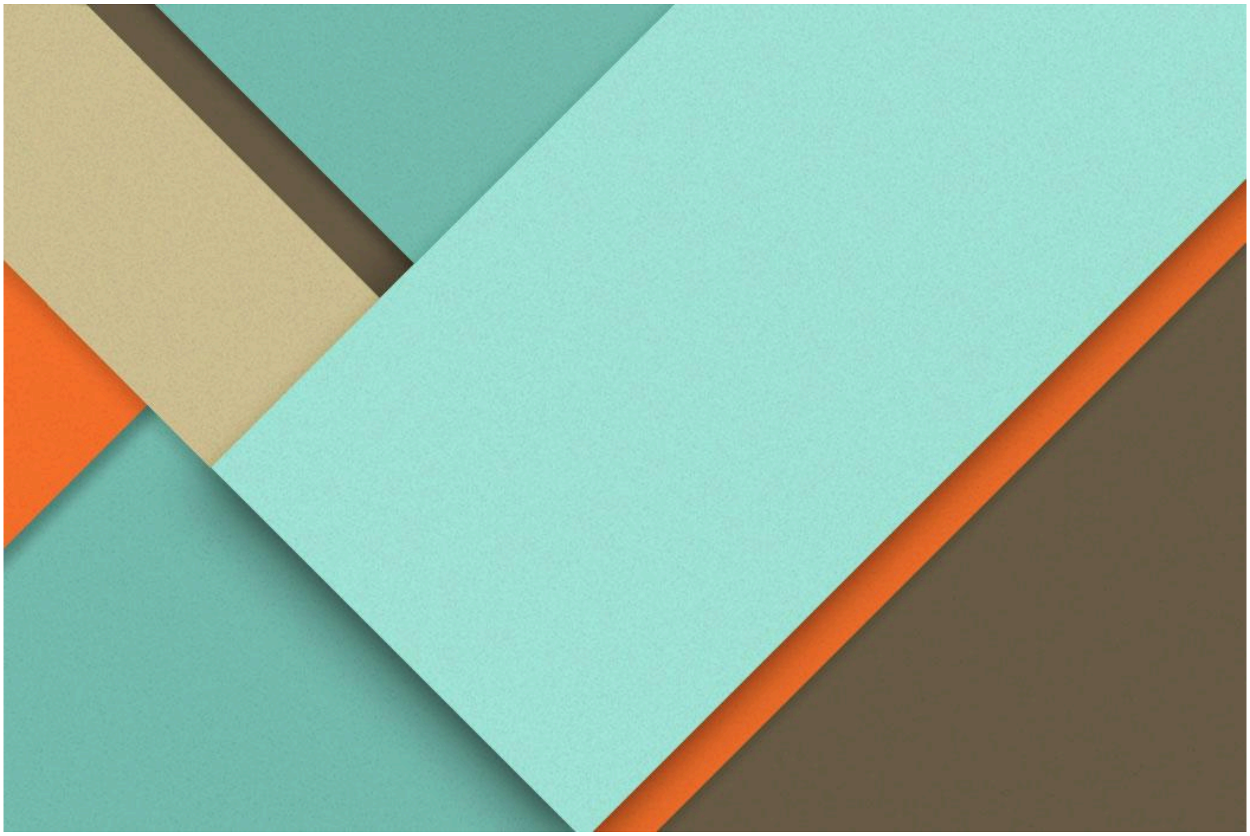


Meniscal Injury Analysis

23.01.2025

Emir Doğan
Data Analyst

Overview

Meniscus injuries are among the most common knee injuries, significantly affecting patients' mobility and quality of life. This project aims to analyze a comprehensive dataset of 597 patients who underwent arthroscopic meniscus injury repair.

Goals

1. Examine the distribution of injuries by gender, age, and injury location (right knee, left knee, or bilateral).
2. Evaluate the prevalence of different injury types, such as medial, lateral, and combined meniscus injuries.
3. Identify relationships between injury types such as medial, lateral and bilateral combined.
4. Develop comprehensive visualizations (e.g., bar charts, pie charts…) to present findings clearly and intuitively.

