Article

Injury Patterns and Impact on Performance in the NBA League Using Sports Analytics

Vangelis Sarlis [ID], George Papageorgiou [ID] and Christos Tjortjis *[ID]

School of Science and Technology, International Hellenic University, 57001 Thessaloniki, Greece; e.sarlis@ihu.edu.gr (V.S.); gpapageorgiou2@ihu.edu.gr (G.P.)
* Correspondence: c.tjortjis@ihu.edu.gr

Abstract: This research paper examines Sports Analytics, focusing on injury patterns in the National Basketball Association (NBA) and their impact on players' performance. It employs a unique dataset to identify common NBA injuries, determine the most affected anatomical areas, and analyze how these injuries influence players' post-recovery performance. This study's novelty lies in its integrative approach that combines injury data with performance metrics and salary data, providing new insights into the relationship between injuries and economic and on-court performance. It investigates the periodicity and seasonality of injuries, seeking patterns related to time and external factors. Additionally, it examines the effect of specific injuries on players' per-match analytics and performance, offering perspectives on the implications of injury rehabilitation for player performance. This paper contributes significantly to sports analytics, assisting coaches, sports medicine professionals, and team management in developing injury prevention strategies, optimizing player rotations, and creating targeted rehabilitation plans. Its findings illuminate the interplay between injuries, salaries, and performance in the NBA, aiming to enhance player welfare and the league's overall competitiveness. With a comprehensive and sophisticated analysis, this research offers unprecedented insights into the dynamics of injuries and their long-term effects on athletes.

Keywords: basketball analytics; data analysis; data mining; data science; injury analytics; musculoskeletal injuries; sports analytics (SA)

1. Introduction

In the domain of professional sports, particularly in the National Basketball Association (NBA), the advent of Sports Analytics (SA), Data Science, Machine Learning (ML), and Data Mining has revolutionized our understanding and prevention of musculoskeletal injuries. These innovative methodologies, which involve the comprehensive collection and analysis of data pertaining to athletes' movements, training patterns, and injury histories, are instrumental in identifying risk factors for knee injuries. This, in turn, facilitates the development of targeted interventions to mitigate such occurrences, a critical advancement in athlete health and performance [1–6].

The musculoskeletal system is a complex structure that plays a vital role in human movement by carrying a significant amount of body weight, allowing for a wide range of motion in six degrees of freedom [7]. Recent advancements in AI and ML algorithms have made significant strides in predicting and diagnosing injuries, marking a paradigm shift in how sports science approaches injury prevention and management. These advancements not only aid in better understanding injury mechanisms, but also play a pivotal role in enhancing athlete outcomes [8–12].

Overall, the use of SA can lead to a better understanding of the causes of musculoskeletal injuries and help to develop more effective strategies for preventing these injuries. By identifying risk factors and protective strategies, it is possible to reduce the incidence of musculoskeletal complications and ultimately improve athletes' health and performance [13–15].

The NBA, a league where physical prowess and high-intensity play are paramount, presents unique challenges in terms of injury prevention and management. Injury risks in the NBA are multifaceted, stemming from factors such as improper technique during physical activities, muscle imbalances, poor flexibility, and weak knee stability. Addressing these risks requires a comprehensive approach combining physical evaluations, strength and flexibility assessments, and a review of injury history. Such detailed assessments pave the way for specific interventions that reduce injury risk and enhance overall performance [16].

In particular, targeted strength training and technique refinement are crucial in improving knee stability and reducing injury risks. Incorporating stretching and flexibility exercises into athletes' routines can further bolster this preventive strategy. Proactively addressing these risk factors is not merely about mitigating injury risks; it is about enhancing overall performance and career longevity in the demanding world of professional basketball [2,8,17].

Athlete injuries are a major concern in the world of sports, as they can significantly impact on an athlete's performance and career. One of the most common types of sports-related injuries is knee injury, which is a tear in the ligament that connects the thigh bone to the shin bone and provides stability to the knee joint. The risk of an injury can be caused by several factors, including improper technique during physical activity, muscle imbalances, poor flexibility, and weak knee stability. These factors can lead to increased stress on the body and increase the likelihood of an injury occurring. To minimize the risk of injuries, it is essential to identify and assess the risk factors associated with these injuries. This can be achieved through a combination of physical evaluations, strength and flexibility assessments, and injury history reviews. In an epidemiological study of basketball injuries during one competitive season in professional and amateur Spanish basketball, the authors aimed to evaluate the incidence and types of injuries in Spanish basketball players and to compare injury rates between professional and amateur athletes. The study found that ankle sprains were the most common injury, followed by knee injuries. Professional players had a higher overall injury rate than amateur players, but amateur players had a higher rate of knee injuries [13].

One study presents a deep learning approach to forecasting injuries in the NBA, highlighting the challenges and nuances of dealing with imbalanced injury datasets. This study emphasizes the importance of correctly splitting data to avoid overfitting and proposes a novel model, METIC (Multiple bidirectional Encoder Transformers for Injury Classification), for assessing injury classification. The model is designed to process sequences of data related to past injuries and games, offering insights into risk factors and the potential for predicting future injuries based on player history and game loads [18].

The knee joint, with its complex anatomy and biomechanical functions, epitomizes the challenges faced in sports medicine. Supporting a significant portion of body weight, it enables a wide range of motion and is crucial for the high-level performance expected of NBA athletes. Understanding the knee's intricate structure—including bones, joint capsules, muscles, tendons, and ligaments—is vital in comprehending injury risks and devising effective prevention strategies [7].

Pivoting to epidemiological insights, studies have revealed intriguing patterns in NBA injuries. Players who sustain injuries tend to be taller and heavier, often playing in forward or center positions. This correlation between physical attributes and injury risks underscores the complexity of sports injuries, necessitating ongoing research for a comprehensive understanding [19–22]. There is some research that suggests there may be a connection between musculoskeletal injury and certain characteristics of NBA players. One study found that NBA players who sustained an injury were more likely to be taller and heavier than players who did not sustain an injury. Additionally, the study found that players who sustained an injury were more likely to play the forward or center positions, which are positions that involve more jumping and pivoting movements [23]. Further studies have highlighted the multifactorial nature of injuries in NBA players, emphasizing the need for continued research to unravel the intricate web of contributing factors [24–26].

