	3 (6.1%)	1 (1.8%)	4 (3.8%)	

¹Chi-Square p-value; ²Cochran-Armitage trend test; ³Fisher Exact p-value;

Table 3. Example of discrete type variables

The code to run the above example is shown below:

```
%TABLEN(data=example, by=arm, var=smoke_st sex dz_g, type=2,
        pvals=1 3 2, dis_order=3 1 2|1|)
```

Survival Type Variable Options

The TABLEN macro performs univariate survival analyses where the variable in the VAR list represents the time variable. The following parameters allow the user to determine the rest of the survival model:

- SURV_STAT: Matches a survival event variable to the time variable listed in the VAR list
- CEN_VL: Determines the value representing the non-event. This must be numeric.
- TIMELIST: Specifies time-points for Kaplan-Meier event-free rates. Unique time-points can be specified for each survival type variable. Other parameters allow the time variable to be transformed (TDIVISOR), a unit label to be applied (TIME_UNITS), and the display format (TIMELISTFMT: percentage vs. proportion).
- SURVDISPLAY: A space delimited list that determines which survival statistics are displayed as well as the order they are displayed in. Certain options can be listed but will not be displayed unless another option is used in conjunction.
 - N: Number of patients
 - EVENTS: Number of events
 - EVENTS_N: Combines N and Events as EVENTS/N
 - MEDIAN: Kaplan-Meier survival median and 95% CI
 - HR: Cox model hazard ratio and 95% CI
 - COXPVAL: Cox model Wald p-value comparing parameters
 - TIMELIST: Kaplan-Meier time-point event-free rates and 95% CI
- Significant digit rounding: The significant digits for the median (SMEDIANDIGITS), hazard ratios (HRDIGITS), and time-point event-free rates (TLDIGITS) can be modified for each survival variable.

Table 4. Example of survival type variables

	Arm			P-value
	A (N=49)	B (N=55)	Total (N=104)	
Followup#Time				0.3613 ¹
Events/N	44/49	51/55	95/104	
Median (95% CI)	4.24 (3.61 - 8.15)	8.28 (4.50 - 10.48)	5.91 (4.24 - 8.44)	
Hazard Ratio (95% CI)	1.21 (0.80 - 1.81)	Reference		
6 months Est (95% CI)	0.37 (0.25 - 0.54)	0.59 (0.48 - 0.74)	0.49 (0.40 - 0.60)	
12 months Est (95% CI)	0.20 (0.11 - 0.35)	0.26 (0.17 - 0.41)	0.23 (0.16 - 0.33)	
Time#to#Progression				0.2611 ²
Events/N	42/49	44/55	86/104	
Median (95% CI)	2.00 (1.84 - 4.50)	3.68 (3.52 - 5.32)	3.48 (2.20 - 3.75)	
Hazard Ratio (95% CI)	1.28 (0.83 - 1.97)	Reference		
3 months Est (95% CI)	0.38 (0.26 - 0.55)	0.70 (0.58 - 0.84)	0.54 (0.45 - 0.65)	
6 months Est (95% CI)	0.20 (0.11 - 0.37)	0.26 (0.16 - 0.42)	0.23 (0.16 - 0.34)	

¹Logrank p-value; ²Type-3 score p-value;

Table 4. Example of survival type variables

The code to run the above example is shown below:

```
%TABLEN(data=example, by=arm, var=fu_time pg_time,type=4, pvals=1 3,
  surv_stat=fu_stat pg_stat, cen_vl=1, tdivisor=30.44 30.44,
  timelist=6 12|3 6, time_units=months,
  survdisplay=events_n median hr timelist)
```

Logistic Type Variable Options

The TABLEN macro performs univariate logistic regression analyses when a BY variable is input where the variable in the VAR list represents the event variable. If a BY variable is not input then the binomial success rate is calculated. The following parameters allow the user to determine the rest of the logistic model:

