Statistical Evaluation of Tommy John Surgery

Ethan Beffert and Chad Fox

University of Minnesota Duluth

Tommy John surgery is a growing issue happening with this new evolution of baseball. Velocity has become such an integral part of being a successful major league pitcher but it has its costs. The ulnar collateral ligament is a sensitive ligament that suffers from the motion of pitching a baseball. A “Tommy John” injury consists of rupturing this ligament due to overuse, not from one throw or one bad pitch. The number of these injuries appears to be going up exponentially every year. This could be due to a number of reasons but increased velocity in MLB pitchers can rapidly degrade the ligament in question. Others hypothesize that specializing in baseball too young and poor youth sports coaches are leading factors to this growing issue. Regardless of the cause it is possible, and in fact common, to come back and pitch in the major leagues after suffering this injury. This led us to our research question; How does the performing surgeon in a Tommy John surgery impact the postoperative pitching statistics and outcomes of major league baseball pitches? There are a lot of doctors who specialize in this specific surgery and we wanted to see who was the best. We hypothesized that the surgeon wouldn’t matter and that all surgeons would provide the same outputs.

This study's literature review and data was collected from multiple sources with two sources used very heavily. Those sources were baseballsavant.com and a Google Sheet maintained by baseball writer Jon Roegele, which documents professional baseball players who have undergone Tommy John surgery. The data gathered from Baseball Savant comprises a list of pitching metrics deemed important for the study, including strikeout percentage, walk percentage, average exit velocity of balls in play, hard-hit percentage, fastball rotations per minute, fastball velocity, earned run average, and walks plus hits per innings pitched. Additionally, information sourced from the Tommy John Google Sheet includes details such as a pitcher’s age at the time of surgery, the surgeon involved, and the recovery time between surgery and returning to Major League Baseball.

Each data group collected included the MLB ID number for each player, enabling the merging of the two groups into one using SQL. Moreover, rows of data lacking key information, such as surgeon and recovery time, were promptly excluded from the dataset, ensuring the analysis was conducted on clean, complete data. Following this, the data was exported to Excel for additional cleaning. This involved refining the dataset to include only the statistics of each pitcher from the year before their surgery and their return year from surgery. Subsequently, the statistics from the year before surgery were subtracted by those after surgery, resulting in 35 distinct rows representing 35 different pitchers. When looking at the data we also wanted to examine potential statistical differences between pitchers who have had Tommy John surgery and those who haven't. Baseball Savant's website was used again to compile an Excel sheet of 35 non-Tommy John pitchers. Differences between the two seasons were also taken from these pitchers to compile data. We hypothesized that the Tommy John pitchers would have a worse difference than non-Tommy John pitchers.

