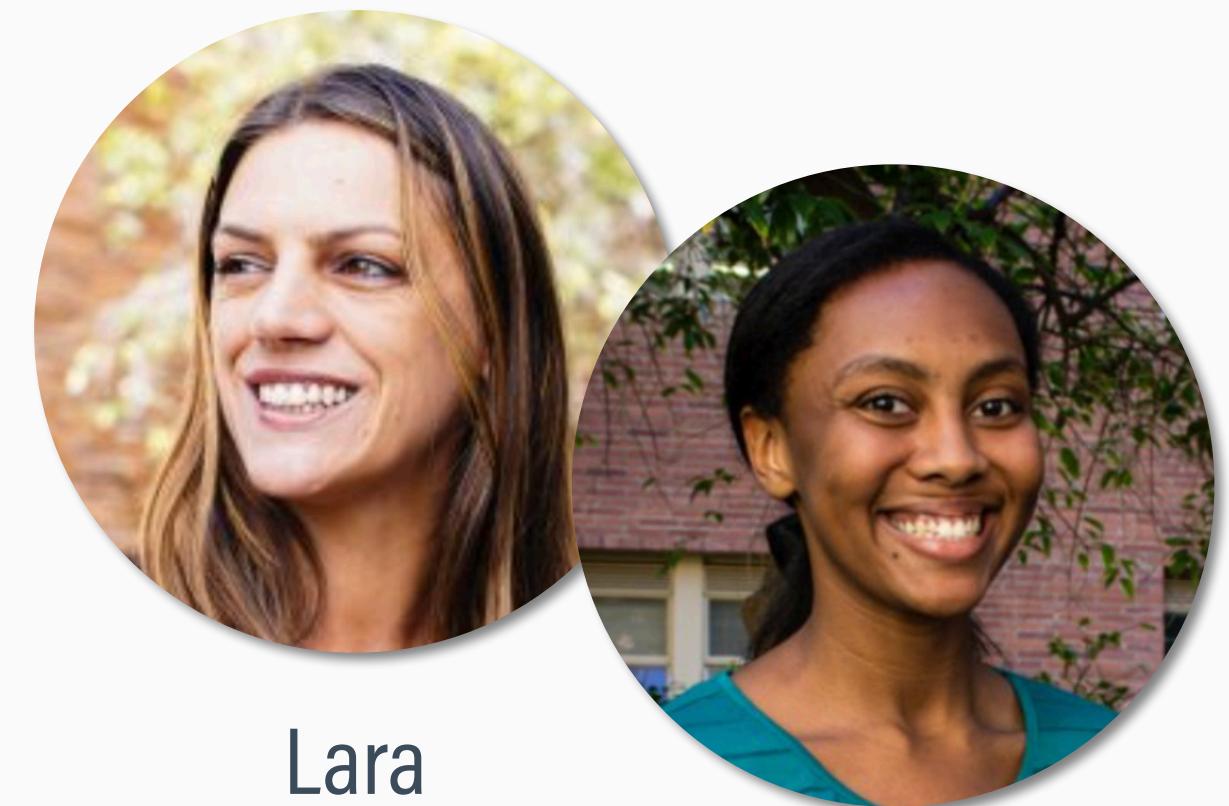


LAB WEEK 7

INDEPENDENT-SAMPLES T-TEST

SMOKING AND DRINKING CESSATION TRIAL

Pharmacological treatments that can concomitantly address cigarette smoking and heavy drinking stand to improve health care delivery for these highly prevalent co-occurring conditions. This superiority trial compared the combination of varenicline and naltrexone against varenicline alone for smoking cessation and drinking reduction among heavy-drinking smokers.

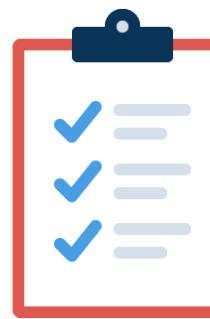


Lara
Ray

ReJoyce
Green

Ray, L.A., Green, R., Enders, C., et al. (2021). Efficacy of combining varenicline and naltrexone for smoking cessation and drinking reduction: A randomized clinical trial. *American Journal of Psychiatry*, 178, 818–828.