- LOG_EVENT: Determines the event value for the binomial variable in the VAR list.
- LOG_BINFMT: Determines if the binomial success rate is calculated as a proportion or percentage.
- LOG_DISPLAY: A space delimited list that determines which statistics are displayed as well as the order they are displayed in. Certain options can be listed but will not be displayed unless another option is used in conjunction.
 - N: Sample size
 - EVENTS: Number of events or successes
 - ODDSRATIO: Logistic regression odds ratio and 95% CI
 - WALDPVAL: Logistic regression Wald p-value comparing parameters
 - BINRATE: Binomial success rate and 95% CI
- Significant digit rounding: The significant digits for the odds ratios (ORDIGITS) and binomial success rates (BINDIGITS) can be modified for each survival variable.

Table 5. Example of binomial type variables

	Arm		Total	P-value
	A (N=49)	B (N=55)	(N=104)	
Grade 3+ AE (Event=1)				0.4218 ¹
Events/N	41/49	49/55	90/104	
Event Rate (95% CI)	0.84 (0.73 - 0.94)	0.89 (0.81 - 0.97)	0.87 (0.80 - 0.93)	
Odds Ratio (95% CI)	0.63 (0.20 - 1.96)	Reference		

¹Type-3 Wald p-value;

Table 5. Example of binomial type variables

The code to run the above example is shown below:

```
%TABLEN(data=example, by=arm, var=gd3,type=5, pvals=1,
  log_event=1,
  log_display=events_n binrate oddsratio)
```

Note about using multiple variable types in one table

The TABLEN macro is designed to make changing options for specific variable types more convenient when there are multiple types of variables in the VAR list. Each variable type has options that can be set for each variable specifically (e.g. DIS_ORDER, CEN_VL, and LOG_EVENT), and to remove confusion these options apply ONLY to their specific variable type. For example:

```
%TABLEN(data=example,
        by=arm,
        var=age date_on
            sex race1 smoke_st
            fu_time pg_time
            gd3,
        type=1 3
            2 2 2
            4 4
            5,
        dis_order=2 1| |2 3 1,
        surv_stat=fu_stat pg_stat, cen_vl=1, tdivisor=30.44 30.44,
        timelist=6 12|3 6, time_units=months,
        survdisplay=events_n median hr timelist,
        log_event=1,
        log_display=events_n binrate oddsratio)
```

The option DIS_ORDER only applies to the discrete variables, so even though the first discrete variable is the third variable in the VAR list there is no need to put two blanks in front of DIS_ORDER to compensate. The first value in DIS_ORDER will apply to the first discrete variable in the VAR list, the second value will correspond to the second discrete variable in the VAR list, and so forth. The same applies to survival options corresponding to survival variables, binomial options corresponding to binomial variables, etc.

CREATING NESTED TABLES

Clinical trials are becoming increasingly focused on subgroup analyses including biomarkers, genomic markers, and tumor staging. This has led to the summary tables becoming more complicated to make, and the TABLEN macro has innovated features to address this.

COLUMN BY (COLBY)

The COLBY option allows the tables to be nested into subgroups in multiple columns. The variable list, BY groupings, and p-values are summarized within each level of the COLBY variable. There is no limit to how many levels the COLBY variable can have, and this feature can work very nicely when combined with the ODS EXCEL destination.

Table 6. Example of TABLEN output using a COLBY variable

	Gender					
	Female			Male		
	Arm		P-value	Arm		P-value
A (N=19)	B (N=19)	A (N=30)		B (N=36)		
Age			0.3348 ¹			0.8064 ¹
N	19	19		30	36	
Mean (SD)	61.2 (11.84)	64.4 (8.49)		60.7 (11.20)	61.4 (8.79)	
Median	61.0	65.0		60.0	61.0	
Range	43.0, 81.0	45.0, 78.0		37.0, 80.0	37.0, 81.0	

¹Kruskal-Wallis p-value;

Table 6. Each level of the COLBY variable will have its own horizontal section. Only the levels of the BY variable that exist within a level of a COLBY will be shown. P-values are calculated within each subgroup individually.

