NURSG 741 Project

Lacey Gleason

May 2, 2018

# Milestone 3

## Basic Information

**Project Title:** Major League Baseball Player Injuries from 2015-2017  
**Author Name:** Lacey Gleason  
**Email Address:** [lpgleas@emory.edu](mailto:lpgleas@emory.edu)

## Introduction

### Background

Despite improvements in technology used for athletic conditioning and medical treatment, injuries to Major League Baseball (MLB) players have not declined over time (Conte et al., 2001). In fact, both the frequency and length of disabled list (DL) assignments increased from 1998-2015 (Conte et al., 2016). In addition to the physical and psychological toll, injuries to MLB players have an enormous financial cost. The annual cost of designating players to the disabled list is estimated at $423 million (Conte et al., 2016). One of the contributing factors to the high cost is that baseball contracts are fully guaranteed, and an injury involves the need to pay a player who is not playing in addition to paying for his replacement. Injuries to star players also likely impact viewership and other metrics important to MLB teams, but these indirect consequences have not been the focus of study.

In addition to the cost to MLB clubs and players, baseball injuries are important to understand because there has been an increase in the number of injuries in youth baseball in recent years as well (Erickson et al., 2016). The majority of previous studies on baseball injuries have evaluated risk factors in youth baseball players, and pitchers are the most frequently studied (Erickson et al., 2016). Among pitchers, some injury risk factors that have been identified include pitching all year, pitching more than 100 innings per year, high pitch counts, pitching for multiple teams, pitching on consecutive days, pitching while fatigued, pitching with higher velocity, and a variety of elbow and shoulder kinetics (Erickson et al., 2016; Fleisig, 2017). Additional risk factors include pitching without adequate rest, poor mechanics, and throwing high numbers of breaking or off-speed pitches (Makhni et al., 2017). Due to these findings, MLB and USA Baseball created the Pitch Smart program to provide recommendations on pitch counts and rest days (Fleisig, 2017; Erickson et al., 2016).

Repetitive microtrauma has been proposed as the physiological pathway for many of these injuries (Erickson et al, 2016). Specifically, studies have used a previous season’s work metrics to predict injury based on the cumulative damage injury model (Karakolis et al., 2013). However, little work has been done to evaluate whether the findings from youth baseball hold true for MLB players. It is difficult to extrapolate results to MLB players due to differences in preparation, length of the seasons, and quality of the athletes (Posner et al., 2011). One study which examined whether or not throwing a complete game in the MLB was a risk factor for being placed on the disabled list found that being a high-volume pitcher in general increased the risk of spending time on the DL (Erickson et al., 2018). However, another study focusing on MLB players found that no cumulative work metric was a significant predictor of future injury (Karakolis et al., 2013). Thus, uncertainty remains about the effect of cumulative work on MLB players’ injuries.

It is known that there are differences in injury patterns based on player position. Pitchers have a greater proportion of disability days compared with position players (Posner, 2011). Pitchers are also more likely to suffer from upper extremity injuries than position players, and about a quarter of those injuries are related to the elbow (Sonne & Keir, 2016).

In order to be placed on the disabled list, a player must be certified as unable to play and receive a diagnosis from the team doctor (Posner et al., 2011). Once on the DL, a player cannot return to active play for a specified number of days depending on the type of disabled list used (i.e., 7 days for the 7-day DL, 10 days for the 10-day DL, and 60 days for the 60-day DL). Once a player is placed on the disabled list, the team may replace him on the 25-man roster, which allows the team to play with a full roster of individuals who may be utilized in a game rather than having an injured player on the bench.

It is particularly important to understand MLB injuries better because recent changes adopted by the MLB in 2015 to increase the pace-of-play may have an impact on fatigue for pitchers (Sonne & Keir, 2016). These changes specifically limit the amount of time that a pitcher may take in between pitches, which is a key component of within-game rest time. MLB players and owners negotiated another recent policy change in their collective bargaining agreement in 2016. The new CBA changed the 15-day disabled list to a 10-day disabled list, which may change the frequency and duration of DL assignments (Petriello, 2016). These changes, along with the fact that published studies have focused on time period up to 2015, leave a gap in understanding about the current utilization of the DL by MLB teams and player characteristics related to injury. The aim of this study is to investigate DL utilization between 2015 and 2017 and determine player characteristics that help explain DL placement in 2017.