The initial methodology identified that age and surgeon would be good y variables used to build models where the x variables would be the differences between seasons of strike out percentage, walk percentage, average exit velocity, how often they are hit hard, fastball velocity and rpm, ERA, Whip, and recovery time. We chose to use age in order to see if that was affecting the statistics and not the surgeon being used. Regression models 1-9 show the relationship between the surgeon used and the 9 pitching statistics. Three surgeon categories were identified based on the 2 most used surgeons (Andrews and Elattache) with an “other” column used for the other surgeons. Regression models 10-18 would be the age of the pitcher predicting the difference in season-long variables from the season before injury to the season after. A regression output and bar graph were then created to amplify the statistically significant findings of our initial model output.

```
                            OLS Regression Results
==============================================================================
Dep. Variable:               kpercent   R-squared:                       0.081
Model:                            OLS   Adj. R-squared:                  0.024
Method:                 Least Squares   F-statistic:                     1.413
Date:                Mon, 15 Apr 2024   Prob (F-statistic):              0.258
Time:                        22:35:12   Log-Likelihood:                -99.710
No. Observations:                  35   AIC:                             205.4
Df Residuals:                      32   BIC:                             210.1
Df Model:                           2
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const         -2.8706      1.060     -2.709      0.011      -5.029      -0.712
andrews        2.8563      1.962      1.455      0.155      -1.141       6.854
neal           2.1706      1.691      1.284      0.208      -1.274       5.615
==============================================================================
Omnibus:                        0.204   Durbin-Watson:                   1.826
Prob(Omnibus):                  0.903   Jarque-Bera (JB):                0.410
Skew:                          -0.047   Prob(JB):                        0.815
Kurtosis:                       2.478   Cond. No.                         3.32
==============================================================================
```

```
                            OLS Regression Results
==============================================================================
Dep. Variable:              bbpercent   R-squared:                       0.030
Model:                            OLS   Adj. R-squared:                 -0.030
Method:                 Least Squares   F-statistic:                    0.5024
Date:                Mon, 15 Apr 2024   Prob (F-statistic):              0.610
Time:                        22:35:13   Log-Likelihood:                -91.857
No. Observations:                  35   AIC:                             189.7
Df Residuals:                      32   BIC:                             194.4
Df Model:                           2
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const          0.3412      0.847      0.403      0.690      -1.384       2.066
andrews       -1.5697      1.568     -1.001      0.324      -4.764       1.624
neal          -0.5230      1.351     -0.387      0.701      -3.275       2.229
==============================================================================
Omnibus:                        0.694   Durbin-Watson:                   2.456
Prob(Omnibus):                  0.707   Jarque-Bera (JB):                0.711
Skew:                           0.072   Prob(JB):                        0.701
Kurtosis:                       2.317   Cond. No.                         3.32
==============================================================================
```

Figure 1: Regression output surgeon predicting strikeout percentage (left) and walk percentage (right)

```
==============================================================================
Dep. Variable:                fastrpm   R-squared:                       0.023
Model:                            OLS   Adj. R-squared:                 -0.039
Method:                 Least Squares   F-statistic:                    0.3697
Date:                Mon, 15 Apr 2024   Prob (F-statistic):              0.694
Time:                        22:35:13   Log-Likelihood:                -220.00
No. Observations:                  35   AIC:                             446.0
Df Residuals:                      32   BIC:                             450.7
Df Model:                           2
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const         -3.4118     32.952     -0.104      0.918     -70.532      63.709
andrews       43.4118     61.015      0.711      0.482     -80.871     167.695
neal         -11.2246     52.573     -0.214      0.832    -118.312      95.863
==============================================================================
Omnibus:                       19.493   Durbin-Watson:                   1.720
Prob(Omnibus):                  0.000   Jarque-Bera (JB):               44.455
Skew:                          -1.146   Prob(JB):                     2.22e-10
Kurtosis:                       8.023   Cond. No.                         3.32
==============================================================================
```

```
==============================================================================
Dep. Variable:                hardhit   R-squared:                       0.091
Model:                            OLS   Adj. R-squared:                  0.034
Method:                 Least Squares   F-statistic:                     1.599
Date:                Mon, 15 Apr 2024   Prob (F-statistic):              0.218
Time:                        22:35:13   Log-Likelihood:                -112.40
No. Observations:                  35   AIC:                             230.8
Df Residuals:                      32   BIC:                             235.5
Df Model:                           2
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const          0.3294      1.523      0.216      0.830      -2.773       3.432
andrews       -4.5008      2.820     -1.596      0.120     -10.245       1.244
neal          -3.1567      2.430     -1.299      0.203      -8.106       1.793
==============================================================================
Omnibus:                        4.872   Durbin-Watson:                   1.906
Prob(Omnibus):                  0.088   Jarque-Bera (JB):                3.450
Skew:                          -0.567   Prob(JB):                        0.178
Kurtosis:                       4.038   Cond. No.                         3.32
==============================================================================
```

