

Comparing Paired (dependent) group's quantitative data

Paired samples t-test &
Wilcoxon-signed rank

Prepared by Dr. Hesham Mousli

MHA 613
Statistics in Healthcare

Lecture 8



Intended Objectives

1. Grasp the importance of comparing paired quantitative data in healthcare.
2. Identify suitable tests for paired data, like paired t-test and Wilcoxon rank test.
3. Interpret test results, focusing on p-values and significance.
4. Recognize the assumptions and limitations of each test.

Paired samples t-Test



Is there a **difference** in a
group between **two**
points in time

Indication of Use

- We have **two samples** that are **related to each other** and **one Quantitative continuous variable** of interest.
- The variable may be measured on each individual in **two circumstances**.

Explaining Example

- **For example,** in a **cross-over trial**, each patient has **two measurements** on the variable, **one while taking active treatment** and **one while taking a placebo**.
- Such data are known as **paired (Dependent) data**.

In a Crossover Study

EACH PATIENT RECEIVES BOTH TREATMENTS

ARM A

ARM A

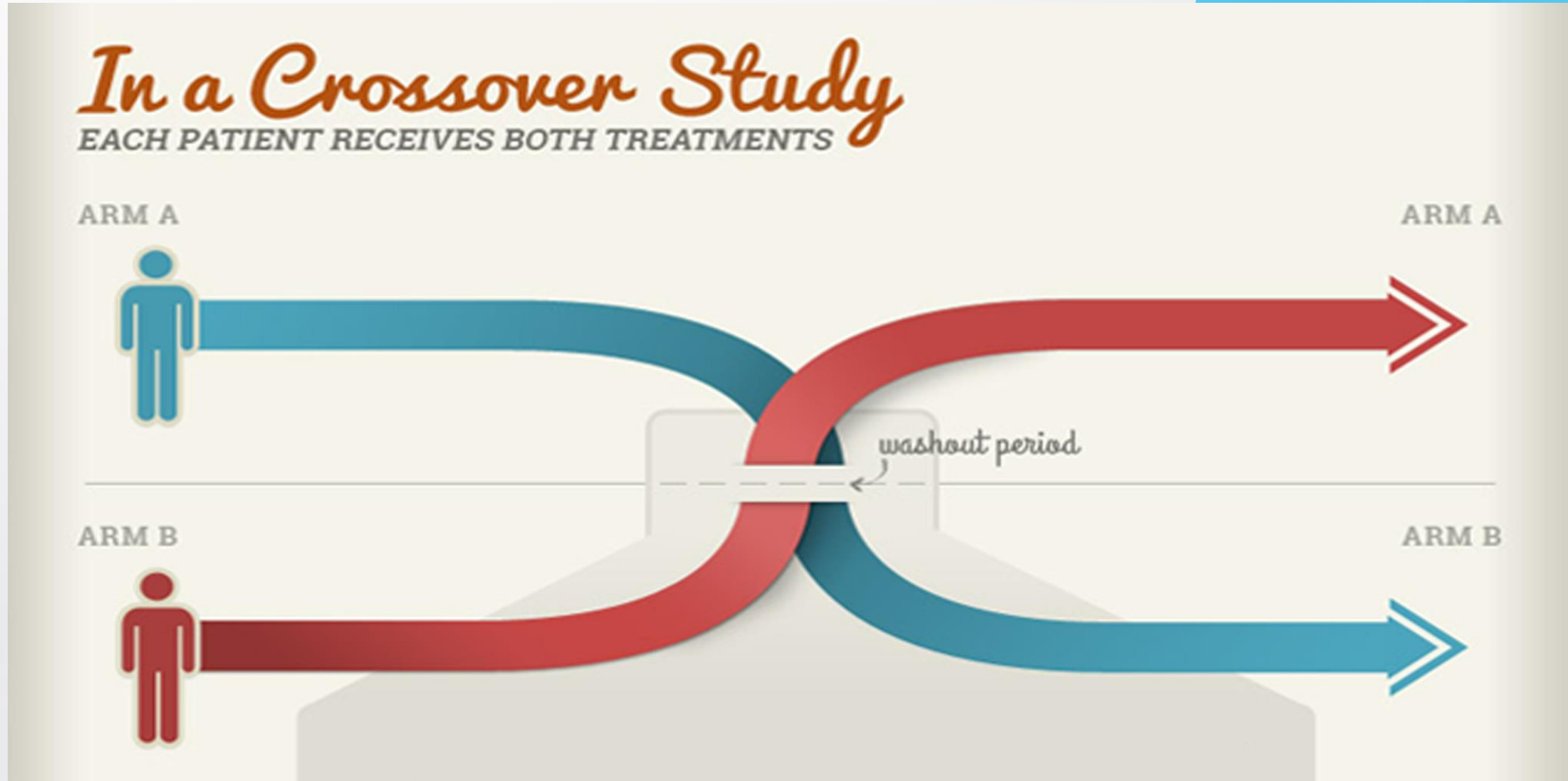


ARM B

ARM B



washout period



Explaining Example

- **Example 2:** a researcher is interested in testing the difference in **Blood Glucose Level (BGL)** for the same group of people after one month and then after two months.



- The Hypothesis test that can solve this problem is either **(Paired samples) t-test OR Wilcoxon-signed rank test.**



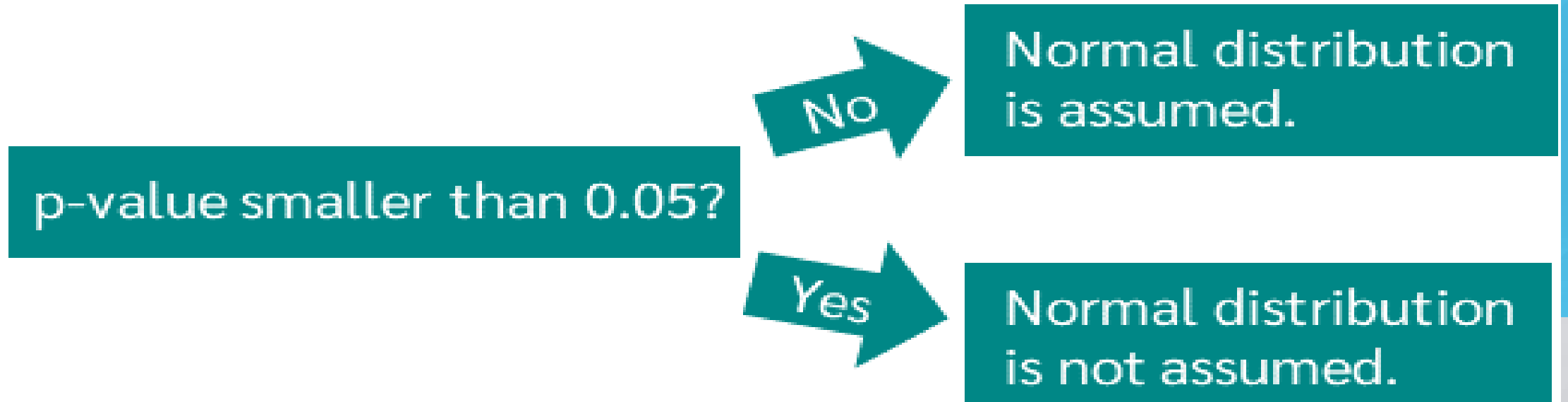
- I can select either of them based on **the Normality status of differences between the repeated measurements.**



- So, I must **first check the Normality of differences between the repeated measurements.**



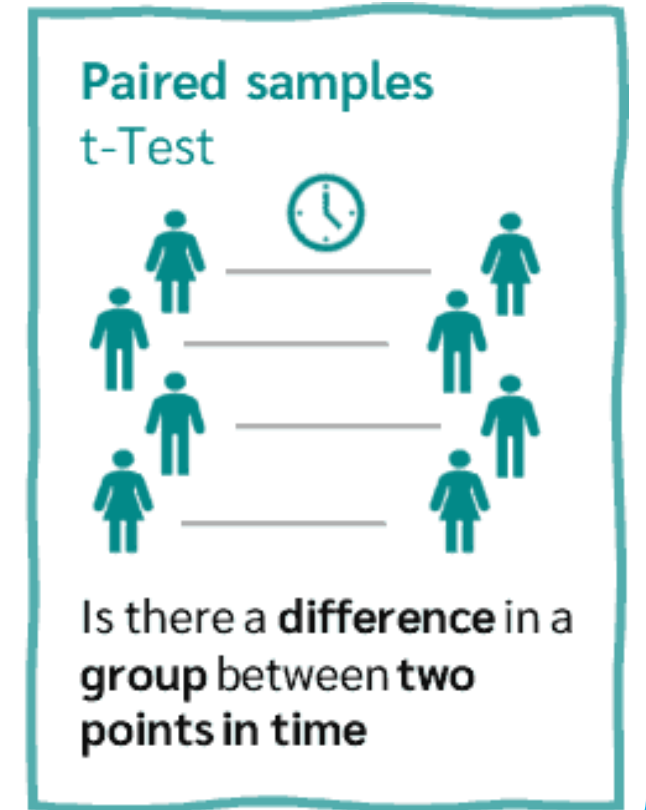
	Statistics	p-Value
Kolmogorov-Smirnov	0.09	0.892
Shapiro-Wilk	0.98	0.862
Anderson-Darling	0.24	0.772



