



Science For A Better Life

Frequency Tables – Aspects to Consider

15.10.2014 / PhUSE - Industry Starters / Katja Glaß



Agenda

- Introduction
- Do the double programming
 - Demography
 - Adverse Events
 - Laboratory
- Summary and further aspects

Introduction

- ❖ Typical tasks for a Statistical Analyst are the creation of
 - ❖ Statistical Analysis Plan
 - ❖ Analysis datasets
 - ❖ Tables, listings and figures

- ❖ Tables can be classified as
 - ❖ Descriptive Tables
 - ❖ Frequency Tables
 - ❖ Listing-Like Tables
 - ❖ Mixtures



Introduction

❖ Frequency Tables

❖ Numerator (x)

❖ Denominator (n)

❖ Percentages (p)

$$x / n (p\%)$$



Do the Double Programming

- ❖ Example:
 - ❖ Receive correct tables (are already validated)
 - ❖ Perform double programming to tune skills and awareness
- ❖ Expecting (nearly) no specifications
- ❖ Study data from PhUSE Scriptathon

- ❖ Examples for:
 - ❖ Demography
 - ❖ Adverse Events
 - ❖ Laboratory



Demography

Table 1: Demography

		Placebo	Total	
n		86 (100.0%)	254 (100.0%)	n
Sex				
F	x	53 (61.6%)	143 (56.3%)	p
M		33 (38.4%)	111 (43.7%)	
Ethnicity				
HISPANIC OR LATINO		3 (3.5%)	12 (4.7%)	
NOT HISPANIC OR LATINO		83 (96.5%)	242 (95.3%)	

Demography

- ❖ Investigate
 - ❖ Data, Variables, SAS Procedure (PROC FREQ)
 - ❖ „Total“ Treatment handling
 - Duplicate observations for „Total“



Program

Compare

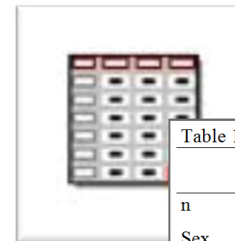
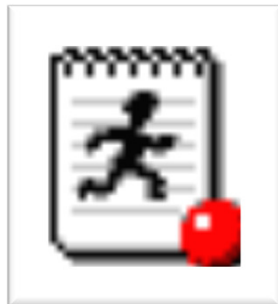


Table 1: Demography			
	Placebo	Xanomeline	
		High Dose	
n	86 (100.0%)	84 (100.0%)	
Sex			
F	53 (61.6%)	40 (47.6%)	
M	33 (38.4%)	44 (52.4%)	
Ethnicity			
HISPANIC OR LATINO	3 (3.5%)	3 (3.6%)	
NOT HISPANIC OR LATINO	83 (96.5%)	81 (96.4%)	

Adverse Events

Table 2: Subjects with Adverse Events

Body System / Reported Term	Placebo N=86 (100%)	Xanomeline High Dose N=84 (100%)
ANY EVENT	69 (80.2%)	79 (94.0%)
...		
VASCULAR DISORDERS	3 (3.5%)	2 (2.4%)
HOT FLUSH	0	0
HYPERTENSION	1 (1.2%)	1 (1.2%)
HYPOTENSION	2 (2.3%)	0
ORTHOSTATIC HYPOTENSION	1 (1.2%)	0
WOUND HAEMORRHAGE	0	1 (1.2%)

Adverse Events

❖ Data, Variables, PROC FREQ, „Total“ duplication



Program

Compare

Adverse Events

Term for “VASCULAR DISORDERS”	Table Result	Check Result
HYPERTENSION	1 (1.2%)	2
HYPOTENSION	2 (2.3%)	3
ORTHOSTATIC HYPOTENSION	1 (1.2%)	2



Adverse Events

Term for “VASCULAR DISORDERS”	Table Result	Check Result
HYPERTENSION	1 (1.2%)	2
HYPOTENSION	2 (2.3%)	3
ORTHOSTATIC HYPOTENSION	1 (1.2%)	2



USUBJID	AETERM
7	HYPERTENSION
7	HYPERTENSION
4	HYPOTENSION
1	HYPOTENSION
1	HYPOTENSION
4	ORTHOSTATIC HYPOTENSION
4	ORTHOSTATIC HYPOTENSION

Adverse Events

Term for “VASCULAR DISORDERS”	Table Result	Check Result
HYPERTENSION	1 (1.2%)	2
HYPOTENSION	2 (2.3%)	3
ORTHOSTATIC HYPOTENSION	1 (1.2%)	2



Observation Count
vs.
Subject Count

Adverse Events



❖ Perform subject count!

