Evaluating the Impact of IL Stint Duration on MLB Pitcher Performance

Angelina Allen

May 14, 2025

1 Introduction

Baseball pitchers are especially prone to arm injury due to the demands of their position. Pitching is very taxing on the body, often leading to surgeries or extended stints on the injured list (IL). While some injuries require months to recover, others seem less severe, with players returning after 10–30 days. Is that short recovery time truly sufficient, or are teams putting their players at a higher risk? This study investigates whether the duration of an IL stint impacts the pitcher's performance. Specifically, it will analyze metrics before and after IL duration to determine if a statistically significant relationship exists between time spent on the IL and key pitching performance metrics.

2 Data

The dataset combines information from two primary sources: FanGraphs' IL roster data and Baseball Savant's pitching metrics for the 2023 and 2024 MLB seasons. From FanGraphs, all pitchers in 2024 who were placed on the IL due to arm-related injuries, including issues with the shoulder, elbow, wrist, and other upper-limb conditions, were used. The Baseball Savant datasets consist of the top-ranked pitchers by xwOBA (expected weighted on-base average) in 2023 and 2024. The three datasets were merged by pitcher's name to create the final dataset. It produced eight different pitchers with a short IL stint and rankable metrics in 2023 and 2024. Specifically, our data includes 16 observations with paired data from 2023 and 2024. By structuring the data across two seasons, we treat 2023 as the pre-injury baseline and 2024 as the post-IL period. The data includes 36 variables covering various information and metrics about the pitcher. Specifically, the metrics selected for this analysis, ERA (Earned Run Average), xwOBA (Expected Weighted On-Base Average), fastball average speed, and walk percentage, are key indicators when evaluating pitcher performance. With a baseball knowledge background, these are the most logical variables to analysis that have a high value in pitcher performance.

- **p__ERA** is the number of earned runs a pitcher allows per nine innings
- **xwOBA** is a more advanced metric involving quality of contact, strikeouts, and walks, helping remove external factors like defense or luck.
- **Fastball average speed** is the velocity the ball is thrown.
- **BB Percentage (walk%)** shows the pitcher's control and command.