1. The (Paired samples) t-test

➤ Assumptions:

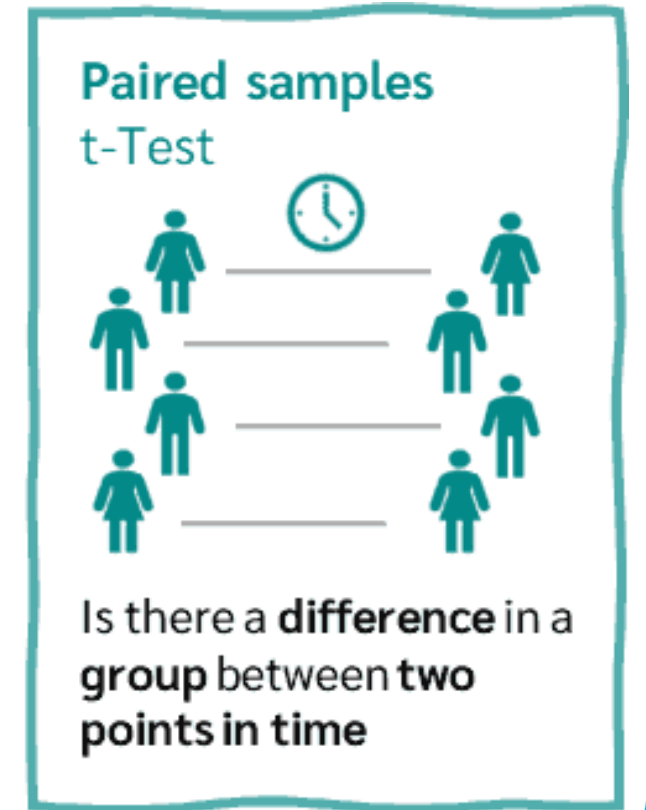
- **Paired Data:** Observations in each group are **paired in a meaningful way** (e.g., measurements **before and after** a treatment on the same subjects).



1. The (Paired samples) t-test

➤ Assumptions:

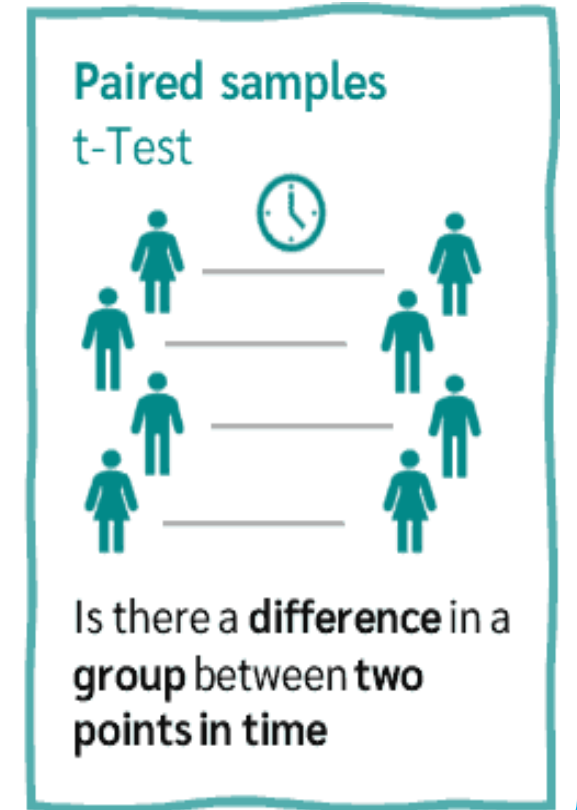
- **Continuous Data:** The differences between pairs should be continuous.
- **Normality of Differences:** The differences between paired observations are approximately normally distributed.