One study conducted an epidemiological retrospective analysis of the NBA seasons from 2017–2018 to 2020–2021, detailing the frequency, characteristics, and impact of injuries on performance, including the effects of the COVID-19 pandemic on the league's schedule and injury rates. The study utilized publicly available data from the NBA's official website, focusing on the official game box scores to track inactive players and injuries. The research aimed to provide new insights into injury trends and their implications for player performance and team strategies [27].

Another study identified the incidence of musculoskeletal injuries in the NBA from the 2006–2007 to 2011–2012 seasons. The study found that there were a total of 24 knee injuries during this time period, with most injuries occurring during games. The study also found that players who sustained an injury were more likely to be taller and heavier than players who did not sustain an injury. In conclusion, while I could not find a specific study that matches the exact criteria you have asked, these studies and others have shown that injuries tend to occur more frequently during games rather than during practices or other team activities; however, it is hard to conclude if the majority of injuries occur during the fourth quarter specifically as the sample size of the studies is relatively small and it is not clear if these studies have that information. Overall, knee injuries in NBA basketball players tend to occur most commonly during games, with a high incidence in the fourth quarter and during the competitive season. While the exact cause of injuries is still not fully understood, it is believed that factors such as fatigue, overuse, and the intense physical demands of the sport may all contribute to the high incidence of these injuries [28].

A pivotal study examining NBA player performance markers before and after severe lower extremity injuries—including ankle, knee, and hip injuries—between 2008 and 2019 provides critical insights. The study found a notable decline in performance levels in less than half of the players within two years post injury, signifying the profound impact such injuries can have on a player's career trajectory [29].

Other experimental designs have included the use of 3D motion analysis and physical data such as sex, body mass index, hamstring flexibility, knee joint laxity, medial knee displacement, height, ankle plantar flexion at initial contact, leg press one-repetition max, and knee valgus at initial contact [8]. Innovative experimental approaches, including 3D motion analysis and the examination of various physical parameters, further contribute to our understanding of injury mechanisms. Data from inertial sensors worn by rugby players, for instance, have been instrumental in differentiating between healthy and post-ACL injury states [9]. Additionally, the application of ML in diagnosing anterior cruciate ligament (ACL) tears through magnetic resonance imaging (MRI) represents a significant leap in injury diagnostics [10,11,30].

A systematic review focusing on the epidemiology of sports injuries in basketball has shed light on injury prevalence and characteristics. This comprehensive analysis revealed that ankle and knee injuries are the most common, with a higher incidence observed in male players. The predominance of injuries during games as opposed to practice sessions highlights the urgent need for effective injury prevention strategies tailored to the high-stakes environment of competitive play [31]. Interestingly, post-ACL surgery performance analysis revealed that while there was an initial decline in performance, many players were able to regain their pre-injury levels, suggesting that ACL reconstruction does not significantly impede NBA players' career longevity or performance [32].

Further emphasizing the importance of epidemiological studies, research tracking basketball players over a season has offered valuable insights into the frequency and types of injuries, with a focus on ankle sprains and overuse knee injuries. Such findings are instrumental in informing targeted prevention efforts [33,34].

In summary, studies on the return to performance after severe ankle, knee, and hip injuries in NBA players, utilizing NBA data to analyze player performance pre and post injury, underscore the significant impact of severe injuries on players' performance and career paths. This research, along with ongoing studies, continues to shape our

understanding of injury patterns and prevention strategies, ultimately aiming to safeguard the health and longevity of athletes in the high-stakes world of professional basketball.

This study examines injury patterns in the NBA and their effects on player performance. Utilizing a unique dataset, it identifies common injuries and their impact on players post recovery. This study's innovation lies in its integration of injury data with performance metrics and salary information, offering new insights into how injuries influence both economic and on-court performance. This approach not only uncovers patterns in injury periodicity and seasonality but also examines the specific impacts of injuries on players' per-game performance. The findings provide valuable contributions to sports analytics, guiding injury prevention strategies and enhancing player welfare in the NBA.

2. Materials and Methods

Addressing three research questions, listed below, using SA has the potential to assist the sports industry in shaping its strategy and making informed investment decisions to minimize costs and achieve championship victories, ultimately enhancing its reputation.

2.1. Research Questions/Hypothesis

1. What are the most common injuries in the NBA, and which anatomical regions are most frequently affected?
2. Is there a tendency for injuries to occur before or after another injury?
3. Do players' per-match statistics and performance change significantly after returning from specific injuries?

This study is centered around three key questions to understand injury dynamics in the NBA. Firstly, it aims to identify the most common injuries and the anatomical regions they predominantly affect. Secondly, this study explores whether there is a pattern in the occurrence of injuries, particularly in relation to the timing of one injury following another. Lastly, it investigates the extent to which specific injuries impact players' per-match statistics and overall performance post recovery. These questions are essential in hypothesizing the relationship between injury patterns and their consequences on playing and economic performance in basketball.

2.2. Aim and Objectives

Basketball is a high-intensity, physically demanding sport. This can lead to injuries, which can have a significant impact on a player's performance and career. To mitigate the risk of musculoskeletal injuries and improve overall performance, it is important to understand the relationship between basketball performance and advanced analytics [35]. Advanced analytics have become an increasingly important tool in the field of sports performance and injury prevention [36].

The purpose of this research, therefore, is to establish a relationship between basketball performance and advanced analytics in musculoskeletal injuries. The aim is to use data-driven methods to analyze performance and injury data and to uncover valuable insights into the factors that contribute to injuries in basketball players. Hence, this research paper can provide recommendations to decision makers on how to reduce the risk of injuries and improve performance.

The goal is to also examine advanced basketball performance (Advanced, 4Factors, Miscellaneous, Scoring, Traditional, Usage, and Player Track), injury, and salary analytics. This requires the use of a variety of advanced analytical tools, including text mining and statistical modelling, to extract meaningful insights from the data.