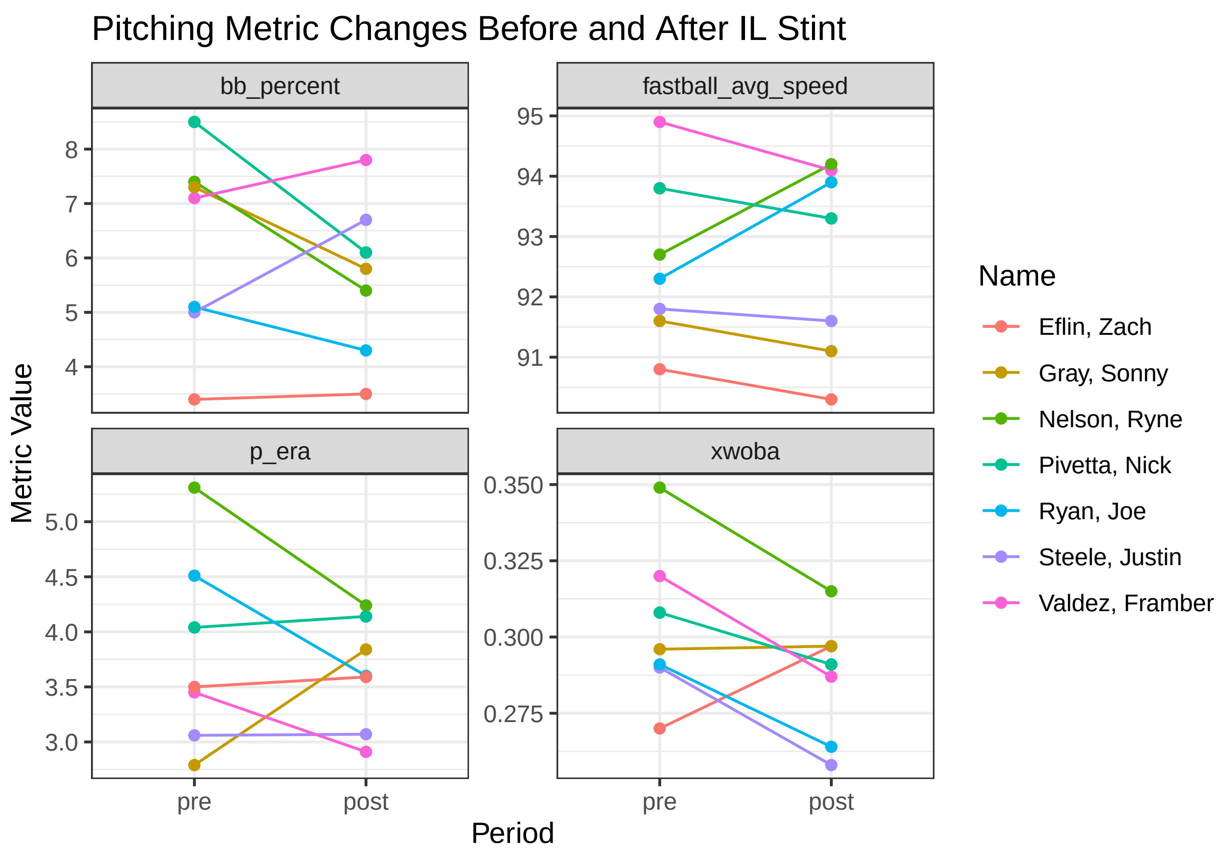

Figure 1. *Pre- and post-IL values for four pitching metrics across seven MLB pitchers. Metrics include ERA, and xwOBA, Fastball Velocity, Walk%,. Lines connect pre-and post-IL status to visualize changes.*

3 Methods

All statistical analyses were done in R (version 4.4.2). The tidyverse package was used for data wrangling and visualization, including dplyr and ggplot2. Hypothesis testing was performed using t.test(), and linear modeling was conducted using the lm() function.

3.1 Paired T-Test

First, this study will conduct a paired t-test to help determine if there is a significant difference in pitchers' performance before and after an injury stint. The paired t-test is appropriate because it compares each pitcher's performance before and after the injury, treating them as their own control to account for individual variability. For the paired t-test to be valid, it is assumed that the differences between pre- and post-measurements are normally distributed, a condition supported by the distribution of our variables. The data is standardized before analyzing to ensure comparability across metrics with different scales.

Let $X_i(pre)$ and $X_i(post)$ represent the metrics of interest for player i. The paired t-test evaluates whether the mean difference $D_i = X_i(post) - X_i(pre)$ significantly differs from zero.

The null and alternative hypotheses for each metric took the following form:

ERA:

- $H_0 : \mu_D(ERA) = 0$

- $H_0 : \mu_D(ERA) \neq 0$

XWOBA:

- $H_0 : \mu_D(xwOBA) = 0$
- $H_0 : \mu_D(xwOBA) \neq 0$

Fastball average speed:

- $H_0 : \mu_D(fastballspeed) = 0$
- $H_0 : \mu_D(fastballspeed) \neq 0$

Walk:

- $H_0 : \mu_D(walkpercent) = 0$
- $H_0 : \mu_D(walkpercent) \neq 0$

3.2 Linear Regression

To further examine the research question: Does the duration of a stint impact performance metrics, a series of simple linear regression models were fit. For each performance metric, the difference between post- and pre-IL values was modeled as the response variable and the length of IL stint as the explanatory variable.

The model took the following form: $Difference_i = \beta_0 + \beta_1 * Stint_i + \epsilon_i$

- $Difference_i$ is the response variable; the change in a performance metric for player i (post - pre IL stint)
- β_0 expected difference when Stint = 0
- β_1 how much the metric difference changes per unit increase in IL.
- $Stint_i$ is the length of the IL stint for player i
- ϵ_i error term

4 Results

4.1 Paired T-Test

A paired t-test was conducted to compare the four performance metrics before and after short IL stints: ERA, xwOBA, fastball average speed, and walk%. The results showed that the 95% confidence interval contained zero across all metrics, indicating no statistically significant change associated with the performance metrics from pre- and post-short IL stint.

ERA:

- 95% Confidence Interval = $(-0.86, 0.86)$
- No statistically significant difference in ERA pre and post short IL stint.

xwOBA:

- 95% Confidence Interval = $(-0.91, 0.91)$
- No statistically significant difference in xwOBA pre and post short IL stint.

Fastball average speed:

- 95% Confidence Interval = $(-0.61, 0.61)$
- No statistically significant difference in fastball average speed pre- and post-short IL stint.

Walk%:

- 95% Confidence Interval = $(-0.84, 0.84)$
- No statistically significant difference in Walk% pre and post short IL stint.