Figure 2: Regression output surgeon predicting fastball rpm (left) and hard hit percentage (right)

```
==============================================================================
Dep. Variable:               fastvelo   R-squared:                       0.077
Model:                            OLS   Adj. R-squared:                  0.020
Method:                 Least Squares   F-statistic:                     1.344
Date:                Mon, 15 Apr 2024   Prob (F-statistic):              0.275
Time:                        22:35:13   Log-Likelihood:                -73.038
No. Observations:                  35   AIC:                             152.1
Df Residuals:                      32   BIC:                             156.7
Df Model:                           2
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const          0.2471      0.495      0.499      0.621      -0.760       1.255
andrews        1.0815      0.916      1.181      0.246      -0.784       2.947
neal          -0.5289      0.789     -0.670      0.508      -2.136       1.079
==============================================================================
Omnibus:                       17.133   Durbin-Watson:                   1.344
Prob(Omnibus):                  0.000   Jarque-Bera (JB):               24.807
Skew:                          -1.263   Prob(JB):                     4.10e-06
Kurtosis:                       6.260   Cond. No.                         3.32
==============================================================================
```

```
                            OLS Regression Results
==============================================================================
 Variable:                   exitvelo   R-squared:                       0.110
:l:                               OLS   Adj. R-squared:                  0.055
iod:                    Least Squares   F-statistic:                     1.984
:                    Mon, 15 Apr 2024   Prob (F-statistic):              0.154
:                            22:35:13   Log-Likelihood:                -65.797
 Observations:                     35   AIC:                             137.6
:esiduals:                         32   BIC:                             142.3
lodel:                              2
iriance Type:               nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
:t             0.5471      0.402      1.360      0.183      -0.272       1.366
'ews          -1.1899      0.745     -1.598      0.120      -2.707       0.327
.             -1.0652      0.642     -1.660      0.107      -2.372       0.242
==============================================================================
.bus:                           0.690   Durbin-Watson:                   2.231
ı(Omnibus):                     0.708   Jarque-Bera (JB):                0.377
r:                             -0.254   Prob(JB):                        0.828
:osis:                          2.975   Cond. No.                         3.32
==============================================================================
```

Figure 3: Regression output surgeon predicting fastball velocity (left) and average exit velocity (right)

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                   whip   R-squared:                       0.131
Model:                            OLS   Adj. R-squared:                  0.077
Method:                 Least Squares   F-statistic:                     2.422
Date:                Mon, 15 Apr 2024   Prob (F-statistic):              0.105
Time:                        22:35:13   Log-Likelihood:                -3.0036
No. Observations:                  35   AIC:                             12.01
Df Residuals:                      32   BIC:                             16.67
Df Model:                           2
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const          0.1494      0.067      2.234      0.033       0.013       0.286
andrews       -0.2351      0.124     -1.899      0.067      -0.487       0.017
neal          -0.1803      0.107     -1.690      0.101      -0.398       0.037
==============================================================================
Omnibus:                        4.066   Durbin-Watson:                   2.236
Prob(Omnibus):                  0.131   Jarque-Bera (JB):                2.672
Skew:                           0.595   Prob(JB):                        0.263
Kurtosis:                       3.643   Cond. No.                         3.32
==============================================================================
```

```
[1] Standard Errors assume that the covariance matrix of the errors is correctly
                            OLS Regression Results
==============================================================================
Dep. Variable:                    era   R-squared:                       0.149
Model:                            OLS   Adj. R-squared:                  0.096
Method:                 Least Squares   F-statistic:                     2.795
Date:                Mon, 15 Apr 2024   Prob (F-statistic):             0.0761
Time:                        22:35:13   Log-Likelihood:                -64.252
No. Observations:                  35   AIC:                             134.5
Df Residuals:                      32   BIC:                             139.2
Df Model:                           2
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const          0.7553      0.385      1.963      0.058      -0.029       1.539
andrews       -0.9696      0.713     -1.361      0.183      -2.421       0.482
neal          -1.3998      0.614     -2.280      0.029      -2.650      -0.149
==============================================================================
Omnibus:                        0.243   Durbin-Watson:                   2.203
Prob(Omnibus):                  0.886   Jarque-Bera (JB):                0.435
Skew:                           0.110   Prob(JB):                        0.804
Kurtosis:                       2.500   Cond. No.                         3.32
==============================================================================
```

Figure 4: Regression output surgeon predicting Whip (left) and Era (right)

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                  recov   R-squared:                       0.194
Model:                            OLS   Adj. R-squared:                  0.144
Method:                 Least Squares   F-statistic:                     3.854
Date:                Mon, 15 Apr 2024   Prob (F-statistic):             0.0316
Time:                        22:35:12   Log-Likelihood:                -127.31
No. Observations:                  35   AIC:                             260.6
Df Residuals:                      32   BIC:                             265.3
Df Model:                           2
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const         20.0588      2.332      8.601      0.000      15.308      24.809
andrews        9.3697      4.318      2.170      0.038       0.574      18.166
neal          -3.3316      3.721     -0.895      0.377     -10.911       4.248
==============================================================================
Omnibus:                       14.669   Durbin-Watson:                   1.845
Prob(Omnibus):                  0.001   Jarque-Bera (JB):               18.040
Skew:                           1.174   Prob(JB):                     0.000121
Kurtosis:                       5.619   Cond. No.                         3.32
==============================================================================
```