### Research Question & Hypotheses

The focal question that I am trying to answer is what Major League Baseball player characteristics were associated with placement on the disabled list during the 2017 season.

I would hypothesize that catchers will have disproportionately high rates of collision injuries since they have more exposure to involvement in run scoring plays. Catcher position has not been shown to be an independent risk factor for injury in MLB. Number of runs scored may be related to injury since there is higher collison risk for base runners during scoring plays.

## Methods

Some researchers have distinguished between pitchers who have played at least 1 MLB game and those who appeared in greater than or equal to 10 MLB games in a single season, which they labeled established players (Makhni et al., 2014). I have included all players who had a single appearance in each season. Disabled list information is occasionally available for individuals who do not end up having a single MLB game appearance in a season if they were placed on the DL early in the season or if they were in transition between the majors and minors. The majority of these are likely carry-over placements from injuries in the previous year or offseason.

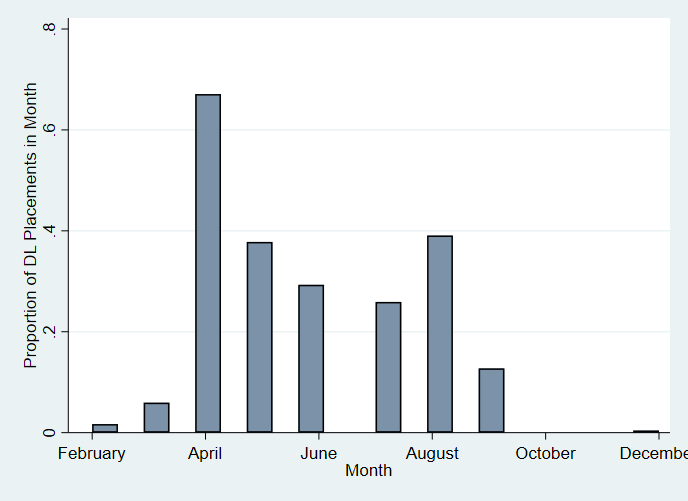
## Results

### Disabled List Transactions in the 2015 MLB Season

#### DL Transactions Overall

In 2015, there were 1,018 MLB Disabled List transactions. Of these, 518 (51%) were placements, 113 (11%) were transfers, and 387 (38%) were activations. Of the DL placements, 286 (55%) were retroactive placements. The median difference between the transaction date and the retroactive date of placement was 3 days. There were 442 unique players involved in these transactions, and they were involved in a median of 2 transactions each (IQR 2-4, range 1-7). When only looking at DL placements, players with any DL placements had a median of 1 placement during the year (IQR 1-2, range 1-3). Placements on the DL were most common in April (31%) and August (18%). Transfers were most common in May (21%) and April (19%). Activations were most common in November (20%) followed by September (18%).

#### Proportion of Total DL Placements Made in Each Month of 2015



#### Position

|  |  |
| --- | --- |
| Position | Frequency and percent of DL placements in 2015 |
| Catcher | 33 (6%) |
| Designated Hitter | 3 (<1%) |
| Infielder | 93 (18.0%) |
| Outfielder | 99 (19%) |
| Pitcher | 290 (56%) |

#### Side of body injured

Only 80% of the DL placements noted the side of injury. Of those that did note a side, 53% involved the right side of the body while 47% involved the left side of the body. It is appropriate that not all DL transactions list side as some injuries, like concussions or groin injuries, do not involve a particular side of the body.

#### Concussions

Fifteen (3%) of the 518 DL placements in 2015 mentioned concussions. No designated hitters were involved in DL placements for concussions in 2015. Catchers and outfielders had highest proportion of concussion-related DL placements with 6% of DL placements for catchers and outfielders involving concussions compared with 5% of DL placements for infielders and <1% of DL placements for pitchers. However, the numbers overall were low. All of the 7-day DL placements involved concussions although it was not noted in one transaction description, and 4 (<1%) of the 15-day DL placements involved concussions.

#### Tommy John Surgery

In 2015, 26 (5%) of the 518 DL placements mentioned Tommy John surgery (TJS) or ulnar collateral ligament (UCL) reconstruction. Twenty-five (96%) of these 26 were DL placements of pitchers. Most of these placements (73%) were made in April in 2015.