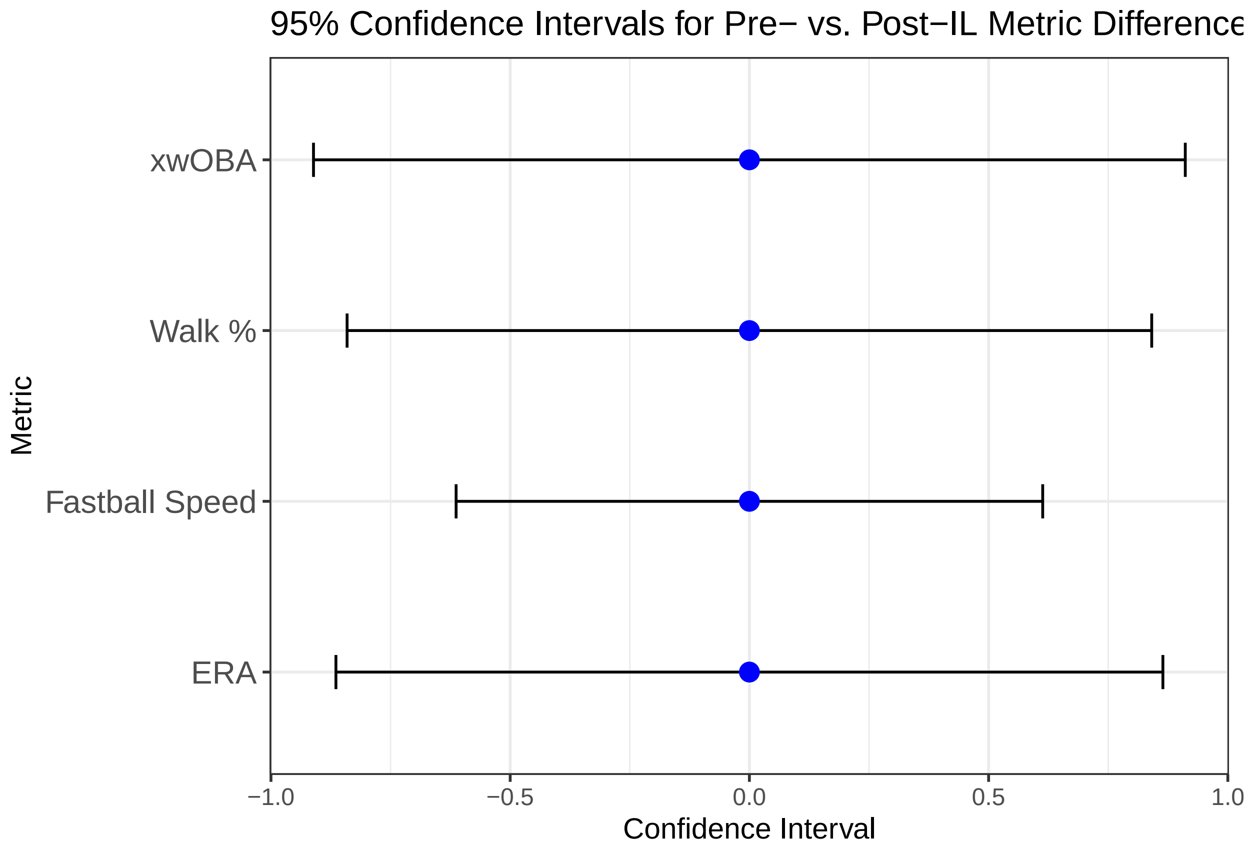

Figure 2. *Visualization of 95% confidence intervals for the mean change in each pitching metric from pre- to post-IL. Blue point indicating interval containing zero.*

4.2 Linear Regression

Multiple simple linear regression models were fitted using IL stint duration to predict changes in the key metrics. The results showed no statistically significant evidence that the performance metric is affected by IL stint.

- **ERA**: $\beta_1 = 0.029$, p = 0.277
- **xwOBA**: $\beta_1 = 0.017$, p = 0.435
- **Fastball Average Speed**: $\beta_1 = -0.019$, p = 0.304
- **Walk%**: $\beta_1 = 0.006$, p = 0.825