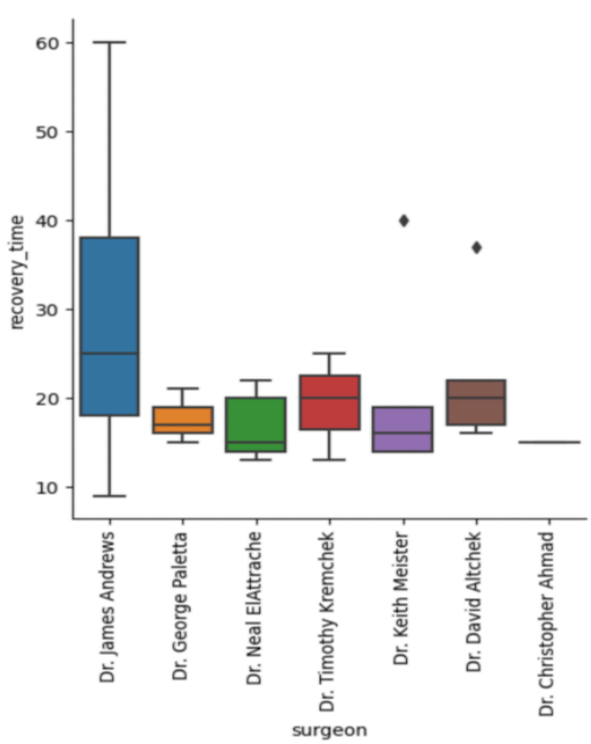
Figure 5: Regression output surgeon predicting recovery time

The results of our initial methodology considering surgeons predicting differences in statistics went how we expected it to for the most part. The surgeons did not seem to have a significant impact on statistics. As seen in figures 1-4 the p values of the first 8 models were all greater than 0.05 showing the surgeon performing the surgery didn't impact the statistics. We did find that the era and whip p values were closer to 0.05 which caused us to explore a second question. Figure 5 shows the only thing statistically significantly impacted by the performing surgeon of the Tommy John surgery. Dr. Neal ElAttrache was shown to reduce recovery time by 3 months when compared to other surgeons. This was validated as being statistically significant as seen by a p-value of 0.03. This conclusion was validated by an output table and a box plot which can be seen in Figures 6 and 7 below.

[12]:

surgeon	count	mean	std	min	25%	50%	75%	max
Dr. Christopher Ahmad	1.0	15.000000	NaN	15.0	15.0	15.0	15.0	15.0
Dr. David Altchek	5.0	22.400000	8.502941	16.0	17.0	20.0	22.0	37.0
Dr. George Paletta	3.0	17.666667	3.055050	15.0	16.0	17.0	19.0	21.0
Dr. James Andrews	7.0	29.428571	17.775049	9.0	18.0	25.0	38.0	60.0
Dr. Keith Meister	5.0	20.600000	11.036304	14.0	14.0	16.0	19.0	40.0
Dr. Neal ElAttrache	11.0	16.727273	3.495452	13.0	14.0	15.0	20.0	22.0
Dr. Timothy Kremchek	3.0	19.333333	6.027714	13.0	16.5	20.0	22.5	25.0

