BACHELOR'S DEGREE THESIS

Degree in Statistics

Title: Methodological Quality and Reporting of Regression Models for Ordinal Responses in Sports Field: A Scoping Review

Author: Mireia Estopañán Moreno

Advisor: Daniel Fernández Martínez; Martí Casals Toquero

Department: DEIO

Academic year: June 2023

UNIVERSITAT DE BARCELONA

UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH
Facultat de Matemàtiques i Estadística

Abstract

The integration of statistical methods in sports science has become essential for decision-making in performance analysis, injury prevention, and athlete outcomes. This work presents a scoping review following PRISMA guidelines to explore the application of ordinal regression models in the sports field. A comprehensive search of articles published, until March 4, 2023, identified 34 included studies. This search included widely recognized databases such as Web of Science, PubMed, and specialized journals of sports statistics, such as Journal of Quantitative Analysis in Sports, and Journal of Sports Analytics. The analysis reveals that 26.5% of these articles were published in statistics and sports statistics journals. However, a significant majority (82.4%) of the studies did not provide data and code repositories. Notably, R emerged as the primary software used for analysis in 38.8% of the studies. Football had the highest representation (28.6%), followed by basketball (17.1%). The most commonly reported ordinal model was the proportional odds model (32.3%), followed by the mixed effects proportional odds model (11.8%), while a relevant proportion (29.4%) did not report the model used. Furthermore, 23.5% of articles proposed novel models. Validation test for proportional odds model were not conducted in 53.3% of cases. This review underscores the importance of improved reporting practices, inclusivity in sport representation, and statistical education in advancing sports analytics. In addition to the scoping review, a real case example to demonstrate the application of the proportional odds model and the mixed effects proportional odds model in the sports field is provided. This case example aims to showcase the practical implementation of these statistical methods and their potential impact on decision-making in sports analytics.

Keywords: *ordinal regression model, ordinal response, data analysis, sport analytics, sports, reporting, scoping review*

Resum

Títol: Qualitat de la metodologia i la informació proporcionada de l'ús de models de regressió per a respostes ordinals en l'àmbit dels esports: Una revisió exploratòria.

Antecedents: La integració dels mètodes estadístics en la ciència de l'esport s'ha convertit en essencial per a la presa de decisions en l'anàlisi del rendiment, la prevenció de lesions i els resultats dels atletes. Els avanços en tecnologia de seguiment de moviments i la gran quantitat de dades disponibles han contribuït a impulsar la investigació estadística en l'àmbit esportiu. Malgrat això, hi ha una manca de directrius específiques per informar adequadament sobre l'ús de models de regressió de respostes ordinals en aquest àmbit. A més, a dia d'avui, no s'ha avaluat l'ús i la qualitat de la informació proporcionada pels models de regressió de respostes ordinals en les ciències de l'esport. És fonamental comprendre aquests models per a una anàlisi adequada de les dades ordinals en l'àmbit esportiu.

Mètodes: Aquest treball presenta una *scoping review* seguint les directrius de l'extensió PRISMA per a *scoping reviews*, per a identificar i descriure l'aplicació de models de regressió ordinal en l'àmbit esportiu. A través d'una cerca exhaustiva d'articles publicats fins al 4 de març de 2023 s'han identificat 34 estudis inclosos. La cerca d'articles ha inclòs reconegudes bases de dades com Web of Science, PubMed, i també revistes d'anàlisi de l'esport com Journal of Quantitative Analysis in Sports i Journal of Sports Analytics.

Resultats: El 26,5% dels articles seleccionats es van publicar en revistes d'estadística i estadística esportiva. Una majoria rellevant (82,4%) dels estudis no va proporcionar repositoris de dades i codi (*data sharing*). És important destacar que el programa R va sorgir com el principal *software* utilitzat per a l'anàlisi en el 38,8% dels estudis. El futbol va tenir la representació més alta (28,6%), seguit del bàsquet (17,1%). El model ordinal informat amb més freqüència va ser el *proportional odds model* (32,3%), seguit del *mixed effects proportional odds model* (11.8%), mentre que una proporció rellevant (29,4%) no va informar sobre el model utilitzat. A més, el 23,5% dels articles va proposar nous models. Els tests de validació per al *proportional odds model* no es van dur a terme en el 53.3% dels casos.