ROW BY (ROWBY)

The ROWBY option allows the tables to be nested into subgroups in multiple rows. The variable list, BY groupings, and p-values are summarized within each level of the ROWBY variable. There is no limit to how many levels the ROWBY variable can have, and this feature is more useful when horizontal space is limited.

Table 7. Example of TABLEN output using a ROWBY variable

Gender		Arm		P-value
		A	B	
Female		(N=19)	(N=19)	
Age				0.3348 ¹
N		19	19	
Mean (SD)		61.2 (11.84)	64.4 (8.49)	
Median		61.0	65.0	
Range		43.0, 81.0	45.0, 78.0	
Male		(N=30)	(N=36)	
Age				0.8064 ¹
N		30	36	
Mean (SD)		60.7 (11.20)	61.4 (8.79)	
Median		60.0	61.0	
Range		37.0, 80.0	37.0, 81.0	

¹Kruskal-Wallis p-value;

Table 7. Each value of the ROWBY variable will have its own vertical section including a new subheader that will show the counts within that subgroup.

DIFFERENT TABLE STYLES

The table produced by TABLEN can have several style modifications to fit the needs of the user. These modifications include shading, spacing, and amount of cell borders.

TABLE SHADING

Table shading can either be disabled, shade alternating rows, or shade alternating variables. This is controlled by the SHADING parameter in the TABLEN macro call. The following are examples of disabling the shading and shading by row. The default of SHADING=1 to have shading by alternating row has been displayed in the prior table examples.

Table 8. Example table without shading

	Arm			P-value
	A	B	Total	
	(N=49)	(N=55)	(N=104)	
Age				0.4116 ¹
N	49	55	104	
Mean (SD)	60.9 (11.34)	62.4 (8.73)	61.7 (10.02)	
Median	60.0	62.0	61.0	
Range	37.0, 81.0	37.0, 81.0	37.0, 81.0	
Gender, n (%)				0.6547 ²
Female	19 (38.8%)	19 (34.5%)	38 (36.5%)	
Male	30 (61.2%)	36 (65.5%)	66 (63.5%)	

¹Kruskal-Wallis p-value; ²Chi-Square p-value;

Table 8. Setting SHADING=0 will remove all shading

Table shading becomes more useful the wider the table gets as the shading assists the reader's eyes with keeping rows aligned. Turning off the shading in a smaller table can create a cleaner and simpler look. When printing a longer table the third option of alternating shading between each variable is useful to help the eyes quickly track from one variable to the next.

Table 9. Example table of shading by each variable

	Arm			P-value
	A (N=49)	B (N=55)	Total (N=104)	
Age				0.4116 ¹
N	49	55	104	
Mean (SD)	60.9 (11.34)	62.4 (8.73)	61.7 (10.02)	
Median	60.0	62.0	61.0	
Range	37.0, 81.0	37.0, 81.0	37.0, 81.0	
Gender, n (%)				0.6547 ²
Female	19 (38.8%)	19 (34.5%)	38 (36.5%)	
Male	30 (61.2%)	36 (65.5%)	66 (63.5%)	

¹Kruskal-Wallis p-value; ²Chi-Square p-value;

Table 9. Setting SHADING=2 will enable shading by alternating variable

TABLE SPACING

Within the table there are three options for the space between variables: adding a blank row, adding a line, and having no space or line between variables. This option is controlled by the SPLIT parameter in the TABLEN macro call. The default of SPLIT=SPACE has been displayed in the prior table examples and will add a space between each variable.