- Keep just one „to-count“ observation per subject
- For AETERM use „NODUPKEY“

USUBJID	AETERM
7	HYPERTENSION
4	HYPOTENSION
1	HYPOTENSION
4	ORTHOSTATIC HYPOTENSION

Adverse Events

Term for “VASCULAR DISORDERS”	Table Result	Check Result	Check Percent
HYPERTENSION	1 (1.2%)	1	25.0
HYPOTENSION	2 (2.3%)	2	50.0
ORTHOSTATIC HYPOTENSION	1 (1.2%)	1	25.0



Adverse Events

Term for “VASCULAR DISORDERS”	Table Result	Check Result	Check Percent
HYPERTENSION	1 (1.2%)	1	25.0
HYPOTENSION	2 (2.3%)	2	50.0
ORTHOSTATIC HYPOTENSION	1 (1.2%)	1	25.0



Population!

- ❖ Use all subjects within the study and not those within the dataset!

Adverse Events

- ❖ Process:
 - ❖ „Total“ treatment duplication
 - ❖ „NODUPKEY“ to have one „to-count“ item per subject
 - ❖ **Calculate population numbers (ADSL)**
 - ❖ Calculate frequencies by „Body System“ and treatment
 - ❖ **Calculate percentages**
- ❖ PROC SUMMARY as percentages are calculated manually



Adverse Events

Term for “VASCULAR DISORDERS”	Table Result	Check Result	Check Percent
HYPERTENSION	1 (1.2%)	1	1.2
HYPOTENSION	2 (2.3%)	2	2.3
ORTHOSTATIC HYPOTENSION	1 (1.2%)	1	1.2



❖ PROC SUMMARY

- Automatic „upper“ summaries

Adverse Events

Term	Table Result	Check Result	Check Percent
VASCULAR DISORDERS	3 (3.5%)	4	4.7
HYPERTENSION	1 (1.2%)	1	1.2
HYPOTENSION	2 (2.3%)	2	2.3
ORTHOSTATIC HYPOTENSION	1 (1.2%)	1	1.2



USUBJID	AETERM	AEBODSYS
7	HYPERTENSION	VASCULAR DISORDERS
4	HYPOTENSION	VASCULAR DISORDERS
1	HYPOTENSION	VASCULAR DISORDERS
4	ORT. HYPO.	VASCULAR DISORDERS

Adverse Events

Term	Table Result	Check Result	Check Percent
VASCULAR DISORDERS	3 (3.5%)	4	4.7
HYPERTENSION	1 (1.2%)	1	1.2
HYPOTENSION	2 (2.3%)	2	2.3
ORTHOSTATIC HYPOTENSION	1 (1.2%)	1	1.2



- ❖ PROC SUMMARY performs observation count!
- ❖ Calculate „upper“ statistics separately!

Adverse Events

Term	Table Result	Check Result	Check Percent
VASCULAR DISORDERS	3 (3.5%)	3	3.5
HYPERTENSION	1 (1.2%)	1	1.2
HYPOTENSION	2 (2.3%)	2	2.3
ORTHOSTATIC HYPOTENSION	1 (1.2%)	1	1.2



- ✓ SUBJECT count (“nodupkey”)
- ✓ POPULATION (ADSL)
- ✓ Perform different level calculations separately

Laboratory

Table 3: Number of subjects with laboratory ranges

Analysis Visit: Baseline		
Parameter	Reference Range Indicator	Total N=252 (100%)
Alanine Aminotransferase (U/L)	LOW	1 (0.4%)
	NORMAL	240 (95.2%)
	HIGH	11 (4.4%)
Alkaline Phosphatase (U/L)	HIGH	8 (3.2%)
	LOW	6 (2.4%)
	NORMAL	236 (93.7%)