Dataset Source Information

The dataset used for this analysis originates from a retrospective study conducted on 597 patients who underwent arthroscopic meniscus injury repair at the Department of Joint Surgery, the First Affiliated Hospital of the University. The study period spans from January 2013 to May 2018.

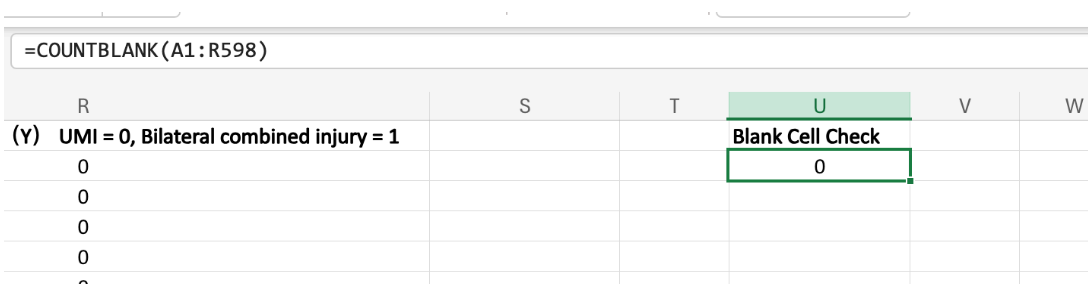
Citation:

Jiang, Peishi (2019). *Meniscal injury (Arthroscopy)*. figshare. Dataset. https://doi.org/10.6084/m9.figshare.11309312.v1

Data Cleaning and Consistency Check

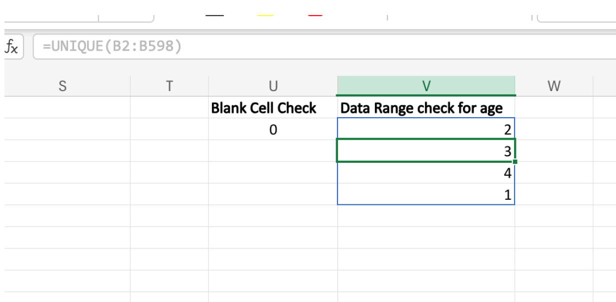
Blank Cell Check:

A comprehensive check for missing values was conducted using the COUNTBLANK function in Excel. The result returned 0, confirming that the dataset contains no missing or blank cells.



Verification of Value Ranges for the Age Column:

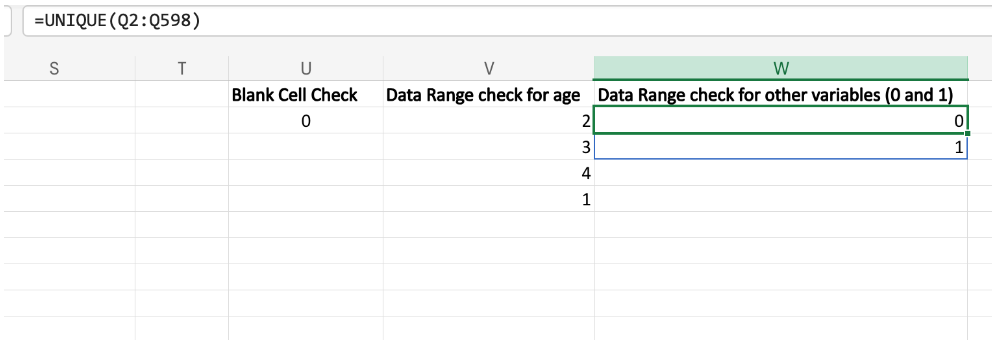
The range of values in the Age column was examined using the `UNIQUE` function in Excel. The values observed were within the expected range of 1 to 4, corresponding to predefined categorical groups. This result confirms that the column adheres to the expected format.