Table 10. Example tables with different separators between variables

	Arm			P-value
	A (N=49)	B (N=55)	Total (N=104)	
Age				0.4116 ¹
N	49	55	104	
Mean (SD)	60.9 (11.34)	62.4 (8.73)	61.7 (10.02)	
Median	60.0	62.0	61.0	
Range	37.0, 81.0	37.0, 81.0	37.0, 81.0	
Gender, n (%)				0.6547 ²
Female	19 (38.8%)	19 (34.5%)	38 (36.5%)	
Male	30 (61.2%)	36 (65.5%)	66 (63.5%)	

¹Kruskal-Wallis p-value; ²Chi-Square p-value;

	Arm			P-value
	A (N=49)	B (N=55)	Total (N=104)	
Age				0.4116 ¹
N	49	55	104	
Mean (SD)	60.9 (11.34)	62.4 (8.73)	61.7 (10.02)	
Median	60.0	62.0	61.0	
Range	37.0, 81.0	37.0, 81.0	37.0, 81.0	
Gender, n (%)				0.6547 ²
Female	19 (38.8%)	19 (34.5%)	38 (36.5%)	
Male	30 (61.2%)	36 (65.5%)	66 (63.5%)	

¹Kruskal-Wallis p-value; ²Chi-Square p-value;

Table 10. Setting SPLIT=NONE will have no lines or spaces between variables, and SPLIT=LINE

will draw a line between variables without a space.

CELL BORDERS

The amount of cell borders is controlled by the BORDERDISPLAY parameter in the TABLEN macro call and has four different options. The default of BORDERDISPLAY=1 only shows border lines at the top and bottom of the table and has been used in the prior table examples.

Table 11. Example tables with different BORDERDISPLAY options

	Arm			
	A (N=49)	B (N=55)	Total (N=104)	P-value
Age				0.4116 ¹
N	49	55	104	
Mean (SD)	60.9 (11.34)	62.4 (8.73)	61.7 (10.02)	
Median	60.0	62.0	61.0	
Range	37.0, 81.0	37.0, 81.0	37.0, 81.0	
Gender, n (%)				0.6547 ²
Female	19 (38.8%)	19 (34.5%)	38 (36.5%)	
Male	30 (61.2%)	36 (65.5%)	66 (63.5%)	

¹Kruskal-Wallis p-value; ²Chi-Square p-value;

	Arm			
	A (N=49)	B (N=55)	Total (N=104)	P-value
Age				0.4116 ¹
N	49	55	104	
Mean (SD)	60.9 (11.34)	62.4 (8.73)	61.7 (10.02)	
Median	60.0	62.0	61.0	
Range	37.0, 81.0	37.0, 81.0	37.0, 81.0	
Gender, n (%)				0.6547 ²
Female	19 (38.8%)	19 (34.5%)	38 (36.5%)	
Male	30 (61.2%)	36 (65.5%)	66 (63.5%)	

¹Kruskal-Wallis p-value; ²Chi-Square p-value;

	Arm			
	A (N=49)	B (N=55)	Total (N=104)	P-value
Age				0.4116 ¹
N	49	55	104	
Mean (SD)	60.9 (11.34)	62.4 (8.73)	61.7 (10.02)	
Median	60.0	62.0	61.0	
Range	37.0, 81.0	37.0, 81.0	37.0, 81.0	
Gender, n (%)				0.6547 ²
Female	19 (38.8%)	19 (34.5%)	38 (36.5%)	
Male	30 (61.2%)	36 (65.5%)	66 (63.5%)	

¹Kruskal-Wallis p-value; ²Chi-Square p-value;

Table 11. BORDERDISPLAY=2 turns on all cell borders, BORDERDISPLAY=3 turns on all cell borders except for title and footnote, and BORDERDISPLAY=4 turns on borders for columns

SAVING THE TABLE INTO VARIOUS DESTINATIONS

The TABLEN macro is designed to be able to output the table in the same style no matter the destination. This includes the LISTING, RTF, PDF, EXCEL, HTML, and POWERPOINT destinations. The macro call can either output a document directly with the OUTDOC option or be included with other items inside ODS tags such as ODS RTF.