Conclusions: L'ús de models de regressió per analitzar respostes ordinals en l'àmbit esportiu ha augmentat significativament en els últims anys. La majoria de publicacions han aparegut en revistes de disciplines diferents a l'estadística. Aquesta revisió destaca la necessitat de millorar la qualitat a l'hora de reportar la informació referent a les característiques dels models, incloent els mètodes analítics, la validació adequada, la selecció precisa dels models i el requeriment de compartir les dades i programes. Cal establir una col·laboració més estreta entre els estadístics i els professionals d'altres àmbits per millorar la qualitat general dels estudis realitzats. També es ressalta la importància de donar més visibilitat als esports femenins i als amateurs en la literatura científica. És recomanable seguir normes i directrius estandarditzades per a la comunicació dels resultats dels models de regressió ordinal, garantint la transparència i la precisió de les troballes.

A més de la revisió realitzada, es proporciona un exemple de cas real per demostrar l'aplicació del *proportional odds model* i el *mixed effects proportional odds model* en l'àmbit esportiu. Aquest exemple de cas té com a objectiu mostrar la implementació pràctica d'aquests mètodes estadístics i el seu impacte potencial en la presa de decisions en l'anàlisi esportiu.

Paraules Clau: *model de regressió ordinal, resposta ordinal, anàlisis de dades, anàlisi de l'esport, esports, revisió exploratòria*

MSC Classification

The main purpose of the classification of items in the mathematical literature using the Mathematics Subject Classification [MSC] scheme is to help users find the items of present or potential interest to them as readily as possible. An item in the mathematical literature should be classified to attract the attention of all those possibly interested in it.

Mathematics Subject Classification for this paper:
62-07 Data analysis
03E10 Ordinal and cardinal numbers
62J12 Generalized linear models
62F07 Ranking and selection
43A95 Categorical methods [See also 46Mxx]

Summary

1. Introduction 10

1.1. The Importance of Statistics in the Sports Field 10

1.2. Ordinal Data 11

1.3. Types of Reviews in Academic Research 16
1.3.1. Overview for Systematic Reviews 19
1.3.2. Overview for Scoping Reviews 20
1.3.3. Key Considerations Systematic Review versus Scoping Review 21

1.4. Objectives 21

2. Methods 22

2.1. Common Ordinal Regression Models 22
2.1.1. Proportional Odds Model Version of the Cumulative Logit Model 24
2.1.2. Continuation-Ratio Logit Model 26
2.1.3. Adjacent-Categories Logit Model 27
2.1.4. Generalized Partial Credit Model 27

2.2. Choosing and Fitting Ordinal Models 28

3. Methodological Procedures in the Scoping Review 32

3.1. Criteria for the Scoping Review 32
3.1.1. Design 32
3.1.2. Sources of Information and Strategy 32
3.1.3. Eligibility Criteria and Selection of Studies 32
3.1.4. Identification of the Studies 32
3.1.5. Data Extraction 34
3.1.6. Statistical Analysis 35

4. Results 36

4.1. Descriptive Analysis 36
4.1.1. General Characteristics of the Articles 39
4.1.2. General Characteristics of the Sport 40
4.1.3. General Characteristics of the Ordinal Regression Models 41
4.1.4. Bivariate Analysis 44

5. Discussion 45

Limitations 47

Conclusions 47

6. Study Case Example 49

6.1. Abstract 49

6.2. Data Source and Outcomes 49

6.3. Analysis and Interpretation 50
6.3.1. Applying Proportional Odds Model in Sports Research 50
6.3.2. Applying Mixed-Effects POM in Sports Research 56