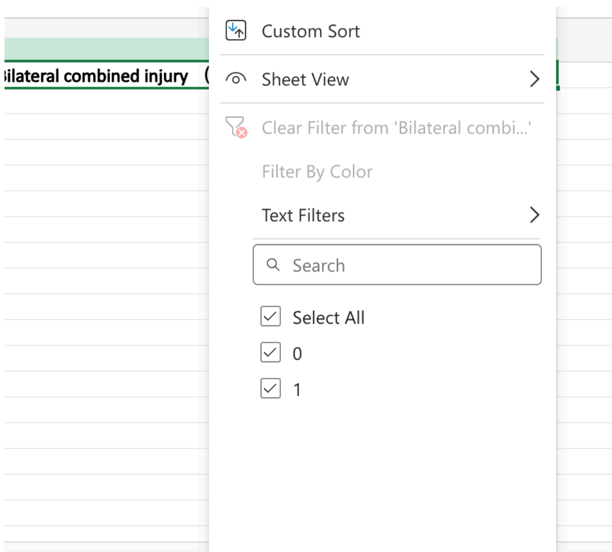


Consistency Validation for Binary Columns:

All remaining columns were verified to contain binary values (`0` and `1`) as expected. This validation was performed using the following methods:

- The `UNIQUE` function to confirm the range of values.
- Filter and selection tools to manually cross-check the binary consistency.
  The results demonstrated that all binary columns maintained the required value consistency.





Descriptive Analysis

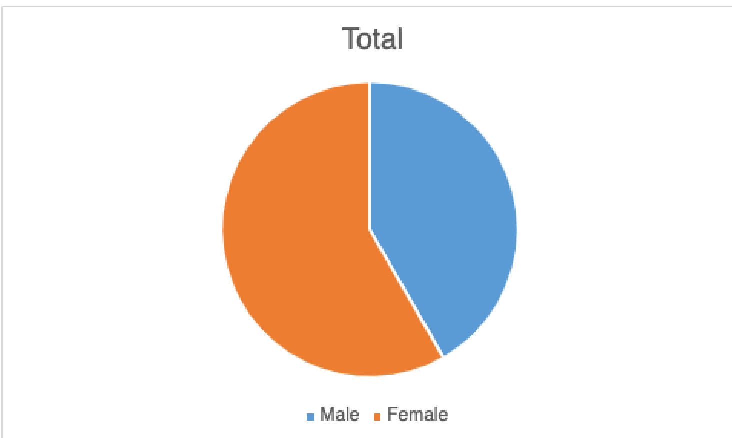
Gender Distribution:

A `PivotTable` was used to calculate the frequency of male (0) and female (1) patients in the dataset. The percentage distribution was computed using the PivotTable again.

This step helped determine the proportion of male and female patients within the dataset. Later pie chart has been inserted with selected male and female percentages.

Gender(X1) Male=0, female=1	Count of Gender(X1) Male=0, female=1
Male	41.71%
Female	58.29%
Grand Total	**100.00%**

Gender(X1) Male=0, female=1	Count of Gender(X1) Male=0, female=1
Male	249
Female	348
Grand Total	**597**



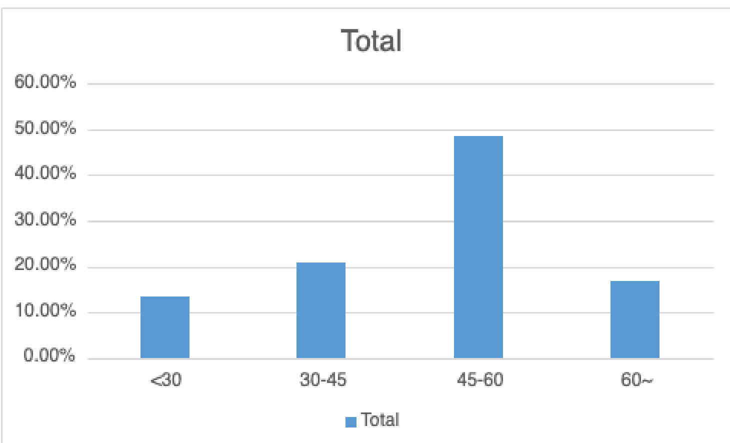
The analysis of the gender distribution reveals that the dataset is composed of 58.29% female patients and 41.71% male patients. This indicates that female patients represent a majority in this study population. The gender ratio (female:male) is approximately **1.4:1**, which suggests that meniscus injuries may have a slightly higher prevalence or treatment rate among females in this sample.