1. The (Paired samples) t-test

➤ Rationale:

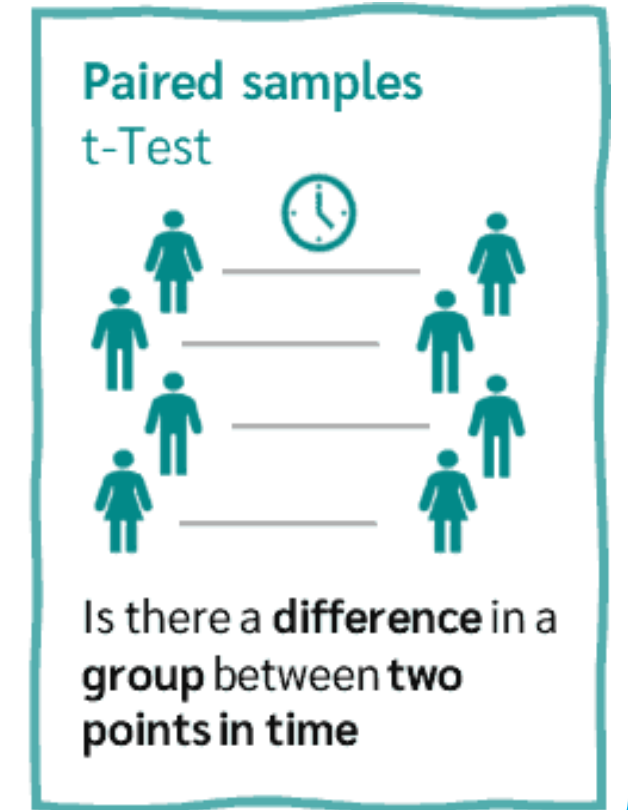
- The **paired samples t-test** is designed to **test whether the mean difference** between paired observations is **significantly different from zero**.



1. The (Paired samples) t-test

➤ Rationale:

- Under the **alternate hypothesis**, the **mean difference** between the **paired observations** is **significantly different from zero**.
- This **implies** that there is a **statistically significant difference** between the **population means** of the **two paired (related) groups**.



$p < 0.05 \rightarrow$ statistically
significant

Difference between the compared
paired samples data

$p > 0.05 \rightarrow$ no statistically
significant

Difference between the compared
paired samples data

Conclusion

- It will be in terms of **Mean & SD** reporting for each time point measurement.
- If **P-value < 0.05** so, there is **a statistically significant difference** between the BGL of patients after one month (Mean=, SD=) and their BGL after 3 months (Mean=, SD=), (P-value =).

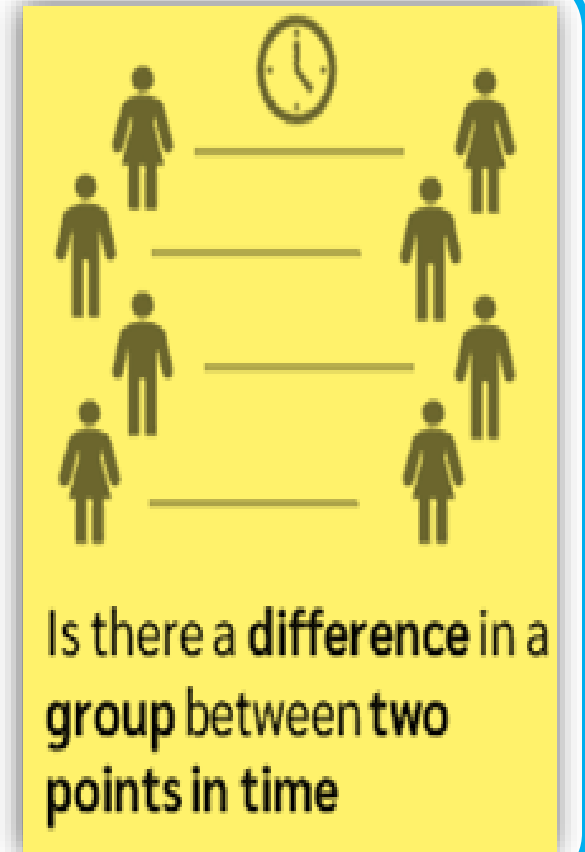
Conclusion

- If **P-value > 0.05** so, there is **No** statistically significant **difference** between the BGL of patients after one month (Mean=, SD=) and their BGL after three months (Mean=, SD=), (P-value =).