Figure 1. TABLEN outputting to multiple ODS destinations

Arm				
	A (N=49)	B (N=55)	Total (N=104)	P-value
Age				0.4116 ¹
N	49	55	104	
Mean (SD)	60.9 (11.34)	62.4 (8.73)	61.7 (10.02)	
Median	60.0	62.0	61.0	
Range	37.0, 81.0	37.0, 81.0	37.0, 81.0	
¹ Kruskal-Wallis p-value;				

Arm				
	A (N=49)	B (N=55)	Total (N=104)	P-value
Age				0.4116 ¹
N	49	55	104	
Mean (SD)	60.9 (11.34)	62.4 (8.73)	61.7 (10.02)	
Median	60.0	62.0	61.0	
Range	37.0, 81.0	37.0, 81.0	37.0, 81.0	
¹ Kruskal-Wallis p-value;				

	A	B	C	D	E
1					
2		Arm			
3		A (N=49)	B (N=55)	Total (N=104)	P-value
4	Age				0.4116 ¹
5	N	49	55	104	
6	Mean (SD)	60.9 (11.34)	62.4 (8.73)	61.7 (10.02)	
7	Median	60	62	61	
8	Range	37.0, 81.0	37.0, 81.0	37.0, 81.0	
9	¹ Kruskal-Wallis p-value;				

Arm				
	A (N=49)	B (N=55)	Total (N=104)	P-value
Age				0.4116 ¹
N	49	55	104	
Mean (SD)	60.9 (11.34)	62.4 (8.73)	61.7 (10.02)	
Median	60.0	62.0	61.0	
Range	37.0, 81.0	37.0, 81.0	37.0, 81.0	
¹ Kruskal-Wallis p-value;				

Arm				
	A (N=49)	B (N=55)	Total (N=104)	P-value
Age				0.4116 ¹
N	49	55	104	
Mean (SD)	60.9 (11.34)	62.4 (8.73)	61.7 (10.02)	
Median	60.0	62.0	61.0	
Range	37.0, 81.0	37.0, 81.0	37.0, 81.0	
¹ Kruskal-Wallis p-value;				

Figure 1. TABLEN outputting (in order) to the RTF, PDF, EXCEL, HTML and POWERPOINT destinations.

Each ODS destination has strengths and weaknesses when combined with the TABLEN macro:

- The RTF and PDF destinations are great for combining TABLEN with other outputs for a report, but have space limited to letter dimensions.

- The EXCEL destination has near unlimited horizontal and vertical space and can handle many different subgroups when using ROWBY and/or COLBY. The capability of freezing the row and column headers makes reading the report simple however the output does not look as professional as other destinations when drafting a report.
- The HTML and HTML5(WEB) destinations work well for creating long tables, but the column headers are not visible if scrolling too far.
- The POWERPOINT destination works well for including small tables into a presentation, but ODS POWERPOINT forces cell padding into the table which makes the table less compressed than other destinations.

The TABLEN macro also creates a table specifically for the LISTING destination that is very clean and compact. Much more information can be fit together in a smaller space than running multiple FREQ, MEANS and other procedural output.

Figure 2. TABLEN outputting to the LISTING destination

	Arm		Total	P-value
	A (N=49)	B (N=55)	(N=104)	
Age				0,4116*
N	49	55	104	
Mean (SD)	60,9 (11,34)	62,4 (8,73)	61,7 (10,02)	
Median	60,0	62,0	61,0	
Range	37,0, 81,0	37,0, 81,0	37,0, 81,0	

*Kruskal-Wallis p-value

Figure 2. The TABLEN macro produces a clean readable table into the listing output that can be combined with other procedure output when saving .LST files.

CONCLUSION

The %TABLEN macro is a flexible and powerful tool for creating demographic summary tables, and has the features required to meet the additional challenges put forth by more complicated clinical trials. The ability to handle multiple variable types, output to several ODS destinations, create nested tables, and perform on the fly data set modifications is useful to any programmer in the healthcare industry.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Name: Jeffrey Meyers
 Enterprise: Mayo Clinic
 Address: 200 First Street SW
 City, State ZIP: Rochester, MN 55905
 Work Phone: 507-266-2711
 E-mail: Meyers.jeffrey@mayo.edu / jpmeyers.spa@gmail.com
 Website (macro available for download): [Communities Page](#)

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

Other brand and product names are trademarks of their respective companies.