Age Group Distribution:

Using a `PivotTable` the frequency of patients in each age group was calculated. The percentage for each age group was calculated using the PivotTable again. With the calculated percentages the column chart's inserted.

Age (years old) (X2) <30=1, 30-45=2, 45-60=3, 60~ =4	Count of Age (years old) (X2) <30=1, 30-45=2, 45-60=3, 60~ =4
<30	13.40%
30-45	20.94%
45-60	48.74%
60~	16.92%
Grand Total	100.00%

Age (years old) (X2) <30=1, 30-45=2, 45-60=3, 60~ =4	Count of Age (years old) (X2) <30=1, 30-45=2, 45-60=3, 60~ =4
<30	80
30-45	125
45-60	291
60~	101
Grand Total	597



The majority of the patients (48.74%) fall within the 45-60 years age group, indicating that middle-aged individuals are most affected by meniscus injuries in this dataset.

This distribution suggests that meniscus injuries predominantly occur in middle-aged individuals, with a noticeable decline in cases among younger and older populations. This insight may reflect activity levels, wear and tear, or age-related vulnerability to knee injuries.

Distribution of Meniscal Injury Types:

In the dataset, there are three types of Medial Injuries: `Medial Posterior Injury`, `Medial Body Injury`, and `Medial Anterior Injury`. To analyze the overall prevalence of Medial Injuries, the three types were aggregated into a single column representing whether a patient has any type of Medial Injury. Lateral Injuries also have three types: Lateral Posterior Injury, Lateral Body Injury, and Lateral Anterior Injury. These were aggregated into a single column representing whether a patient has any type of Lateral Injury.

```
=IF(OR(L2=1, M2=1, N2=1), 1, 0)
```

X	Y	Z
		Total Medial Injury
		1
		1
		1
		1
		1
		0
		1
		1
		1
		1
		1
		1
		1
		1

```
=IF(OR(O2=1, P2=1, Q2=1), 1, 0)
```

AC	
Total Lateral Injury	Sum the Total
0	
0	
0	
0	
0	
1	
1	
0	

Combined injuries are already represented in the dataset under the column Bilateral Combined Injury. This step focuses on calculating all of the injuries prevalence.

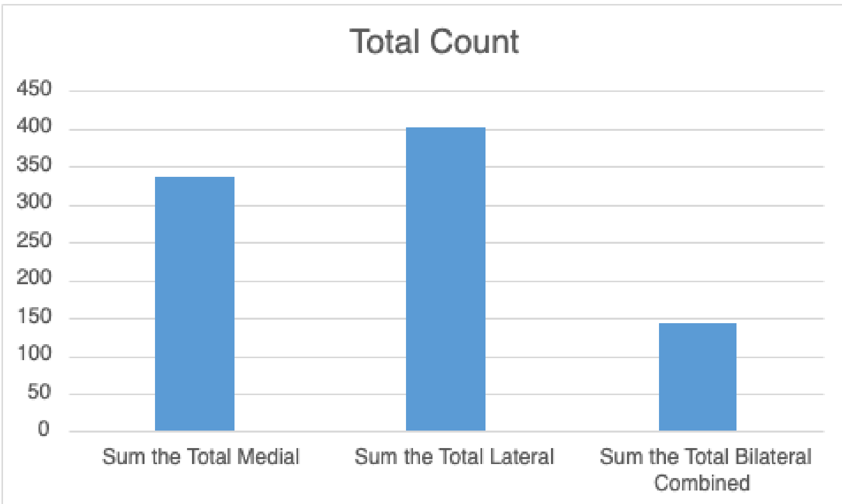
```
=SUM(Z2:Z1000) / COUNT(Z2:Z598)
```

AB	AC	AD	AE	AF	AG
Prevalence of Medial Injuries	Total Lateral Injury	Sum the Total Lateral	Prevalence of Lateral Injuries	Sum the Total Bilateral Combined	Prevalence of Bilateral Combined Injuries
0.563	0	401	0.672	143	0.240
	0				
	0				
	0				
	0				
	1				
	1				
	0				
	0				
	1				

To understand the distribution of injury types, we analyzed the prevalence of medial, lateral, and bilateral combined injuries in the dataset. The results are as follows:

- Medial Injuries: Represent 56.3% of all cases, making them a highly common type of meniscal injury.
- Lateral Injuries: The most prevalent type, occurring in 67.2% of cases.
- Bilateral Combined Injuries: Represent 24.0% of cases, showing significantly lower prevalence compared to medial and lateral injuries.