In conclusion, the purpose of this research is to establish a relationship between basketball performance and advanced analytics in the anatomical subareas of musculoskeletal injuries, with the aim of uncovering valuable insights for players and teams and furnishing recommendations to decision makers. By using data-driven methods to analyze performance and injury data, this research hopes to contribute to the broader effort to reduce the risk of injuries in basketball and improve overall performance.

2.3. Methodology

The data for this study were primarily sourced from various available sources [37–39], with the goal of acquiring the most comprehensive information possible to facilitate a robust analysis. The most challenging aspects of the process involved not only data retrieval but also the pre-processing stage, which entailed consolidating the information into a supervised data model and prioritizing data quality. Our methodology comprised analyses of NBA player performance, injuries, and salaries. It encompasses three key areas: data collection, where extensive player data from 2000–2001 to 2022–2023 seasons were gathered; data engineering, which refined and integrated these data for clarity and consistency; and data analysis and statistical methodology, applying both descriptive analytics to assess the impact of injuries on player performance and also statistical tests based on the impact of the findings and players' salary changes, as outlined in the following sub-sections, and involved the stages of data collection, pre-processing, analysis, and result evaluation [40].

2.3.1. Data Collection

The data collection stage had an unorganized and heterogeneous structure. To comprehend the information and derive valuable insights, the methodology involved collecting, text mining, pre-processing, analyzing, and evaluating the results. The data were acquired through a combination of Python scripts and the KNIME Analytics Platform flows. It underwent a pre-processing stage that included identifying and removing missing values and irrelevant data. Subsequently, the data underwent an Extract, Transform, and Load (ETL) process for standardization and homogenization, using NBA API, PostgreSQL, and Python scripts for data analysis.

In this study, a comprehensive dataset was acquired and analyzed. It encompassed player performance, injuries, and salary data for all NBA players from the 2000–2001 season to the 2022–2023 season. This subsection outlines the sources from which the data were acquired, the types of data, and the different shapes of each corresponding dataset that was scraped.

The player performance and injury data for this study were obtained from nba_api, a robust API that directly retrieves data from the NBA's official website and database. Data collection was conducted over a specific timespan from 2000–2001 to 2022–2023 through web scraping [39]. In total, 2296 players were included in this comprehensive dataset, representing the entire pool of players who participated in the NBA across the selected seasons.

Two separate scrape runs were conducted using the nba_api. One targeted the regular season per game_date player performance results, and the other focused on playoffs. Nine distinctive datasets, each highlighting different player performance aspects and rich in both demographic and performance-related metrics, were assembled: leaguegamelog, players, boxscoreadvanced, boxscorefourfactors, boxscoremisc, boxscoreplayertrack, boxscorescoring, boxscoretraditional, and boxscoreusage.

These diverse datasets were merged by leveraging identical player_id and game_date fields, enabling their integration into a complete dataset, which was stored directly in a PostgreSQL database. Additionally, the datasets contained comment-based metadata, explaining why a player was absent from a game, including several reasons, with the most common being due to coaching decisions or injuries.

A summary of the schemas for these datasets is provided in Table 1.

Table 1. Players' performance analytics (regular and playoffs), injury analytics, and salary data.

Name (Type)	# Records	# Columns
Player performance statistics (regular)	733,193	132
Player performance statistics (playoffs)	48,213	132
Injury data (on and off game)	58,151	4
Contracts\Salaries data (signed over seasons)	7257	6

2.3.2. Data Engineering

After raw data acquisition, the subsequent phase involved data engineering, which encompassed data cleansing, structuring, and enrichment of the collected datasets to enable in-depth analysis.

The player performance dataset provided a comprehensive overview of advanced player statistics on a per-game-date basis. It encompassed various performance key performance indicators (KPIs), including general player information, advanced box score statistics, four factors, miscellaneous metrics, player tracking, traditional scoring, and usage statistics. While not all performance KPIs were consistently available throughout the entire scraping time frame (for example, specific advanced performance statistics were not recorded in earlier years like 2000), the decision was made to retain as much pertinent information as possible for our comprehensive analysis. Duplicate records, primarily pertaining to primary player reference information such as player_name, were identified and subsequently removed.

The injuries dataset was acquired through nba_api and integrated with our performance dataset. This involved incorporating a comment column within each type of performance sub-dataset, which provided textual descriptions of players' injuries. Due to the lack of standardized formatting in these comments, we applied text-mining techniques to extract the necessary injury information. Subsequently, a customized dictionary was developed to classify injuries into predefined categories. For example, a comment like "torn ACL in left knee (out for season)" would be categorized under "Torn ACL".

Additionally, the dataset contained multiple duplicates stemming from cases where a player missed more than one game due to injury. To address this, we identified and retained only the initial instance along with the date of the initial injury occurrence, applying specific conditions. Consequently, after the textual descriptions were mapped, records were flagged as "duplicate = TRUE" if they referred to the same type of injury for the same player within a 15-day window from the last reported occurrence of that particular injury type.

The contracts dataset required transformation into a more analytically valuable salaries dataset. Once again, we applied text-mining techniques to extract contract lengths and amounts from the textual descriptions within the scraped data. For instance, a contract described as "signed restricted free agent (from Clippers) to a 6-year, $51M contract" was parsed to discern its length (6 years) and amount (USD 51 million).

Furthermore, we utilized inflation rate data obtained from [14,41] to standardize salary figures. This adjustment enabled more meaningful comparisons of player salaries, considering the year the contract was signed and the economic context in the U.S., where the NBA operates.

2.3.3. Data Analysis and Statistical Methodology

The final dataset used for our study was derived from a complex process of data integration, which included advanced performance metrics, injury records, and salary data spanning from the 2000–2001 to the 2022–2023 regular and playoff seasons. This comprehensive dataset consisted of 749,631 records across 158 columns, capturing diverse aspects such as player performance, injury history, and financial information.