KEY VARIABLES



Breath (alveolar) carbon monoxide

A measure of carbon monoxide in the lungs.
Breath carbon monoxide is a biomarker of smoking behavior common in clinical trials.



Medication arm

Participants were randomly assigned to receive one of two meds: varenicline plus naltrexone or varenicline plus placebo pills

RESEARCH QUESTION

- Question: Does the combination of two medications improve smoking cessation beyond a single medication alone?
- The study used a between-subjects design where smoking levels were evaluated in two distinct groups
- An independent t-test is appropriate for comparing two means from different subsamples

LOAD PACKAGES AND IMPORT DATA

- = data frame name
- = variable name
- = raw data file name

```
# LOAD R PACKAGES ----  
  
# load R packages  
library(ggplot2)  
library(Hmisc)  
library(psych)  
library(summarytools)  
  
# READ DATA ----  
  
# github url for raw data  
filepath <-  
  'https://raw.githubusercontent.com/craigenders/psych250a/main/data/ClinicalTrialData.csv'  
  
# create data frame called ClinicalTrial from github data  
ClinicalTrial <- read.csv(filepath, stringsAsFactors = T)
```

SUMMARIZING DATA

- = data frame name
- = variable name

```
# INSPECT DATA ----
```

```
# summarize entire data frame (summarytools package)
dfSummary(ClinicalTrial)
```

```
# DESCRIPTIVE STATISTICS ----
```

```
# descriptive statistics for entire data frame (psych package)
describe(ClinicalTrial)
```

R OUTPUT

Data Frame Summary

ClinicalTrial

Dimensions: 165 x 2

Duplicates: 125

| No | Variable | Stats / Values | Freqs (% of Valid) | Graph | Valid | Missing |
|----|-----------------------|---|--------------------------|-------------------------|-----------------|-------------|
| 1 | Condition [factor] | 1. Varenicline 2. Varenicline + Naltrexone | 82 (49.7%) 83 (50.3%) | IIIIIIIII IIIIIIIIII | 165 (100.0%) | 0 (0.0%) |
| 2 | COWeek8 [integer] | Mean (sd) : 5.5 (6) min < med < max: 0 < 3 < 29 IQR (CV) : 8 (1.1) | 25 distinct values | : | 165 (100.0%) | 0 (0.0%) |