7. References 62

Appendix 68

List of Figures

Figure 1: Motion tracking markers attached to an athlete's body during training to capture their movements in 3D space. The data compiled from these markers is used for statistical research in sports, such as analyzing technique and designing individualized training programs. Source: (13) 10
Figure 2: Likert scale: This scale measures how much an individual agrees or disagree with a particular statement. Source: (21). 11
Figure 3: The figure displays a comparison between the advantages of treating ordinal data as ordinal and the disadvantages of treating ordinal data as numeric. 12
Figure 4: Ranking The Best NBA Teams In Orlando By Tiers. Source: (29) 13
Figure 5: Ranking The Best NBA Shooting Guards By Tiers. Source: (30) 13
Figure 6: Soccer pitch divided into 18 zones. Source: (32) 13
Figure 7: Showcases several examples of ordinal variables commonly found in the field of sports medicine, which could provide valuable insights for research and analysis. 14
Figure 8: Schematic visualization for some common response variables used in sports medicine research, including the Time-Loss (TL) injuries, the Oslo Sports Trauma Research Center (OSTRC) severity score, the Copenhagen Hip and Groin Outcome Score (HAGOS), the Rating of Perceived Exertion (RPE), the Injury Severity Score (ISS) and the Abbreviated Injury scale (AIS). 15
Figure 9: 'Hierarchy of Evidence Pyramid' adapted from EBP & the Medical Librarian training manual, Duke University 2019, and Online EBM Page Generator, Dartmouth College and Yale University 2019, under the license CC-BY-NC and 113P. Source: (49). 18
Figure 10: Indications for conducting a Systematic Review. Source: (50) 20
Figure 11: Indications for conducting a Scoping Review. Source: (50) 20
Figure 12: Illustration of the proportional odds assumption for 5 response categories. Source: (54). 24
Figure 13: PRISMA-SCR: Flow diagram of the scoping review of the use of ordinal regression models in original articles in the field of sports. 33
Figure 14: **Evolution of Articles Published per Year**. This graph presents the temporal trend of published articles in the sports field that utilize ordinal regression models. The y-axis represents the number of articles published and the x-axis shows the years. The graph provides an overview of the popularity of using ordinal regression models in the sports field over time. 40
Figure 15: **Frequency of Sport Type in which the Study Focuses**. This bar chart displays the percentage and frequency of studies focusing on different types of sports used in the analysis. The length of each bar represents the number of studies in which the corresponding type of sport was analyzed. (Football refers to soccer and futsal, Rugby refers to Australian football). 41
Figure 16: **Frequency of Type of Ordinal Regression Model in which the Article Focuses**. This chart displays the percentage and frequency of different types of ordinal regression model used in the selected articles to perform the analysis. The Proposed Models are with a frequency of 1: Dynamic Cumulative model; Functional Ordinal Logistic Regression model; General Latent Trait model; Generalization to CUB model; Generalized Partial Credit model; Isotonic Least Squares Pairwise Comparison model; Latent Factor Distributed Lag model. And with a frequency of 2: Ordinal Logistic Regression and Expectation. 42
Figure 17: **Type of Validation Method Used per Type of Ordinal Regression Model.** The chart displays the distribution of validation methods across different ordinal regression model types. For this stacked bar plot, each bar represents a specific model type and is segmented based on the types of validation methods used. The height of each segment indicates the frequency of the validation method used for the corresponding model type. 43
Figure 18: **Frequency of Not Reported/Reported validations in POMs**. This chart displays the frequency of POMs that did not report any use of validation method for the proportional odds assumption. 44
Figure 19: Predicted Probabilities of each outcome. This chart displays the predicted probabilities of each outcome category based on the fitted Proportional Odds Model (POM). The x-axis represents

the groin problem prevalence variable, while the y-axis represents the predicted probabilities ranging from 0 to 1. Each block in the chart corresponds to a different outcome category, representing different levels of severity. Red light indicates the player should 'stop' current football activity. Yellow light indicates 'attention', the player should be clinically reviewed by a health professional before football participation is decided on, and green light indicates 'go' for football participation. 53
Figure 20: Predicted Probabilities of each outcome. This chart displays the predicted probabilities of each outcome category based on the fitted Proportional Odds Model (POM). The x-axis represents the time-loss groin problem variable, while the y-axis represents the predicted probabilities ranging from 0 to 1. Each block in the chart corresponds to a different outcome category, representing different levels of severity. Red light indicates the player should 'stop' current football activity. Yellow light indicates 'attention', the player should be clinically reviewed by a health professional before football participation is decided on, and green light indicates 'go' for football participation. 55
Figure 21: This chart displays the predicted probabilities of each outcome category based on the fitted Mixed-Effects Proportional Odds Model. The x-axis represents the groin problem variable, while the y-axis represents the predicted probabilities ranging from 0 to 1. Each box and number correspond to a different outcome category, representing different levels of severity. 1 indicates the player should 'stop' current football activity. 2 indicates 'attention', the player should be clinically reviewed by a health professional before football participation is decided on, and 3 indicates 'go' for football participation. 58
Figure 22: This chart displays the predicted probabilities of each outcome category based on the fitted Mixed-Effects Proportional Odds Model. The x-axis represents the time-loss groin problem variable, while the y-axis represents the predicted probabilities ranging from 0 to 1. Each box and number correspond to a different outcome category, representing different levels of severity. 1 indicates the player should 'stop' current football activity. 2 indicates 'attention', the player should be clinically reviewed by a health professional before football participation is decided on, and 3 indicates 'go' for football participation..... 59