#### Overall Surgery

Overall, 50 (10%) of the 518 DL placements mentioned surgery. Thirty-seven (74%) of these 50 were DL placements of pitchers. In addition to Tommy John Surgery mentioned above, other types of surgery included shoulder/labrum/UC joint (7), other elbow (6), knee (4), core muscle (2), ankle (2), hip (2), back (1), and thoracic outlet syndrome surgery (1). Most of these placements happened in April (74%). The Texas Rangers stood out as a team that had 6 surgery-related DL placements during 2015, but all 30 teams had at least one.

#### Team

In 2015, the San Francisco Giants had the most DL placements at 26. The next highest were the Los Angeles Dodgers (25), New York Mets (24), and Texas Rangers (24). The team with the fewest DL placements in 2015 was the Seattle Mariners (7).

#### Word Cloud for DL Transactions in 2015

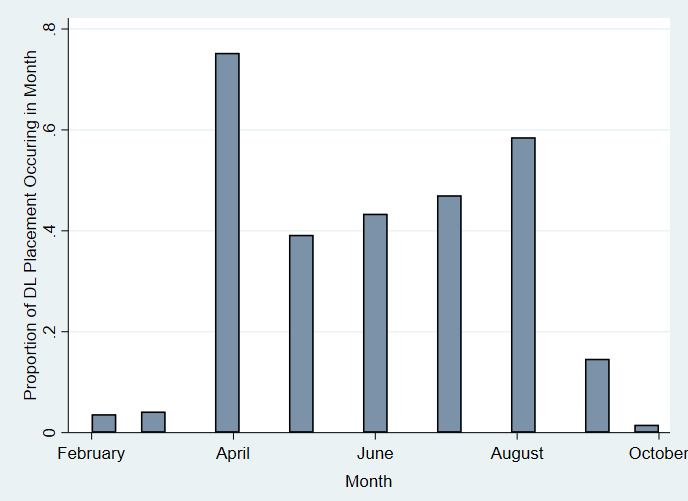


### Disabled List Transactions in the 2016 MLB Season

#### Number and Type of DL Transactions

In 2016, there were 1,139 MLB Disabled List transactions. Of these, 550 (48%) were placements, 135 (12%) were transfers, and 454 (40%) were activations. Of the 550 DL placements, 310 (56%) were retroactive placements. The median difference between the transaction date and the retroactive date of placement was 3 days. There were 483 unique players involved in these transactions, and they were involved in a median of 2 transactions each (IQR 2-4, range 1-8). When only looking at DL placements, players had a median of 1 placement during the year (IQR 1-2, range 1-4). Placements on the DL were most common in April (26%) and August (20%). Transfers were most common in August (22%) and July (19%). Activations were most common in September (18%) followed by August (17%).

#### Proportion of Total DL Placements Made in Each Month of 2016



#### Side of body injured

Only 80% of the DL transactions noted the side of injury. Of those that did note a side, 53% involved the right side of the body while 47% involved the left side of the body. These proportions exactly matched those from 2015.

#### Position

|  |  |
| --- | --- |
| Position | Frequency and percent of DL placements in 2015 |
| Catcher | 44 (8%) |
| Designated Hitter | 4 (<1%) |
| Infielder | 90 (16%) |
| Outfielder | 113 (21%) |
| Pitcher | 299 (54%) |

#### Concussions

Ten (2%) of the 550 DL placements in 2016 mentioned concussions. No pitchers or DHs were involved in DL placements for concussions in 2016. Catchers had the highest proportion of concussion-related DL placements with 7% of DL placements for catchers involving concussions compared with 3% of DL placements for infielders and 4% of DL placements for outfielders. However, the numbers overall were low. As expected, all of the 7-day DL placements involved concussions and 2 (<1%) of the 15-day DL placements involved concussions.

#### Tommy John Surgery

In 2016, 20 (4%) of the 550 DL placements mentioned Tommy John surgery (TJS) or ulnar collateral ligament (UCL) reconstruction. Nineteen (95%) of these 20 were for DL placements of pitchers. Most of these placements actually happened in the offseason, presumably following surgery.