The dataset underwent a meticulous two-phase preprocessing operation prior to the commencement of analytical procedures. In the first phase, records were organized in ascending order, first by the "PLAYER_NAME" attribute and then by "GAME_DATE". This organization ensured a chronological sequence of games for each player, facilitating a more structured analysis. Subsequently, the dataset was divided into two distinct subsets. The "Performance" dataset included instances with non-null game dates and corresponding performance metrics. Conversely, the "Injury" dataset contained instances marked with non-null injury dates, focusing specifically on the players' injury histories.

Our statistical analysis comprised a principal test applied to the aggregated performance data. A paired sample t-test was conducted to statistically compare the means of players' salaries before and after injuries. This test generated a t-statistic and an associated

p-value for each anatomical subarea injury category in our dataset associated with players' pre- and post-injury salaries.

The results of these analyses were systematically stored in a PostgreSQL database. This organization into unique categories was designed to simplify further exploratory data analyses and hypothesis testing, thus ensuring a robust and thorough examination of the dataset.

3. Results

Basketball involves substantial uncertainty and interdependence within a multivariate framework. Musculoskeletal injuries make up a significant portion of the health problems in basketball, with a total count of 15,500, which is 65.54% of all reported problems, as presented in Table 2.

Table 2. Number of grouped health problems and percentage allocation.

Health Problems	# Health Problems	% Health Problems
General health problems	7532	31.85%
Head injuries	618	2.61%
Musculoskeletal Injuries	15,500	65.54%
Total	**23,650**	**100.00%**

Based on previous research [2], detailed musculoskeletal injuries span multiple major anatomical regions (e.g., ankle, abdomen, and upper and lower limbs). Injuries were further categorized into anatomical subareas as described in Table 3: the trunk (including the chest, abdomen, and thoracolumbar), the upper extremity (encompassing the shoulder to fingers), and the lower extremity (covering the hips to toes).

Basketball places significant demands on the musculoskeletal system, especially the lower extremities. The data suggest that preventive measures, training modifications, and targeted therapies could be especially beneficial for the knee, ankle, and thigh regions. Furthermore, given the high incidence of musculoskeletal injuries, strength and conditioning programs that focus on the entire musculoskeletal system, as well as proprioception and balance training, may help in reducing the prevalence of these injuries among basketball players.

The knee area stands out as the most injury-prone sub-area, accounting for 23.64% of injuries. The ankle area (18.96%), thigh area (8.08%), and back/thoracolumbar area (10.55%) are also significant. In contrast, areas like the toes (1.73%), upper arm and forearm (0.17%), and fibular area (0.21%) had lower incidences of injuries.

The lower extremities (knee, ankle, and thigh) are the most common areas prone to injury, likely due to the physical demands of the sport/activity in question. These areas are involved in weight-bearing, balance, and propulsion, making them susceptible to injuries. Musculoskeletal injuries are the predominant health problem, reinforcing the physical nature of the sport/activity. It would be worth focusing on strengthening, conditioning, and preventive measures for these areas to reduce injury risks. Although the neck area constitutes a smaller percentage of injuries, any injury in this region can be severe, so preventive measures and safety precautions are crucial. The data provide valuable insights for healthcare professionals, trainers, and athletes to prioritize preventive strategies and interventions to minimize injuries in the most affected areas.

Table 3. Number of anatomical sub-areas of musculoskeletal injuries and percentage allocation split.

Anatomical Sub-Areas	# Anatomical Sub-Areas	% of Anatomical Sub-Areas
Knee area	3664	23.64%
Ankle area	2939	18.96%
Thoracolumbar area	1635	10.55%
Thigh area	1253	8.08%
Foot area	1137	7.34%
Hand, thumb, and fingers area	743	4.79%
Shoulder area	655	4.23%
Hip area	635	4.10%
Abdominal area	603	3.89%
Calf area	516	3.33%
Heel area	385	2.48%
Wrist area	361	2.33%
Toes area	268	1.73%
Elbow area	227	1.46%
Neck	168	1.08%
Shin area	114	0.74%
Chest area	98	0.63%
Pelvic area	39	0.25%
Fibular area	33	0.21%
Upper arm and forearm arm	27	0.17%
Total	**15,500**	**100.00%**

Figure 1 provides data on different anatomical sub-areas, covering metrics such as statistical significance (t-stat and *p*-value), average salary percentage difference post injury, average recovery time, number of players, number of injuries, positive and negative salary changes, average salary change, and correlation with salary change.

The majority of the anatomical sub-areas show "Significant" results based on their *p*-values, indicating that injuries in these areas have a significant impact on the metrics under consideration. However, areas such as the thigh, thoracolumbar area, toes, upper arm and forearm area, and wrist area are deemed "Not Significant".

Injuries in the chest area result in the largest average salary reduction (-25.6%). Surprisingly, some areas, such as the pelvic area and upper arm and forearm area show a positive average salary change after injury ($+13.9\%$ and $+12.6\%$, respectively). This might be due to contracts, insurance, or other external factors not detailed in the table.

Ankle area injuries have the shortest average recovery time (41.9 days), while injuries in the hand, thumb, and fingers area have the longest average recovery time (70.5 days).

In terms of actual salary amounts, the upper arm and forearm area sees the highest increase (USD 1,077,624.8), whereas the chest area observes the most substantial decrease (USD $-155,840.0$).