R OUTPUT

| | vars | n | mean | sd | median | trimmed | mad | min | max | range | skew | kurtosis | se |
|----------------|------|-----|-------|-------|--------|---------|-------|-----|-----|-------|-------|----------|------|
| Participant | 1 | 165 | 83.00 | 47.78 | 83 | 83.00 | 60.79 | 1 | 165 | 164 | 0.00 | -1.22 | 3.72 |
| Condition* | 2 | 165 | 1.50 | 0.50 | 2 | 1.50 | 0.00 | 1 | 2 | 1 | -0.01 | -2.01 | 0.04 |
| Gender* | 3 | 165 | 1.61 | 0.49 | 2 | 1.63 | 0.00 | 1 | 2 | 1 | -0.43 | -1.83 | 0.04 |
| COWeek0 | 4 | 165 | 10.55 | 6.86 | 9 | 9.82 | 5.93 | 0 | 46 | 46 | 1.41 | 3.60 | 0.53 |
| COWeek4 | 5 | 165 | 5.46 | 5.10 | 4 | 4.69 | 4.45 | 0 | 24 | 24 | 1.24 | 0.96 | 0.40 |
| COWeek8 | 6 | 165 | 5.53 | 5.96 | 3 | 4.55 | 2.97 | 0 | 29 | 29 | 1.75 | 3.35 | 0.46 |
| QuitCigsWeek4* | 7 | 165 | 1.62 | 0.49 | 2 | 1.65 | 0.00 | 1 | 2 | 1 | -0.48 | -1.78 | 0.04 |
| QuitCigsWeek8* | 8 | 165 | 1.64 | 0.48 | 2 | 1.68 | 0.00 | 1 | 2 | 1 | -0.59 | -1.66 | 0.04 |
| DrinksWeek0 | 9 | 165 | 6.40 | 4.42 | 6 | 5.77 | 2.97 | 1 | 35 | 34 | 2.44 | 10.46 | 0.34 |
| DrinksWeek4 | 10 | 165 | 3.59 | 2.98 | 3 | 3.30 | 2.97 | 0 | 13 | 13 | 0.80 | 0.30 | 0.23 |
| DrinksWeek8 | 11 | 165 | 3.23 | 2.68 | 3 | 2.97 | 2.97 | 0 | 13 | 13 | 0.89 | 0.67 | 0.21 |
| CigsWeek0 | 12 | 165 | 14.22 | 8.22 | 12 | 13.05 | 5.93 | 3 | 51 | 48 | 1.59 | 3.20 | 0.64 |
| CigsWeek4 | 13 | 165 | 4.18 | 5.70 | 2 | 3.08 | 2.97 | 0 | 41 | 41 | 3.01 | 13.13 | 0.44 |
| CigsWeek8 | 14 | 165 | 3.16 | 4.75 | 2 | 2.25 | 2.97 | 0 | 35 | 35 | 3.44 | 16.81 | 0.37 |

DESCRIPTIVE STATISTICS BY GROUP

- = data frame name
- = variable name
- = grouping variable

```
# DESCRIPTIVE STATISTICS BY GROUP ----  
  
# summary statistics separately by group (psych package)  
describeBy(COWeek8 ~ Condition, data = ClinicalTrial)
```

R OUTPUT

Descriptive statistics by group

Condition: Varenicline

| | vars | n | mean | sd | median | trimmed | mad | min | max | range | skew | kurtosis | se |
|---------|------|----|------|------|--------|---------|------|-----|-----|-------|------|----------|------|
| COWeek8 | 1 | 82 | 5.02 | 4.88 | 3.5 | 4.33 | 3.71 | 0 | 26 | 26 | 1.57 | 3.09 | 0.54 |

Condition: Varenicline + Naltrexone

| | vars | n | mean | sd | median | trimmed | mad | min | max | range | skew | kurtosis | se |
|---------|------|----|------|------|--------|---------|------|-----|-----|-------|------|----------|------|
| COWeek8 | 1 | 83 | 6.02 | 6.86 | 3 | 4.81 | 4.45 | 0 | 29 | 29 | 1.62 | 2.31 | 0.75 |

INDEPENDENT-SAMPLES T-TEST

- = data frame name
- = variable name
- = grouping variable

- The mean difference is computed by subtracting groups in alphabetical order (e.g., Varenicline - Varenicline + Naltrexone)

```
# T-TEST WITH TWO-TAILED ALTERNATE HYPOTHESIS ----
```

```
# independent t-test with default two-tailed alternate hypotheses (base R)
results <- t.test(COWeek8 ~ Condition, data = ClinicalTrial)
results
```

```
# print standard error
cat('standard error of mean difference:', results$stderr)
```

R OUTPUT

```
data: COWeek8 by Condition  
t = -1.0802, df = 148.19, p-value = 0.2818  
alternative hypothesis: true difference in means between group Varenicline and  
group Varenicline + Naltrexone is not equal to 0
```

t-statistic and p-value

```
95 percent confidence interval:  
-2.8285605 0.8291482
```

95% confidence interval for mean difference

```
sample estimates:  
mean in group Varenicline mean in group Varenicline + Naltrexone  
5.024390 6.024096
```

Means

```
standard error of mean difference: 0.9254862
```

Standard error of mean difference

STANDARDIZED MEAN DIFFERENCE

- = data frame name
- = variable name
- = grouping variable

```
# STANDARIZED MEAN DIFFERENCE EFFECT SIZE ----
```

```
# standardized mean difference effect size (psych package)
cohen.d(COWeek8 ~ Condition, data = ClinicalTrial)
```