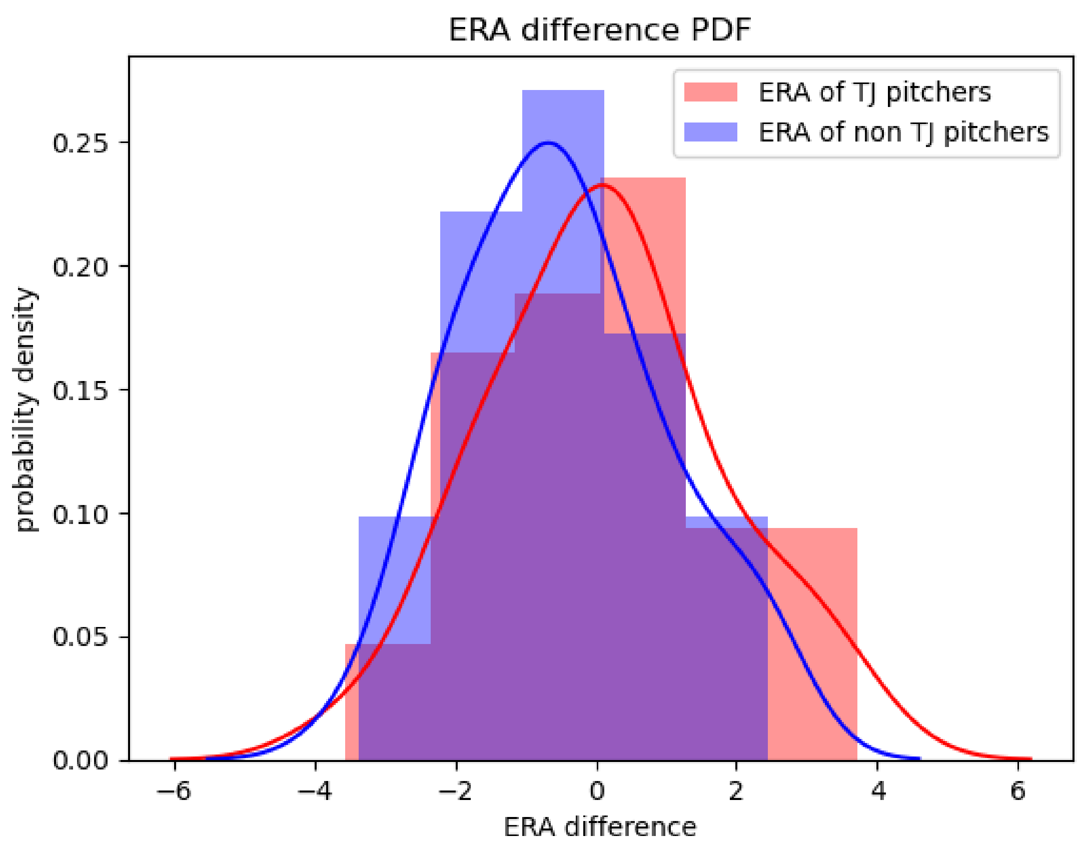
Figure 6: Output table showing surgeons statistics as it pertains to recovery time.



Both figures 6 and 7 confirm that Dr. Neal is the best surgeon to pick if you want to recover the fastest. He has treated the most patients (11) and still has the lowest mean recovery time of almost 17 months. His standard deviation is also one of the smallest which is impressive given the number of patients treated. Interestingly Dr. James Andrews has one of the highest means in part due to his extremely high standard deviation.