2. The Wilcoxon-signed rank test

➤ Rationale:

- The **Wilcoxon signed-rank test** is a **non-parametric alternative** to the paired samples t-test, used **when the data does not meet the normality assumption**.



$p < 0.05 \rightarrow$ statistically
significant

Difference between the compared
paired samples data

$p > 0.05 \rightarrow$ no statistically
significant

Difference between the compared
paired samples data

Conclusion

- It will be in terms of **Median & Range/IQR** reporting for each time point measurement.
- If **P-value < 0.05** so, there is **a statistically significant difference** between the BGL of patients after one month (Median=, Range=), and their BGL after 3 months (Median=, Range=), (P-value =).

Conclusion

- If **P-value > 0.05** so, there is **No statistically significant difference** between the BGL of patients after one month (Median=, Range=) and their BGL after 3 months (Median=, Range=), (P-value =).



Please don't forget: Before
dealing with any
Quantitative data, first,
test for its normality.

Case Scenario (1)

- **Comparing Costs of Medication Management Before and After Centralizing Pharmacy Services:**
 - **Scenario:** A hospital centralizes its pharmacy services to reduce costs associated with medication management. Monthly medication costs for the same patients are compared for periods before and after centralization.

Case Scenario (1)

- **Null Hypothesis (H_0):** There is no difference in monthly medication costs **before** and **after** centralizing pharmacy services.
- **Alternative Hypothesis (H_1):** There is a difference in monthly medication costs **before** and **after** centralizing pharmacy services.

T	 Cost (\$) Before	 Cost (\$) After
1	33.4	30.8
2	36	35.4
3	35.4	30
4	22.5	25.3
5	33.2	22.4
6	42.2	32.4
7	19.3	42.7
8	22.5	22.5
9	31.5	33.2



Difference

-2.6

12.9

-3.2

-16.9

3.1

9.9

11.2

-0.2

-6.8

Descriptive Statistics

Descriptive Statistics

	Difference
Valid	36
Missing	0
Mean	-16.597
Std. Deviation	21.728
Shapiro-Wilk	0.892
P-value of Shapiro-Wilk	0.002
Minimum	-63.400
Maximum	18.700

The screenshot displays a statistical software interface with a top navigation bar and a main menu. The top bar contains icons for 'Edit Data' (spreadsheet), 'Descriptives' (bar chart), 'T-Tests' (T-test icon), 'ANOVA' (ANOVA icon), and 'Mixed' (mixed model icon). An orange arrow points from the 'ANOVA' icon to the 'T-Tests' tab. Below the top bar, the 'Descriptive Statistics' section is expanded, showing a list of variables: 'No-show (Before)', 'No-show (After)', and 'Difference'. To the right, the 'T-Tests' menu is open, showing two categories: 'Classical' and 'Bayesian'. Each category lists 'Independent Samples T-Test', 'Paired Samples T-Test', and 'One Sample T-Test'. An orange arrow points from the 'Paired Samples T-Test' option in the 'Classical' section to the right.

Top navigation bar: Edit Data, Descriptives, T-Tests, ANOVA, Mixed

Descriptive Statistics:

- No-show (Before)
- No-show (After)
- Difference

T-Tests menu:

- Classical**
 - Independent Samples T-Test
 - Paired Samples T-Test
 - One Sample T-Test
- Bayesian**
 - Independent Samples T-Test
 - Paired Samples T-Test
 - One Sample T-Test

▼ Paired Samples T-Test

- Cost (\$) Before
- Cost (\$) After
- Difference

↓
A
Z



Variable Pairs

Cost (\$) After

Cost (\$) Before

Tests

- 
- ☒ Student
 - ☐ Wilcoxon signed-rank

Alternative Hypothesis

- ☒ Measure 1 $\neq$ Measure 2
- ☐ Measure 1 $>$ Measure 2
- ☐ Measure 1 $<$ Measure 2

Additional Statistics

- ☐ Location parameter
 - ☐ Confidence interval %
- ☐ Effect size
 - ☐ Confidence interval %
 - ☐ Correct for correlation
- ☐ Descriptives
- ☐ Vovk-Sellke maximum p-ratio