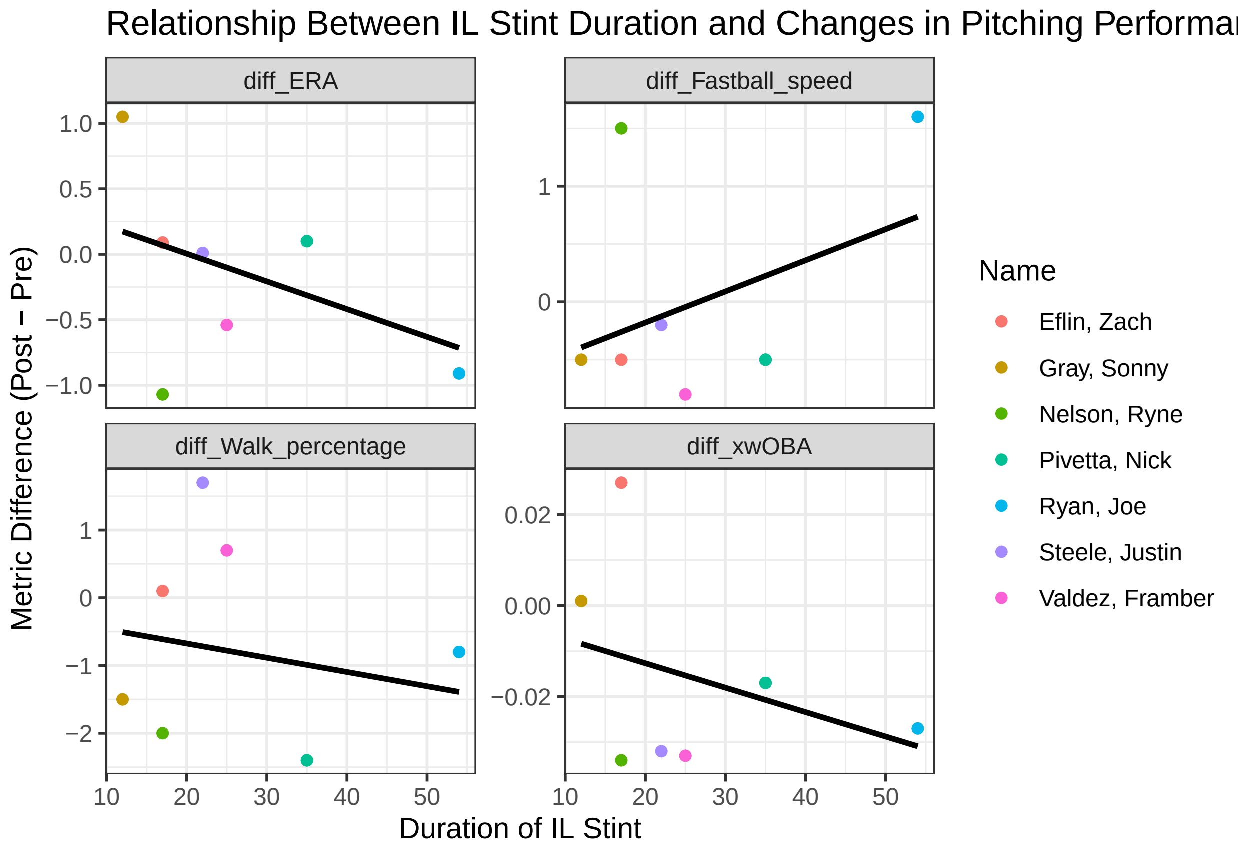

Figure 3. *Relationship between IL stint duration and changes in pitching performance metrics. Each point represents a pitcher, and trend lines show the linear regression fit for each metric.*

Looking at the visualization in **Figure 3**, we see patterns of a slight decrease in ERA and walk% post-IL stint. However, the high p-values and near-zero slope estimates suggest IL stint duration does not reliably predict changes in any of the four performance metrics.

5 Discussion

This study investigates whether a short IL stint's duration impacts the pitcher's performance metrics using paired t-tests and simple linear regression models applied to key performance metrics: ERA, xwOBA, fastball velocity, and walk percentage. Although no statistical evidence of an impact was found in these analyses, it highlighted potential trends in our metrics. For example, ERA and Walk percentage showed a slight decrease post-IL stint, which sparks interest in further investigation.

Due to the limited availability of data and small sample size, this study is limited in the strength and generalization of any conclusions. Additionally, multiple other factors, such as injury type, past injuries, and age, can significantly impact a pitcher's performance. So, it is recognized that other factors were overlooked.

These findings justify further exploration in regression modeling with a larger dataset to detect meaningful effects in future work. Including more pitchers, specific injuries, starter versus relievers, and possibly teams' recovery work can enhance this study and show more statistical significance. With this, a study can produce more complex models, like mixed-effect models or time series methods, to analyze.

Overall, this analysis is not meant to provide definitive answers but to highlight potential trends in the impact of short IL stints on pitcher performance. Then, when offered sufficient data, the exploration can continue to provide more generalizations for pitchers.