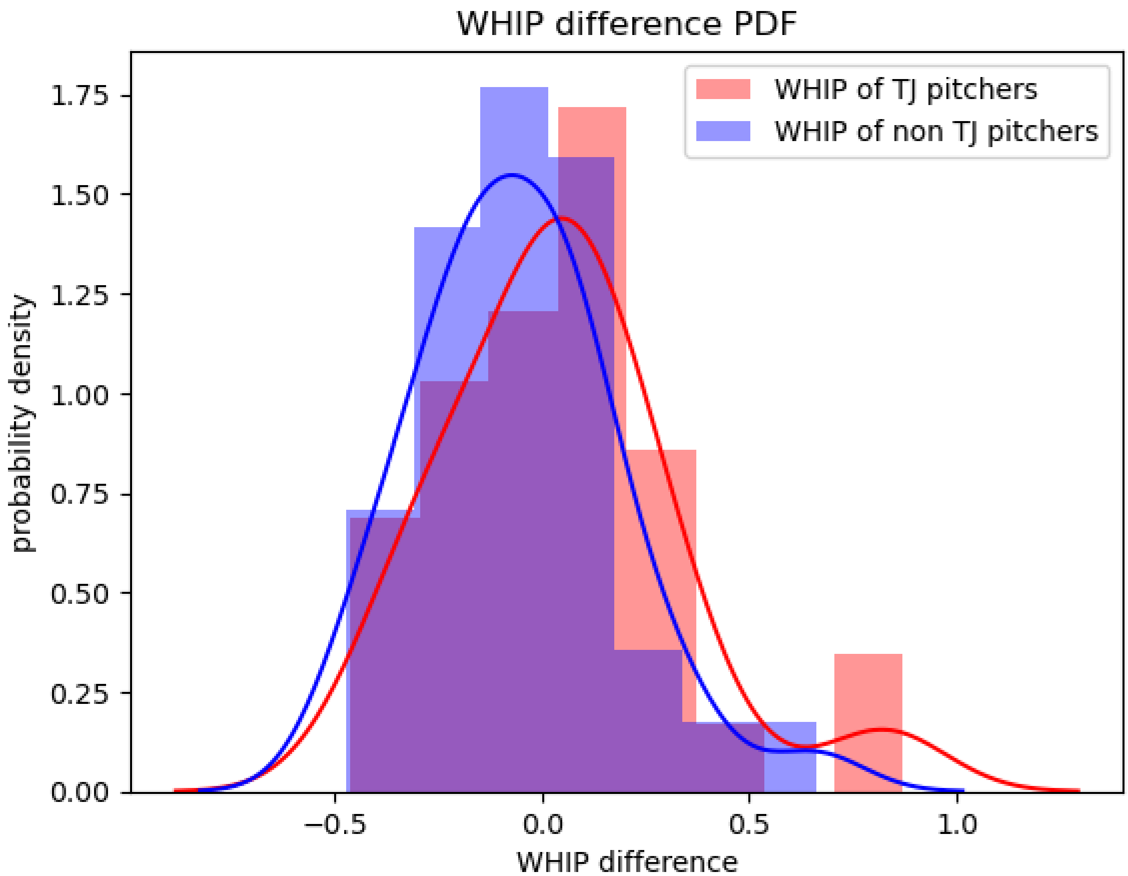
Figure 7: Box plot of output table

The study's second research question aimed to investigate whether there exists a statistically significant difference between the statistics of pitchers who underwent Tommy John surgery and those who did not. Given that the most significant disparity was observed in the difference between a pitcher's ERA and WHIP from the year before surgery to their return year, conducting a test to compare non-Tommy John pitchers with those who underwent the surgery was important for the study's objectives. To achieve this, two independent t-tests were conducted to identify any statistical significance. Test number one involved comparing a list of Tommy John pitchers' ERA differences with a list of non-Tommy John pitchers' ERA differences. The second test is the same but performed on WHIP instead of ERA. The results of these tests are as follows.



T-Test Result: statistic=1.591418198901537, p value=0.11615403130184827, df=68.0

The result of the ERA T-Test shows no statistical significance because of the p-value greater than 0.05. However, by analyzing the graph of the normal distribution of the two data sets it is an interesting observation that pitchers post-Tommy John appear to be improving in performance more than pitchers without the surgery.



T-Test Result: statistic=1.6424610891642755, p value=0.10511137950195451, df=68.0

The result of the WHIP T-Test also indicates no statistical significance, with a p-value greater than 0.05. However, similar to the ERA T-Test, the observation remains consistent: pitchers post-Tommy John surgery seem to exhibit greater performance improvement compared to pitchers without the surgery.