The column chart below illustrates the prevalence of different meniscal injury types:



Chi-Square Test: Relationship Analysis

Chi-Square Test Between Injury Types:

A chi-square test was conducted to assess the independence of the following injury types:

- Medial Injuries (Total Medial).
- Lateral Injuries (Total Lateral).
- Bilateral Combined Injuries.

Observed and expected frequencies were calculated using a contingency table. Excel's CHISQ.TEST function was used to compute the chi-square statistic and p-value.

Contingency tables were created for the following pairs:

- Medial and Lateral injuries.
- Medial and Bilateral injuries.
- Lateral and Bilateral injuries.

(Excel's COUNTIFS function was used to calculate observed frequencies for each combination.)

```
=COUNTIFS(S2:S598,0,T2:T598,0)
```

Contingency Tables

Medial vs Lateral	Lateral 0	Lateral 1	**Total**
Medial 0	2	259	261
Medial 1	194	142	336
Total	196	401	**597**

Medial vs Bilateral	Bilateral 0	Bilateral 1	**Total**
Medial 0	260	1	261
Medial 1	194	142	336
Total	454	143	**597**

Lateral vs Bilateral	Bilateral 0	Bilateral 1	**Total**
Lateral 0	195	1	196
Lateral 1	259	142	401
Total	454	143	**597**

Also Expected Frequency tables were created for the following pairs:

- Medial and Lateral injuries.
- Medial and Bilateral injuries.
- Lateral and Bilateral injuries.

(Expected frequencies were calculated using the formula: Expected Frequency = (Row Total * Column Total) / Grand Total) fx =(AM5*AK7)/AM7

	AI	AJ	AK	AL	AM	AN
		Expected Frequencies				
		Medial vs Lateral	Lateral 0	Lateral 1	**Total**	
		Medial 0	85.688442	175.3115578	261	
		Medial 1	110.31156	225.6884422	336	
		Total	196	401	**597**	
		Medial vs Bilateral	Bilateral 0	Bilateral 1	**Total**	
		Medial 0	198.48241	62.51758794	261	
		Medial 1	255.51759	80.48241206	336	
		Total	454	143	**597**	
		Lateral vs Bilateral	Bilateral 0	Bilateral 1	**Total**	
		Lateral 0	149.05193	46.9480737	196	
		Lateral 1	304.94807	96.0519263	401	
		Total	454	143	**597**	

Excel's CHISQ.TEST function was used to compare observed and expected frequencies.

Chi-Square test for Medial vs Lateral
6.07202E-49
Chi-Square test for Medial vs Bilateral
1.29396E-32
Chi-Square test for Lateral vs Bilateral
6.42412E-21

Findings:

Medial vs. Lateral:

- p-value < 0.05
- Several observed frequencies in the contingency table were below 5, violating a critical assumption of the Chi-Square test.
- Additionally, there were substantial discrepancies between observed and expected frequencies, which may have inflated the test statistic and rendered the result potentially unreliable.

Medial vs. Bilateral:

- p-value < 0.05
- Similar to the first pair, observed frequencies below 5 and large differences between observed and expected values raise concerns about the validity of this result.

Lateral vs. Bilateral:

- p-value < 0.05
- The test indicated a highly significant relationship, but the presence of low observed frequencies (<5) and large deviations between observed and expected frequencies suggest that the results may be misleading.

Alternative Method:

Fisher's Exact Test: For datasets with low observed frequencies, Fisher's Exact Test is a more appropriate statistical method, as it does not rely on large-sample assumptions.