Anatomical sub-areas	T-Stat	*P*-Value	Significance	Avg Inf Salary Percentage Diff	Average Recovery Time	Number of Players	Number of Injuries	Positive Salary Changes	Negative Salary Changes	Average Salary Change	Correlation with Salary Change
Abdominal area	−7.777	2.2×10^{-14}	Significant	3.4	53.1	824	603	426	398	$ 234,842.8	−8.2%
Ankle area	−7.647	6.0×10^{-14}	Significant	0.0	41.9	788	2939	394	394	$ 264,236.2	−9.6%
Calf area	−5.939	5.2×10^{-9}	Significant	−3.4	44.0	530	516	256	274	$ 13,451.3	−5.3%
Chest area	−5.163	3.6×10^{-7}	Significant	−7.4	23.0	473	98	219	254	$ −155,840.0	−25.6%
Elbow area	−4.353	1.7×10^{-5}	Significant	−11.6	56.5	466	227	206	260	$ 113,171.4	0.5%
Fibular area	−4.129	4.7×10^{-5}	Significant	−11.1	88.0	306	33	136	170	$ −561,391.8	−11.1%
Foot area	−4.068	6.1×10^{-5}	Significant	−6.9	63.1	288	1137	134	154	$ 124,662.3	−5.6%
Hand, thumb and fingers area	−3.518	0.00051	Significant	6.2	70.5	273	743	145	128	$ 258,003.2	−4.4%
Heel area	−3.213	0.00182	Significant	−13.0	56.5	92	385	40	52	$ 265,047.3	−11.0%
Hip area	−2.789	0.00677	Significant	0.0	34.0	72	635	36	36	$ 112,353.9	−5.7%
Knee area	−2.661	0.00813	Significant	−12.9	59.3	365	3664	159	206	$ 114,169.1	−6.8%
Neck	−2.313	0.02194	Significant	−15.2	36.9	165	168	70	95	$ 105,293.8	−12.5%
Pelvic area	−2.305	0.02220	Significant	−13.0	30.8	200	39	87	113	$ 558,358.5	13.9%
Shin area	−2.323	0.02305	Significant	−22.2	69.2	72	114	28	44	$ −230,668.6	−1.1%
Shoulder area	−2.258	0.03294	Significant	15.4	56.3	26	655	15	11	$ 348,183.7	−0.7%
Thigh area	−1.811	0.07241	Not Significant	−9.5	42.6	137	1253	62	75	$ 304,370.7	−9.0%
Thoracolumbar area	−1.613	0.10855	Not Significant	−16.4	39.0	177	1635	74	103	$ 98,353.5	−9.6%
Toes area	−1.637	0.11907	Not Significant	−36.8	45.5	19	268	6	13	$ 318,067.4	−4.3%
Upper arm and forearm area	−1.218	0.22446	Not Significant	−16.0	43.5	238	27	100	138	$ 1,077,624.8	12.6%
Wrist area	0.782	0.44244	Not Significant	−30.4	53.8	23	361	8	15	$ 509,262.0	−9.0%

Figure 1. Musculoskeletal anatomical sub-areas' statistical significance and salary correlation.

Most of the anatomical sub-areas exhibit a negative correlation with salary change post injury.
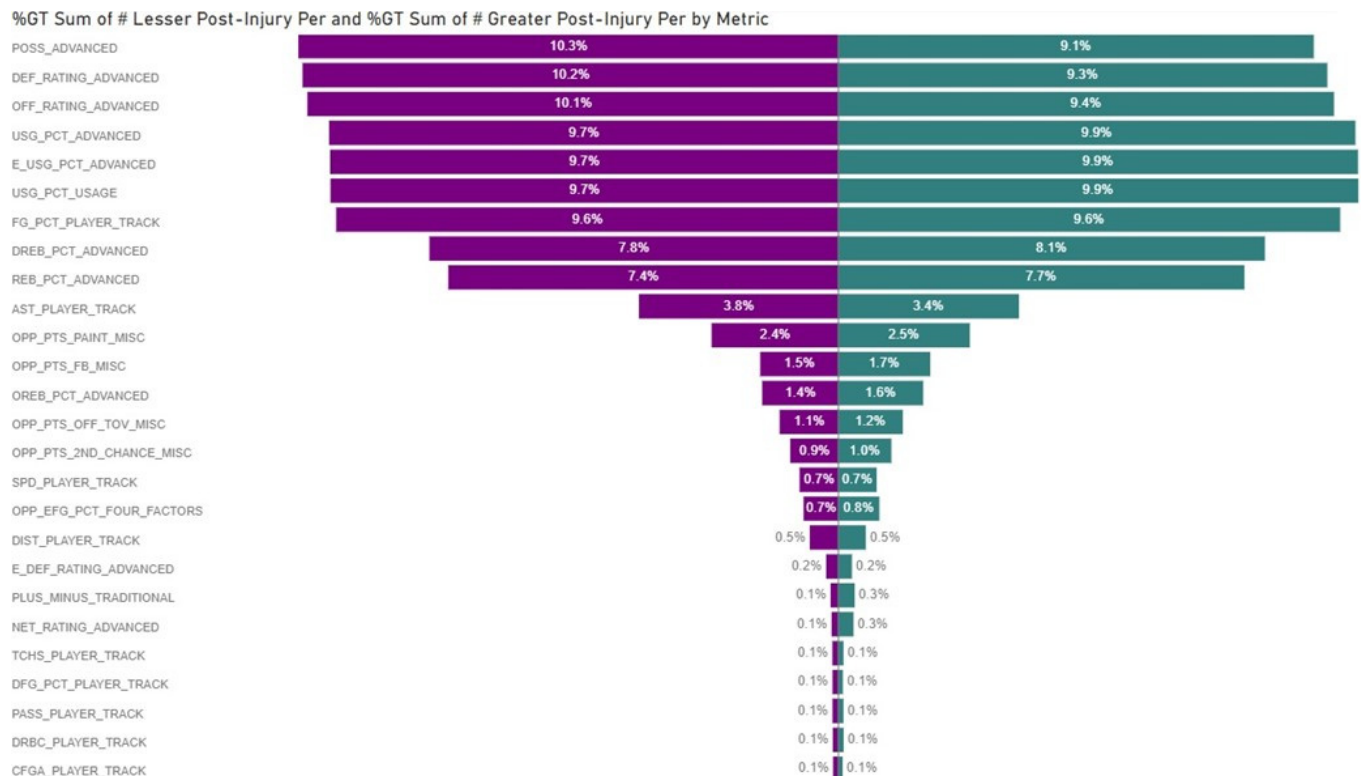
The upper arm and forearm area had the highest positive correlation (+12.6%), while the abdominal area had the most negative correlation (−8.2%).