List of Tables

Table 1: Description about the main review types. Source: (42) 17
Table 2: Summary table of common ordinal regression models, including their advantages, disadvantages, and recommended use cases. 28
Table 3: Summary table providing a list of R packages that can be used to fit ordinal models. 30
Table 4: Descriptive table of the general characteristics for the selected articles 34
Table 5: Descriptive table of the general characteristics of the sport 34
Table 6: Descriptive table of the general characteristics of the ordinal regression models 35
Table 7: Frequency distribution of general characteristics of the selected articles by group. 36
Table 8: Frequency distribution of general characteristics of the sport in which focuses the articles by group. 37
Table 9: Frequency distribution of general characteristics of the ordinal regression model used in the selected articles by group. 38
Table 10: Cross Table of Statistical Software Not Reported in the selected articles per Type of Model Not Reported 44
Table 11: Proportional Odds Model estimates of pain monitoring between players reporting and player not reporting groin problems. 51
Table 12: Cumulative probabilities for players reporting groin problem prevalence and players not reporting groin problems. 52
Table 13: Odds for players reporting groin problem prevalence and players not reporting groin problems. 52
Table 14: Predicted probabilities for players reporting groin problem prevalence and players not reporting groin problems. 53
Table 15: Proportional Odds Model estimates of pain monitoring between players reporting groin problems with and without time-loss. 54
Table 16: Cumulative probabilities for players reporting groin problems with and without time loss. 54
Table 17: Odds for players reporting groin problems with and without time loss. 54
Table 18: Predicted probabilities for players reporting groin problems with and without time loss. . 55
Table 19: Test result for assessing the proportional odds assumption. 57
Table 20: Mixed Effects Proportional Odds Model estimates of pain monitoring between players reporting and player not reporting groin problems. 58
Table 21: Test result for assessing the proportional odds assumption. 59
Table 22: Mixed-Effects POM estimates of pain monitoring between players reporting groin problems with and without time-loss. 59
Table 23: Extended summary table for model 3 (Mixed Effects Proportional Odds Model of HAGOS, Sport/Rec subscale scores between players reporting and player not reporting groin problems) and model 4 (Mixed-Effects POM of HAGOS, Sport/Rec subscale scores between players reporting groin problems with and without time-loss). 68
Table 24: Chart visualizations of the alternative testing for the proportional odds assumption, plotting the observed mean of the covariate against the expected mean implied by model 3 (Mixed Effects Proportional Odds Model of HAGOS, Sport/Rec subscale scores between players reporting and player not reporting groin problems) and model 4 (Mixed-Effects POM of HAGOS, Sport/Rec subscale scores between players reporting groin problems with and without time-loss). 69

1. Introduction

1.1. The Importance of Statistics in the Sports Field

The relationship between the statistical community and sports science is increasingly close and necessary to make decisions on performance-related goals, injury prevention, athlete outcomes and also the development of new statistical techniques (1–4). Over the years, the use of statistics in sports has been of great interest, leading to the development of a field of research that has gained significant attention. In recognition of the growing importance of sports statistics, organizations such as the American Statistical Association (ASA) and the International Statistical Institute (ISI) have created sections and committees dedicated to this field.

The recent surge in opportunities for statistical research in sports can be attributed to advancements in motion tracking technology (**Figure 1**), and the explosion of available data (5). This is reflected in the establishment of dedicated journals such as the *Journal of Quantitative Analysis in Sports* (founded in 2006) and the *Journal of Sports Analytics* (founded in 2014). Regular meetings and conferences, such as MathSport International, New England Symposium on Statistics in Sports (SSiS), Carnegie Mellon Sports Analytics Conference (CMSAC), and Joint Statistical Meetings (JSM), provide opportunities for researchers to present their work and advance the field. In the sports statistics literature, the *Handbook of Statistical Methods and Analyses in Sports* (6), offers a comprehensive overview of current research in the field of sports analytics. The growth of this field has been further fueled by the establishment of sports analytics departments and groups in universities, such as Harvard (7), Berkeley (8), Syracuse (9) and Carnegie Mellon (10). Additionally, the popularity of films like *Moneyball* (11) and *Concussion* (12) have helped to raise awareness and interest in sports analytics among a wider audience, including fans, coaches, and team owners.