Fisher's Exact Test

Fisher's Exact Test for Relationships Between Injury Types Using R:

Due to the limitations of the Chi-Square test (e.g., low observed frequencies in some cells), Fisher's Exact Test was performed using R to ensure greater reliability.

The fisher.test() function in R was applied to the following 2x2 contingency tables:

1. Medial vs. Lateral
2. Medial vs. Bilateral

3. Lateral vs. Bilateral

```r
# Medial vs. Lateral
table_medial_lateral <- matrix(c(2, 259, 194, 142),
                               nrow = 2,
                               byrow = TRUE)

# Fisher's Exact Test
fisher_test_medial_lateral <- fisher.test(table_medial_lateral)
fisher_test_medial_lateral

# Medial vs. Bilateral
table_medial_bilateral <- matrix(c(260, 1, 194, 142),
                                 nrow = 2,
                                 byrow = TRUE)

# Fisher's Exact Test
fisher_test_medial_bilateral <- fisher.test(table_medial_bilateral)
fisher_test_medial_bilateral

# Lateral vs. Bilateral
table_lateral_bilateral <- matrix(c(195, 1, 259, 142),
                                  nrow = 2,
                                  byrow = TRUE)

# Fisher's Exact Test
fisher_test_lateral_bilateral <- fisher.test(table_lateral_bilateral)
fisher_test_lateral_bilateral
```

Results:

```
data:  table_medial_lateral
p-value < 2.2e-16
alternative hypothesis: true odds ratio is not equal to 1
95 percent confidence interval:
 0.0007173073 0.0214093308
sample estimates:
 odds ratio
0.005706572
```

```
data:  table_medial_bilateral
p-value < 2.2e-16
alternative hypothesis: true odds ratio is not equal to 1
95 percent confidence interval:
   32.82051 7240.07882
sample estimates:
odds ratio
  189.2606
```

```
data:  table_lateral_bilateral
p-value < 2.2e-16
alternative hypothesis: true odds ratio is not equal to 1
95 percent confidence interval:
   18.44384 4172.56279
sample estimates:
odds ratio
  106.4773
```

1. Medial vs. Lateral:
    - P-value: < 2.2e-16
    - Odds Ratio (OR): 0.0057
    - Confidence Interval (95%): [0.0007, 0.0214]
    - Interpretation:
        - The extremely low p-value indicates a statistically significant relationship between Medial and Lateral injuries.
        - The odds ratio of 0.0057 (95% CI: [0.0007, 0.0214]) suggests a very strong negative association, meaning the likelihood of a Medial injury occurring together with a Lateral injury is extremely low.
2. Lateral vs. Bilateral:
    - P-value: < 2.2e-16
    - Odds Ratio (OR): 106.48
    - Confidence Interval (95%): [18.44, 4172.56]
    - Interpretation:
        - The extremely low p-value indicates a statistically significant relationship between Lateral and Bilateral injuries.

        - The odds ratio of 106.48 (95% CI: [18.44, 4172.56]) suggests a very strong positive association, meaning the presence of a Lateral injury greatly increases the likelihood of a Bilateral injury.
3. Medial vs. Bilateral:
    - P-value: < 2.2e-16
    - Odds Ratio (OR): 189.26
    - Confidence Interval (95%): [32.82, 7240.08]
    - Interpretation:
        - The extremely low p-value indicates a statistically significant relationship between Medial and Bilateral injuries.
        - The odds ratio of 189.26 (95% CI: [32.82, 7240.08]) suggests an extremely strong positive association, meaning Medial injuries are highly associated with Bilateral injuries.

Conclusion:

Fisher's Exact Test revealed statistically significant relationships between all injury type pairs (Medial vs. Lateral, Lateral vs. Bilateral, and Medial vs. Bilateral).

The results indicate:

1. Medial vs. Lateral: A very strong negative association, suggesting these injury types rarely co-occur.
2. Lateral vs. Bilateral: A very strong positive association, indicating these injuries frequently co-occur.
3. Medial vs. Bilateral: An extremely strong positive association, showing a high likelihood of co-occurrence.

Implications:

The findings highlight the distinct patterns of co-occurrence between injury types, which may provide valuable insights for further clinical or biomechanical studies.