Paired Samples T-Test ←

Paired Samples T-Test

Measure 1	Measure 2	t	df	p	←
Cost (\$) After	- Cost (\$) Before	-0.038	35	0.970	←

Note. Student's t-test. ←

Case Scenario (1)

- With a **p-value of 0.970**, which is **much higher than the typical significance level (0.05)**, Since the test showed **no significant cost reduction**, the hospital may **re-evaluate centralizing pharmacy services, consider alternative cost-saving strategies, or conduct further analysis to improve the system.**



**Now, let's go to the 2nd
Case Scenario**

Case Scenario (2)

- **Evaluating the Impact of a New Digital Appointment Scheduling System on Patient No-Show Rates:**
 - A clinic implements a digital appointment scheduling system designed to reduce patient no-show rates. The no-show rates are measured for the same set of patients before and after the implementation.

Case Scenario (2)

- **Null Hypothesis (H_0):** There is no difference in patient no-show rates **before** and **after** implementing the digital scheduling system.
- **Alternative Hypothesis (H_1):** There is a difference in patient no-show rates **before** and **after** implementing the digital scheduling system.

T	 No-show (Before)	 No-show (After)	 Difference
1	3.5	1.2	2.3
2	4.7	2.1	2.6
3	6.5	1.3	5.2
4	4.2	1.3	2.9
5	7.3	3.2	4.1
6	4.5	1.2	3.3
7	8.2	3.2	5
8	4.9	1.4	3.5

Descriptive Statistics

Descriptive Statistics

	Difference
Valid	36
Missing	0
Mean	-16.597
Std. Deviation	21.728
Shapiro-Wilk	0.802
P-value of Shapiro-Wilk	0.002
Minimum	-63.400
Maximum	18.700

The screenshot displays a statistical software interface. At the top, there is a navigation bar with icons for a menu, data editing, descriptive statistics, T-Tests, ANOVA, and Mixed Models. The 'T-Tests' tab is selected, and an orange arrow points to it. Below the navigation bar is a data table with three columns: a filter icon, 'No-show (Before)', and 'No-show (After)'. The table contains 9 rows of data. A dropdown menu is open under the 'T-Tests' tab, showing two sections: 'Classical' and 'Bayesian'. Each section lists three options: 'Independent Samples T-Test', 'Paired Samples T-Test', and 'One Sample T-Test'. An orange arrow points to the 'Paired Samples T-Test' option in the 'Classical' section.

	No-show (Before)	No-show (After)	
1	3.5	1.2	
2	4.7	2.1	
3	6.5	1.3	
4	4.2	1.3	
5	7.3	3.2	
6	4.5	1.2	
7	8.2	3.2	
8	4.9	1.4	3.5
9	6.3	2.4	3.9

T-Tests

- Classical**
 - Independent Samples T-Test
 - Paired Samples T-Test
 - One Sample T-Test
- Bayesian**
 - Independent Samples T-Test
 - Paired Samples T-Test
 - One Sample T-Test

▼ Paired Samples T-Test



- No-show (Before)
- No-show (After)
- Difference



Variable Pairs

No-show (Before)

No-show (After)



Tests

- ☐ Student
- ☒ Wilcoxon signed-rank

Additional Statistics

- ☐ Location parameter
- ☐ Confidence interval 95.0 %



Paired Samples T-Test

Paired Samples T-Test

Measure 1	Measure 2	W	z	df	p	
No-show (Before)	- No-show (After)	461.500	2.400		0.017	

Note. Wilcoxon signed-rank test.

Case Scenario (2)

- With a **p-value of 0.017**, we **reject the null hypothesis**, indicating a **significant reduction in no-show rates**.
- The clinic **should consider fully adopting the digital scheduling system**, as the evidence suggests it **effectively reduces patient no-show rates**.

Wrap up

1. Paired tests like the paired samples t-test or Wilcoxon-signed rank test are essential for comparing dependent data, such as pre- and post-treatment measurements.
2. Correct test selection (parametric vs. non-parametric) depends on the normality of the data.
3. Practical applications include scenarios like assessing changes in patient health metrics over time or evaluating healthcare interventions.
4. Interpreting results with a focus on p-values allows healthcare managers to make data-informed decisions about treatment effectiveness and operational improvements.



Thank You