The Tornado diagram visually represents the variance in basketball performance metrics post injury. Each row signifies a distinct performance metric. On the left side, we have percentages corresponding to "Lesser Post-Injury" performances, while the right-side displays percentages for "Greater Post-Injury" performances in Figure 2. This diagram indicates a variance in basketball performance metrics post injury. Metrics such as possession, defensive and offensive ratings, and usage percentage showed the highest variance, indicating changes in players' performance post injury.

The Figure 2 illustrates a comparative bar chart with two sets of data metrics, shown in purple and green bars. The purple bars represent the percentage of a certain performance metric that is lesser post injury (Lesser Post injury). On the other hand, the green bars likely represent the percentage of metrics that are greater post-injury (Greater Post injury).

The metrics POSS_ADVANCED, DEF_RATING_ADVANCED, OFF_RATING_ADVA NCED, USG_PCT_ADVANCED, and E_USG_PCT_USAGE have the highest variance post injury, with nearly 10% of players showing either lower or greater performances in these areas. The basketball performance analytics terminologies are explained in a previous study [14].

This study reveals that musculoskeletal injuries are prevalent in basketball, constituting 65.54% of all health problems (Table 2). The major anatomical regions affected include the knee, ankle, and thigh. Knee injuries are most common, accounting for 23.64% of the total (Table 3). The data also show significant impacts of injuries on players' salaries in Figure 1, with the largest average reduction observed with injuries in the chest area (−25.6%) and some areas like the pelvic area and upper arm and forearm areas showing a positive average salary change post injury. Recovery times varied across injuries, with ankle injuries having the shortest average recovery time (41.9 days).

Figure 2. Tornado diagram that analyzes the percentage variance in basketball performance analytics in lesser/greater post-injury cases.

4. Discussion

Basketball, as a sport, is characterized by a high degree of uncertainty and the presence of numerous inter-related parameters within a multivariate framework.

The most prominent injuries are localized to areas that bear significant weight or are heavily involved in movement and stability, namely, the knee, ankle, thigh, and back. While some areas, such as the neck and fibular region, have lower incidences, they should not be ignored, as such injuries can be particularly debilitating. Preventive measures, such as strengthening exercises, stretching, and protective gear, might be beneficial, especially for highly impacted areas.

Teams may use these data to invest in specialized medical care and training routines to prevent injuries in high-risk anatomical areas. Player agents might be interested in these data to negotiate contracts, especially ensuring protection and clauses related to injuries in areas that show significant salary implications. Medical staff can prioritize and tailor recovery plans based on average recovery times for different injury types.

Since the highest number of injuries is reported in the knee area, teams should consider specialized training or protective gear to minimize such injuries. Players and agents can use these data during contract negotiations, ensuring protective clauses, especially for areas with significant negative salary implications. The reasons behind the positive salary changes associated with certain injury areas should be explored.

The low variance in metrics at the bottom of the graph may indicate that these areas of performance are less affected by injuries. For instance, metrics related to player tracking, such as distance covered, might not be as influenced by injuries as other performance indicators.

The data suggest that injuries predominantly affect metrics related to possession, defensive and offensive ratings, and usage percentage. This can be attributed to players perhaps being more cautious post injury, leading to a decrease in their on-court activity in these domains.

Interestingly, while some players show a decline in performance post injury in certain metrics, there is an almost equivalent set of players who demonstrate enhanced performances in those same areas. This could indicate a compensatory mechanism where players adapt their playing style post injury, focusing on areas they find more comfortable or less taxing.

This research underscores the physical demands of basketball, particularly on the lower extremities, and the consequent injury patterns. The findings suggest a need for focused preventive measures and training modifications, especially for high-risk areas like the knee, ankle, and thigh.

The impact of injuries on salary and performance metrics highlights the economic and professional implications of these injuries. Teams and player agents can use these data for contract negotiations and tailored injury prevention strategies. The variance in post-injury performance suggests that some players may adapt their playing style to compensate for physical limitations post injury, indicating a potential area for further research.

These results contribute valuable insights into injury patterns and their implications in professional basketball, aiding in the development of more effective injury prevention and management strategies.

5. Conclusions

This study represents a significant advancement in sports analytics, highlighting the criticality of injuries in basketball, particularly in high-risk areas such as the knee, ankle, and foot. It underscores the importance of preventive measures and the strategic management of player injuries.

A notable aspect of our analysis revealed a direct correlation between injuries in specific anatomical regions and significant variations in players' salaries, underscoring the profound economic implications these injuries bear. A particular focus was placed on their prevalence, the anatomical regions most affected, and their consequential impact on player performance metrics. We revealed that injuries predominantly occur in the knee, ankle, and thigh areas, reflecting the intense physical demands placed on lower extremities in basketball.

In particular, the "DEF RATING_ADVANCED" and "OFF RATING_ADVANCED" metrics are significant, suggesting that they are crucial performance indicators. The novelty of this research lies in its comprehensive analysis that integrates injury data with player performance metrics, offering new insights into the impact of injuries on performance.

While exploring the sequence of injuries, our study did not identify a consistent pattern that would suggest a predisposition for subsequent injuries in specific anatomical regions. However, this observation calls for further research with a more extensive dataset to dive deeper into the interconnections of injuries.

The data also revealed intriguing variations in players' performance metrics post injury. While some players exhibited a decline in certain metrics, an equal number demonstrated enhanced performance metrics in the same areas. This indicates a potential adaptive mechanism, wherein players modify their playing style post injury, potentially shifting focus to less physically taxing elements of the game.

Our findings significantly contribute to the expanding field of sports analytics, offering valuable insights into the patterns and consequences of injuries in professional basketball. These insights can assist teams, medical professionals, and players in formulating more effective injury prevention strategies and targeted rehabilitation programs, and making informed decisions about player health and performance. Furthermore, this study lays the groundwork for future research, especially in exploring the long-term effects of injuries on players' careers and the sport's economic dynamics.