#### Overall Surgery

Overall, 33 (6%) of the 550 DL placements mentioned surgery. Twenty-five (76%) of these 33 were DL placements of pitchers. In addition to Tommy John Surgery mentioned above, other types of surgery included shoulder/labrum (8), other elbow (2), core muscle (1), ankle (1), rib (1), groin (1), and back surgery (1). Most of these placements happened in April (76%) and 15% happened in the offseason (February and March).

#### Team

In 2016, over 6% of DL placements were for Dodgers players. The Dodgers had 8 more DL placements than the next highest team (OAK) and 26 more DL placements than the lowest team (HOU).

#### Word Cloud for DL Transactions in 2016

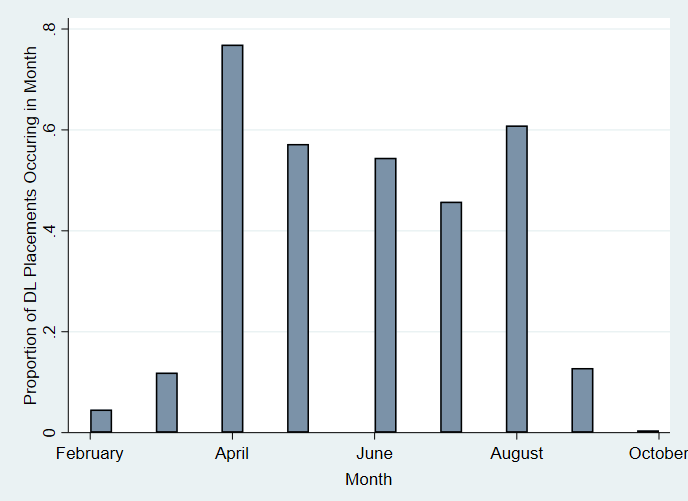


### Disabled List Transactions in the 2017 MLB Season

#### Number and Type of DL Transactions

In 2017, there were 1,386 MLB Disabled List transactions. Of these, 710 (51%) were placements, 133 (10%) were transfers, and 543 (39%) were activations. Of the 710 DL placements, 306 (43%) were retroactive placements. The median difference between the transaction date and the retroactive date of placement was 2 days. There were 537 unique players involved in these transactions and they were involved in a median of 3 transactions each (IQR 2-4, range 1-9). When only looking at DL placements, players had a median of 1 placement during the year (IQR 1-2, range 1-5). Placements on the DL were most common in April (24%) and August (19%). Transfers were most common in May (23%) and June (19%). Activations were most common in June (17%) followed by August (15%) and September (15%). There were no disabled list transactions in January or December of 2017.

#### Proportion of Total DL Placements Made in Each Month of 2017



#### Side of body injured

Only 73% of the DL transactions noted the side of injury. Of those that did note a side, 60% involved the right side of the body while 40% involved the left side of the body.

#### Position

|  |  |
| --- | --- |
| Position | Frequency and percent of DL placements in 2017 |
| Catcher | 51 (7.2%) |
| Designated Hitter | 8 (1.1%) |
| Infielder | 131 (18.5%) |
| Outfielder | 133 (18.7%) |
| Pitcher | 387 (54.5%) |

#### Concussions

Forty (6%) of the 710 DL placements in 2017 mentioned concussions. No designated hitters were involved in DL placements for concussions in 2017. Catchers and outfielders had highest proportion of concussion-related DL placements with 29% of DL placements for catchers and 11% of DL placements for outfielders involving concussions compared with 7% of DL placements for infielders and <1% of DL placements for pitchers. All of the 7-day DL placements involved concussions, and 14 (2%) of the 10-day DL placements involved concussions.

#### Tommy John Surgery

In 2017, 12 (2%) of the 710 DL placements mentioned Tommy John surgery (TJS) or ulnar collateral ligament (UCL) reconstruction. Eleven (92%) of these 12 were for DL placements of pitchers. Forty-two percent of TJS-related DL placements happened in April.

#### Overall Surgery

Overall, 23 (3%) of the 710 DL placements mentioned surgery. Sixteen (70%) of these 23 were DL placements of pitchers. In addition to Tommy John Surgery mentioned above, other types of surgery included shoulder (3), other elbow (3), hip (2), unspecfied arm (2), foot (1), thoracic outlet syndrome (1), appendix (1), achilles (1), colon (1), knee (1), and biceps (1) surgeries. One of the surgeries was listed as both shoulder and biceps. Most of these placements happened in April. Five of the Tommy John surgeries did not mention surgery in the DL transaction. Thus, the number of surgeries should be at least 28 (4% of DL placements).