R OUTPUT

Cohen d statistic of difference between two means

lower effect upper

COWeek8 -0.14 0.17 0.47

Multivariate (Mahalanobis) distance between groups

[1] 0.17

r equivalent of difference between two means

COWeek8

0.08



SMALL GROUP EXERCISE

Download two files from Bruin Learn: "Week 7 Lab. Independent t-Test.R" and "Week 7 Small Group Exercise.R". The Lab script contains the R code we just discussed. The Exercise script contains only the URL for a different data set, CancerData.csv. In groups of two or three, you will complete a series of R tasks that provide practice for the next assignment. There is no need to write code from scratch; instead, you can copy and paste code chunks from the Lab file into your Exercise script, modifying the data and variable names as needed. The CancerData.csv file for this exercise contains data from study investigating the impact of a cancer diagnosis on psychological outcomes like depression.

UVEAL MELANOMA AND DEPRESSION

Uveal melanoma, a rare eye cancer, presents potential vision loss and life threat. This prospective, longitudinal study interrogated the predictive utility of visual impairment, as moderated by optimism/pessimism, on depressive symptoms in 299 adults undergoing diagnostic evaluation.

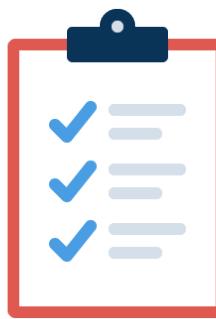


Annette
Stanton

James
MacDonald

MacDonald, J.J., Jorge-Miller, A., Enders, C.K., McCannel, T., Beran, T., & Stanton, A.L. (2021). Perceived and objective visual impairment predicting depressive symptoms across one year in uveal melanoma diagnostic biopsy: Optimism and pessimism as moderators. *Health Psychology, 40*, 408-417.

KEY VARIABLES



Depression

The CES-D is a 20-item inventory that asks people to rate how often they experience depressive symptoms such as restless sleep, poor appetite, and feeling lonely.



Cancer Diagnosis

Based on a diagnostic clinical evaluation for a possible intraocular malignancy, participants were classified as having malignant or nonmalignant diagnoses.



SMALL GROUP EXERCISE TASK 1

- Use the provided URL to import the CancerData.csv file into an R data frame (import method #3 from the Week 0 lab script).
- Use the dfSummary function to get numeric and visual summaries of the data frame's variables.



SMALL GROUP EXERCISE TASK 2

- Use the describeBy function to get descriptive statistics for the Depression dependent variable within each of the two Diagnosis groups (Non-Malignant vs. Malignant).



SMALL GROUP EXERCISE TASK 3

- State and justify the hypotheses. Clearly write out the null hypothesis (H_0) and the alternative (research) hypothesis (H_1) in both statistical notation and plain language. Explain why a two-tailed test is appropriate for this study, even if the expected direction of change might seem obvious.
- Use the `t.test` function to perform an independent-samples t-test to determine whether depression levels differ between participants to received positive versus negative diagnoses.



SMALL GROUP EXERCISE TASK 4

- Interpret the standard error of the mean difference. Explain what this value tells you about the precision of your estimated mean difference. How does it relate to the concept of sampling variability in repeated studies.
- Explain what the magnitude of the t-value tells you about how far the observed mean difference is from the null hypothesis.
- Explain what the magnitude of the p-value tells you about how far the observed mean difference is from the null hypothesis.



SMALL GROUP EXERCISE TASK 5

- Interpret the 95% confidence interval for the mean difference. Use the numeric values in your explanation. Relate your interpretation to the hypothesis test results—does the CI include the null hypothesis mean of 0?
- Provide an interpretation of the standardized mean difference effect size. Explain what the numeric value means and how it compares to Cohen's "off-the-shelf" benchmarks.