This study provides vital information for teams and player agents, informing enhanced medical care, injury prevention strategies, and contract negotiations that consider the potential financial effects of injuries. Particularly, the findings about the knee area, which experiences the most injuries with significant salary implications, demand further ex-

ploration to understand underlying factors, potentially including insurance and post-injury performance.

In conclusion, this research emphasizes the need for a data-informed approach in basketball for injury prevention, strategic contract discussions, and tailored recovery plans, paving the way for future studies to explore the real-world implications of these findings.

Author Contributions: Conceptualization: V.S.; methodology: V.S. and G.P.; software: V.S. and G.P.; validation: C.T., V.S. and G.P.; formal analysis: V.S. and G.P.; investigation: V.S. and G.P.; resources: C.T., V.S. and G.P.; data curation: V.S. and G.P.; writing—original draft preparation: V.S.; writing—review and editing: C.T., V.S. and G.P.; visualization: V.S.; supervision: C.T.; project administration: C.T. and V.S. All authors have read and agreed to the published version of the manuscript.

Data Availability Statement: Data are partly available within the manuscript, further inquiries can be directed to the corresponding author.

Conflicts of Interest: The authors declare no potential conflicts of interest with respect to the research, authorship, and/or publication of this article. This manuscript complies with ethical standards.

References

1. Sarlis, V.; Tjortjis, C. Sports analytics—Evaluation of basketball players and team performance. *Inf. Syst.* **2020**, *19*, 101562. [CrossRef]
2. Sarlis, V.; Chatziilias, V.; Tjortjis, C.; Mandalidis, D. A Data Science approach analysing the Impact of Injuries on Basketball Player and Team Performance. *Inf. Syst.* **2021**, *99*, 16. [CrossRef]
3. Richter, C.; O'Reilly, M.; Delahunt, E. Machine learning in sports science: Challenges and opportunities. *Sport. Biomech.* **2021**, *1*, 1–7. [CrossRef]
4. Kato, S.; Yamagiwa, S. Predicting Successful Throwing Technique in Judo from Factors of Kumite Posture Based on a Machine-Learning Approach. *Computation* **2022**, *10*, 175. [CrossRef]
5. Frank, M.; Drikakis, D.; Charissis, V. Machine-learning methods for computational science and engineering. *Computation* **2020**, *8*, 15. [CrossRef]
6. Mirmozaffari, M.; Alinezhad, A.; Gilanpour, A. Data Mining Apriori Algorithm for Heart Disease Prediction. *Int. J. Comput. Commun. Instrum. Eng.* **2017**, *4*, 1–5. [CrossRef]
7. Abulhasan, J.F.; Grey, M.J. Anatomy and physiology of knee stability. *J. Funct. Morphol. Kinesiol.* **2017**, *2*, 34. [CrossRef]
8. Jauhiainen, S.; Kauppi, J.-P.; Leppänen, M.; Pasanen, K.; Parkkari, J.; Vasankari, T.; Kannus, P.; Äyrämö, S. New Machine Learning Approach for Detection of Injury Risk Factors in Young Team Sport Athletes. *Int. J. Sports Med.* **2021**, *42*, 175–182. [CrossRef] [PubMed]
9. Claudino, J.G.; de Oliveira Capanema, D.; De Souza, T.V.; Serrão, J.C.; Pereira, A.C.M.; Nassis, G.P. Motion sensors-based machine learning approach for the identification of anterior cruciate ligament gait patterns in on-the-field activities in rugby players. *Sensors* **2020**, *20*, 3029. [CrossRef]
10. Claudino, J.G.; de, O. Capanema, D.; de Souza, T.V.; Serrão, J.C.; Pereira, A.C.M.; Nassis, G.P. Current Approaches to the Use of Artificial Intelligence for Injury Risk Assessment and Performance Prediction in Team Sports: A Systematic Review. *Sport. Med. Open* **2019**, *5*, 1–12. [CrossRef] [PubMed]
11. Lao, Y.; Jia, B.; Yan, P.; Pan, M.; Hui, X.; Li, J.; Luo, W.; Li, X.; Han, J.; Yan, P.; et al. Diagnostic accuracy of machine-learning-assisted detection for anterior cruciate ligament injury based on magnetic resonance imaging: Protocol for a systematic review and meta-analysis. *Medicine* **2019**, *98*, e18324. [CrossRef] [PubMed]
12. Dritsas, E.; Trigka, M. Efficient Data-Driven Machine Learning Models for Cardiovascular Diseases Risk Prediction. *Sensors* **2023**, *23*, 1161. [CrossRef] [PubMed]
13. Conde, J.M.; Morán, M.T.C.; Pascual, C.M. Prospective epidemiological study of basketball injuries during one competitive season in professional and amateur Spanish basketball. *Phys. Sportsmed.* **2022**, *50*, 349–358. [CrossRef] [PubMed]
14. Sarlis, V.; Papageorgiou, G.; Tjortjis, C. Sports Analytics and Text Mining NBA Data to Assess Recovery from Injuries and their Economic Impact. *Computers* **2023**, *12*, 261. [CrossRef]
15. Papageorgiou, G.; Sarlis, V.; Tjortjis, C. Unsupervised Learning in NBA Injury Recovery: Advanced Data Mining to Decode Recovery Durations and Economic Impacts. *Information* **2024**, *15*, 61. [CrossRef]
16. Allen, T.; Goff, J.E. Resources for sports engineering education. *Sport. Eng.* **2018**, *21*, 245–253. [CrossRef]
17. Terner, Z.; Franks, A. Modeling player and team performance in basketball. *Annu. Rev. Stat. Its Appl.* **2021**, *8*, 1–23. [CrossRef]

18. Cohan, A.; Schuster, J.; Fernandez, J. A deep learning approach to injury forecasting in NBA basketball. *J. Sport. Anal.* **2021**, *7*, 277–289. [CrossRef]

19. Hewett, T.E.; Myer, G.D.; Ford, K.R.; Paterno, M.V.; Quatman, C.E. Mechanisms, prediction, and prevention of ACL injuries: Cut risk with three sharpened and validated tools. *J. Orthop. Res.* **2016**, *34*, 1843–1855. [CrossRef]