#### Team

In 2017, 6% of DL placements (41 total) were for Dodgers players. The Dodgers had 8 more DL placements than the next highest team (TOR) and 29 more DL placements than the lowest team (CHC).

#### Word Cloud for DL Transactions in 2017



### Logistic regression for DL placement in 2017

library(readxl)  
  
combined <- read\_excel("2016\_2017\_combined\_all.xls",   
 sheet = "Sheet1")  
  
# logistic regression - placed on the DL in 2017 or not  
DL17 <- glm(place\_17 ~ place\_16 + surg\_16 + INN\_16 + pitch\_16 + concordance, family=binomial(link="logit"), data = combined)  
summary(DL17)

##   
## Call:  
## glm(formula = place\_17 ~ place\_16 + surg\_16 + INN\_16 + pitch\_16 +   
## concordance, family = binomial(link = "logit"), data = combined)  
##   
## Deviance Residuals:   
## Min 1Q Median 3Q Max   
## -1.7468 -1.0453 -0.8234 1.2311 1.7369   
##   
## Coefficients:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) -1.2614977 0.1978517 -6.376 1.82e-10 \*\*\*  
## place\_16 0.7922599 0.1356082 5.842 5.15e-09 \*\*\*  
## surg\_16 0.8319822 0.6802128 1.223 0.221284   
## INN\_16 0.0006326 0.0002088 3.030 0.002448 \*\*   
## pitch\_16 0.5915692 0.1764680 3.352 0.000802 \*\*\*  
## concordance 0.3184757 0.1393398 2.286 0.022277 \*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## (Dispersion parameter for binomial family taken to be 1)  
##   
## Null deviance: 1457.7 on 1061 degrees of freedom  
## Residual deviance: 1402.2 on 1056 degrees of freedom  
## (662 observations deleted due to missingness)  
## AIC: 1414.2  
##   
## Number of Fisher Scoring iterations: 4

exp(DL17$coefficients)

## (Intercept) place\_16 surg\_16 INN\_16 pitch\_16 concordance   
## 0.2832295 2.2083814 2.2978690 1.0006328 1.8068215 1.3750302

exp(confint(DL17))

## Waiting for profiling to be done...

## 2.5 % 97.5 %  
## (Intercept) 0.1910723 0.4152875  
## place\_16 1.6946726 2.8844539  
## surg\_16 0.6655085 10.5642661  
## INN\_16 1.0002250 1.0010449  
## pitch\_16 1.2820964 2.5621637  
## concordance 1.0475672 1.8094290

## Discussion

These results are in line with previous literature that suggests that April is the most common month for MLB injuries and that there is a steep drop off in DL placements in September (Erickson et al., 2016). The greater number of injuries in April is thought to be due to deconditioning and overload (Posner, 2011). Also, fewer players may be placed on the DL in September because teams are allowed to expand their rosters to 40 men (Posner, 2011).

The number of DL placements per year seems higher than reported for previous years, which was 438 players per year (Posner, 2011). Mean of 464 designations annually between 1998-2015 (Conte et al., 2016).

Position can certainly influence type of injury. For example, pitchers are more likely to have upper extremity injuries, but catchers and base runners are more likely to have acute collison injuries, including concussions (Erickson et al, 2016).

### Changes in DL trends

One season is related to the next due to the fact that transactions cross over between the years. For example, if there is a very high proportion of placements in one year, there may be a higher than average number of activations in the following year depending on the type and duration of injury and when the injuries occurred.

### Data Checking

There were several challenges that made it difficult to use the transaction data from MLB.com. One challenge is that retroactive transactions seem to have occasionally produced duplicate transactions both on the date of the actual transaction and on the date to which the DL was retroactive. I am not sure if this is a bug in the system for generating the transaction table, but it did result in duplicates. Any time that there were consecutive placements or activations, the record was manually checked the player’s MLB.com page to verify the information. Depending on the year, 3%-10% of transactions had some sort of inconsistency that required verification.