❖ By time point analysis

Laboratory

- ❖ Data, Variables, Timepoint

- ❖ Process:

- ❖ Duplicate observations for „Total“ treatment group
- ❖ Reduce observation to have one „to-count“ item per subject („nodupkey“)
- ❖ Calculate population numbers
- ❖ Calculate frequencies
- ❖ Merge population and calculate percentages
- ❖ Reduce and compare the result



Laboratory

- ❖ Duplicate observations for „Total“ treatment group
 - ❖ Data modification
 - ❖ Lower performance

USUBJID		TRTA
1	1	Placebo
1	2	TOTAL
2	1	Xanomeline High Dose
2	2	TOTAL

- ❖ Working with summary functions for „Total“ treatment
 - ❖ Issues with observation vs. subject count
 - ❖ Issue with crossover studies

USUBJID		TRTA
1	1	Placebo
1	2	TOTAL
1	3	Xanomeline High Dose
1	4	TOTAL

- ❖ Solution: perform two calculations
 - ❖ Per treatment and
 - ❖ Overall treatments



Laboratory

- ❖ „NODUPKEY“ reduction might be problematic

SUBJID	AVISITN	PARAMCD	AVAL
1186	.	ALT	95
1186	.	ALT	92
1186	.	ALT	73



```
PROC SORT DATA=ads.adlbc OUT=adlbc DUPOUT=err NODUPKEY;
```

```
    BY avisitn avisit param usubjid;
```

```
RUN;
```

NOTE: The data set WORK.ADLBC has 74138 observations and 46 variables.

NOTE: The data set WORK.ERR has **126 observations** and 46 variables.

- ❖ Only use observations where AVISIT is provided -> no duplicated records

Laboratory

- ❖ Calculations

- ❖ PROC SUMMARY

- ❖ Population according ADSL

Visit = “Baseline” and Param = “Alanine Aminotransferase (U/L)” – Treatment Total **N=252 (100%)**

Reference Range Indicator	Table Result:	Check Result	Check Percent	Check Population
LOW	1 (0.4%)	1	0.4	254
NORMAL	240 (95.2%)	240	94.5	254
HIGH	11 (4.4%)	11	4.3	254



- ❖ Population differs!!!

- ❖ To be used here:
Number of subjects available at the specified time point

Laboratory

**Visit = “Baseline” and Param = “Alanine Aminotransferase (U/L)” – Treatment
Total N=252 (100%)**

Reference Range Indicator	Table Result:	Check Result	Check Percent	Check Population
LOW	1 (0.4%)	1	0.4	252
NORMAL	240 (95.2%)	240	95.2	252
HIGH	11 (4.4%)	11	4.4	252



Summary and Further Aspects

Major points to consider:

- ❖ Subject count vs. observation count
- ❖ Population / Denominator
- ❖ Reduction of duplicates
- ❖ Total treatment handling



Summary and Further Aspects

Further aspects:

- ❖ Missing handling
- ❖ Multidimensional frequencies

Total N=254 (100%)	
Female	143 (56.3%)
65-80 years	78 (30.7%)
<65 years	19 (7.5%)
>80 years	46 (18.1%)

- ❖ Cross frequencies

	NORMAL	ABNORMAL	TOTAL
NORMAL	222 (92.5%)	1 (0.4%)	225 (93.8%)
ABNORMAL	2 (0.8%)	7 (2.9%)	9 (3.8%)
TOTAL	226 (94.2%)	8 (3.3%)	240 (100.0%)

- ❖ „At-risk“ frequencies
with shifting denominator

Alanine Aminotransferase (U/L)	6/227 (2.6%)
Albumin (g/L)	16/234 (6.8%)
Alkaline Phosphatase (U/L)	4/226 (1.8%)
Aspartate Aminotransferase (U/L)	5/223 (2.2%)

Summary and Further Aspects

- ❖ Due to the various aspects of frequency tables it is of crucial importance to understand the possible frequency options as well as the underlying data.
- ❖ For a full understanding of requirements detailed specifications must be available.



Questions?





Science For A Better Life

Thank you!