20. Cole, B.; Arundale, A.J.H.; Bytomski, J.; Amendola, A. *Basketball Sports Medicine and Science*; Springer: Berlin/Heidelberg, Germany, 2020. [CrossRef]

21. Trojian, T.H.; Cracco, A.; Hall, M.; Mascaro, M.; Aerni, G.; Ragle, R. Basketball injuries: Caring for a basketball team. *Curr. Sports Med. Rep.* **2013**, *12*, 321–328. [CrossRef]

22. Khan, M.; Ekhtiari, S.; Burrus, T.; Madden, K.; Rogowski, J.P.; Bedi, A. Impact of Knee Injuries on Post-retirement Pain and Quality of Life: A Cross-Sectional Survey of Professional Basketball Players. *HSS J.* **2020**, *16*, 327–332. [CrossRef]

23. Petushek, E.J.; Sugimoto, D.; Stoolmiller, M.; Smith, G.; Myer, G.D. Evidence-Based Best-Practice Guidelines for Preventing Anterior Cruciate Ligament Injuries in Young Female Athletes: A Systematic Review and Meta-analysis. *Am. J. Sports Med.* **2019**, *47*, 1744–1753. [CrossRef]

24. Caplan, N.; Kader, D.F. Knee injury patterns among men and women in collegiate basketball and soccer: NCAA data and review of literature. *Class. Pap. Orthop.* **2014**, 153–155. [CrossRef]

25. Krosshaug, T.; Nakamae, A.; Boden, B.P.; Engebretsen, L.; Smith, G.; Slauterbeck, J.R.; Hewett, T.E.; Bahr, R. Mechanisms of anterior cruciate ligament injury in basketball: Video analysis of 39 cases. *Am. J. Sports Med.* **2007**, *35*, 359–367. [CrossRef]

26. Sundemo, D.; Senorski, E.H.; Karlsson, L.; Horvath, A.; Juul-Kristensen, B.; Karlsson, J.; Ayeni, O.R.; Samuelsson, K. Generalised joint hypermobility increases ACL injury risk and is associated with inferior outcome after ACL reconstruction: A systematic review. *BMJ Open Sport Exerc. Med.* **2019**, *5*, e000620. [CrossRef] [PubMed]

27. Torres-Ronda, L.; Gámez, I.; Robertson, S.; Fernández, J. Epidemiology and injury trends in the National Basketball Association: Pre- and perCOVID-19 (2017–2021). *PLoS ONE* **2022**, *17*, e0263354. [CrossRef]

28. Powers, C.M. The influence of abnormal hip mechanics on knee injury: A biomechanical perspective. *J. Orthop. Sports Phys. Ther.* **2010**, *40*, 42–51. [CrossRef] [PubMed]

29. Bullock, G.S.; Ferguson, T.; Arundale, A.H.; Martin, C.L.; Collins, G.S.; Kluzek, S. Return to performance following severe ankle, knee, and hip injuries in National Basketball Association players. *PNAS Nexus* **2022**, *1*, 1–7. [CrossRef] [PubMed]

30. Hughes, G.T.G.; Camomilla, V.; Vanwanseele, B.; Harrison, A.J.; Fong, D.T.P.; Bradshaw, E.J. Novel technology in sports biomechanics: Some words of caution. *Sport. Biomech.* **2021**, *1*, 1–9. [CrossRef] [PubMed]

31. Andreoli, C.V.; Chiaramonti, B.C.; Buriel, E.; Pochini, A.D.C.; Ejnisman, B.; Cohen, M. Epidemiology of sports injuries in basketball: Integrative systematic review. *BMJ Open Sport Exerc. Med.* **2018**, *4*, e000468. [CrossRef]

32. Kester, B.S.; Behery, O.A.; Minhas, S.V.; Hsu, W.K. Athletic performance and career longevity following anterior cruciate ligament reconstruction in the National Basketball Association. *Knee Surg. Sport. Traumatol. Arthrosc.* **2017**, *25*, 3031–3037. [CrossRef] [PubMed]

33. Cumps, E.; Verhagen, E.; Meeusen, R. Prospective epidemiological study of basketball injuries during one competitive season: Ankle sprains and overuse knee injuries. *J. Sport. Sci. Med.* **2007**, *6*, 204–211.

34. Vidal-Codina, F.; Evans, N.; El Fakir, B.; Billingham, J. Automatic event detection in football using tracking data. *Sport. Eng.* **2022**, *25*, 1–15. [CrossRef]

35. Pérez-Chao, E.A.; Lorenzo, A.; Scanlan, A.; Lisboa, P.; Sosa, C.; Gómez, M.Á. Higher Playing Times Accumulated Across Entire Games and Prior to Intense Passages Reduce the Peak Demands Reached by Elite, Junior, Male Basketball Players. *Am. J. Mens. Health* **2021**, *15*, 1–15. [CrossRef] [PubMed]

36. Short, S.M.; MacDonald, C.W.; Strack, D. Hip and Groin Injury Prevention in Elite Athletes and Team Sport—Current Challenges and Opportunities. *Int. J. Sports Phys. Ther.* **2021**, *16*, 270–281. [CrossRef]

37. W. NBA Basketball Reference. NBA Basketball Reference. 2021. Available online: https://www.basketball-reference.com/ (accessed on 20 August 2023).

38. W. ESPN NBA Stats. ESPN NBA Stats. 2023. Available online: https://www.espn.com/nba/stats (accessed on 10 January 2023).

39. NBA.com. 2023. Available online: https://stats.nba.com (accessed on 1 November 2023).

40. Jansen, B.J.; Moore, K.; Carman, S. Evaluating the performance of demographic targeting using gender in sponsored search. *Inf. Process. Manag.* **2013**, *49*, 286–302. [CrossRef]

41. Vartholomatou, K.; Pendaraki, K.; Tsagkanos, A. Corporate bonds, exchange rates and business strategy. *Int. J. Banking, Account. Financ.* **2021**, *12*, 97–117. [CrossRef]