Names Names with a “de” like “Rubby De La Rosa” are denoted differently. De showed up as the full last name on one data source and as only part of the last name (merged with larosa) in another data source. A number of players shared the same first initial, last name, and position. A similar issue occured with Scott Van Slyke, who showed up as two different players as did Tommy La Stella.

Name of team - CWS vs. CHW in different datasets.

Another limitation of the transaction data is that the injury information is not structured consistently. For example, one of the 7-day disabled list transactions in 2015 did not mention concussion even though the 7-day disabled list is exclusively for concussions. This transaction did not contain any information on the type of injury. When I looked up the player to check the type of injury, the player did in fact have a concussion after being struck in the neck with a fly ball. Thus, while the data from MLB.com transactions appear to be very accurate, there are instances in which information is missing. As a result, the proportions of different types of injury mentioned in this report should be interpreted with some skepticism.

Another challenge related to the injury information being unstructured is that there are many different ways of denoting each injury. For example, many transactions simply said shoulder injury, while others said labrum surgery or AC joint surgery. The labrum is a ring of cartiledge that supports the shoulder, and the acromioclavicular (AC) joint is at the top of the shoulder. Thus, it would require substantial knowledge of these injuries to allow for automatic classification of the injuries that captures all the possible iterations. It is also possible that some of these surgeries are described incorrectly. For example, some of those listed as elbow surgeries may in fact have been Tommy John Surgeries.

## Limitations

A significant portion of DL transactions were not complete within a single calendar year. For example, 29% of the DL transactions in 2015, 28% of the DL transactions in 2016, and 27% of the DL transactions in 2017 were incomplete. This means that the last DL transaction for the player in the year was a placement or that the first DL transaction for the player in the year was an activation. I addressed this limitation by focusing on placement on the DL as the outcome of interest. In the future, if I want to include the number of days spent on the DL, I will need to complete the transactions that cross years or consider defining the year in a way that better suits baseball transaction activity rather than the calendar year.

Missing data on injury type - diagnosis not always listed or specfic. Lack of data on work metrics and player characteritics from minor leagues or from winter leagues.

Some of the changes in type of injury might have to do with changes in the ability to detect certain injuries. For example, the spike in concussions between 2016 and 2017 may be due to a change in the concussion protocol. In a more thorough study, policy changes would be incorporated into the analysis to better understand difference between seasons. There are also other changes between years that may be of significant interest, such as a change in the playing style from small-ball eras to home-run eras or an increase in the prevalence of “flamethrower” pitchers or changes in bullpen utilization patterns. Changes in medical technology could affect how many players with the same injury end up getting surgery.

While this is not thought to be due to improvements in ability to detect injuries, it may be partially due to the success of reutrning players to competition after procedures like Tommy John surgery and then subsequent reinjury (Conte et al., 2001; Makhni et al., 2014).

MIght want to exclude injuries during the offseason?

## Future Research

Include information on pitch veloity and pitch type from Statcast. Also, it would be helpful to be able to parse out the pitcher information in more detail with regard to rest between appearances, especially for relief pitchers.

Look at days on the disabled list as an outcome in a two-part model.

Availability of new statistics The Pitch f/x video track system now analyzes over 60 variables on each of the approximatley 660,000 pitches thrown in the MLB each year (Conte et al., 2016).

Wearable devices & pitchers The mTHROW sleeve is an elastic sleeve worn by pitchers that measures torque, velocity, and workload on their dominant arm through use of an accelerometer (Erickson et al., 2016). This sleeve sends the data to a smart device (e.g., tablet, cell phone). New developments allow for widespread use of these devices in research because they have reduced the need for expansive equipment associatd with high-speed video analysis using markers (Makhni et al., 2017).

inertial measurement unit with triaxial accelerometer and triaxial gyroscope - questions about precision (Fleisig, 2017). The accuracy was validated through comparison with optical motion capture, which is the current gold standard. (Fleisig, 2017).

Makerless optical motion tracking - emerging for use during competitions. One elbow sensor was recently approved for use during MLB games. (Fleisig, 2017).

Investigating Twitter APIs and tweets surrounding time of injury. Twitter handle is available in MLB.com player profile.

## Link

The link for this file is located at <https://github.com/lpgleason/Project.git>.

## References