***Figure 1**: Motion tracking markers attached to an athlete's body during training to capture their movements in 3D space. The data compiled from these markers is used for statistical research in sports, such as analyzing technique and designing individualized training programs. Source: (13)*

As Baumer et al. (14) pointed out, “sports analytics has expanded its footprint both in the professional sports industry and in academia over the past 30 years”, leading to strategic changes in various sports. This work has mainly been promoted with four ideas that are common across multiple sports, such as the expected value of a game, win probability, measures of team strength, and the use of sports betting market data. Insights derived from the analysis of data have transformed the world of sports over the last few decades, with significant developments. In fact, it is now difficult to find a professional sport in which sports analytics is not having an impact.

Academia and industry both play a role in the development of these insights. Researchers use public data to produce high-quality, peer-reviewed scientific work, while analysts work with players, coaches, and team officials to put these new ideas into immediate effect thanks to high-resolution data. Performing analytical work requires to compute with data and rely on computational tools that enable the processing and analysis of data (14). These tools are constantly evolving, making it necessary for sports analysts to stay up to date with the latest resources available. R has emerged as a popular open-source software for statistical analysis and sports analytics (15). The Comprehensive R Archive Network (CRAN) is a valuable resource for sports analytics practitioners, as it provides a vast selection of R packages for different sports (16). A recent review (17) of the sport-related packages on CRAN revealed a total of 81 packages. Most of the packages focus on sports performance analysis, with web scraping being a common technique used.

1.2. Ordinal Data

An ordinal variable is described as such when there is a clear ordering of the categories, but the absolute distances among them are unknown. Distinct levels of such variable differ in degree of dissimilarity more than in quality (18,19). Ordinal variables are of a very different nature than nominal variables, which distinct levels differ in quality, not in quantity, and thus the order of listing of the categories is unimportant (18). Some examples of ordinal variables include the Pain scale, the Likert scale, or the Braun-Blanquet scale in vegetation sciences. The Pain scale is a scale that measures pain intensity levels, ranging from no pain to the worst possible pain. It is typically scored from 0 to 10, with 0 indicating no pain and 10 indicating the worst possible pain. The Pain scale is often used to measure pain levels in sports injuries or in rehabilitation programs. Additionally, the Likert Scale is used to measure attitudes or opinions. It consists of a series of statements that are rated on a scale from not satisfied at all to extremely satisfied (e.g., see **Figure 2**). In Sports Sciences, the Likert scale is commonly used to measure subjective perceptions or opinions of athletes in relation to sports performance or training programs. This scale typically ranges from 1 to 5 or 1 to 7, with higher scores indicating greater agreement with the statement.

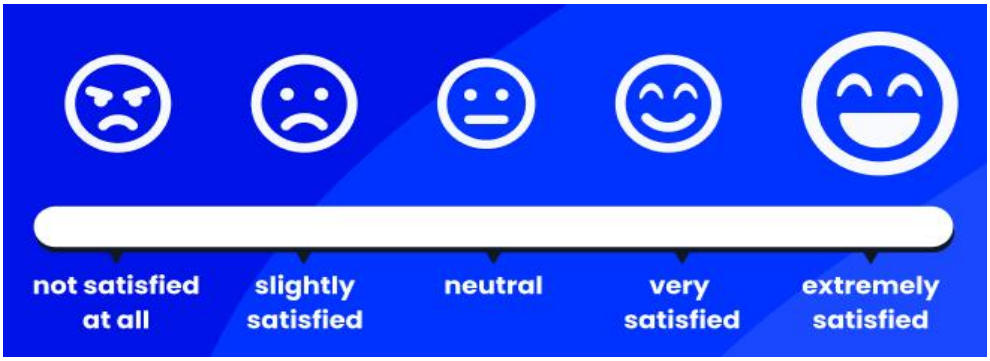


***Figure 2**: Likert scale: This scale measures how much an individual agrees or disagree with a particular statement. Source: (21).*

In the past decade, there has been a significant rise in the utilization of methods for analyzing ordinal data, as reflected in the growing number of literature published on the subject (18). Multiple studies (20–22) have confirmed this trend, revealing an increasing use of ordinal variables and outcomes across various fields, including the health and social sciences. These studies also mention the emergence of new statistical methods for analyzing ordinal data, emphasizing the need for a better

understanding of how these outcomes are used in practice. These serves as an example of the ongoing efforts to develop and improve the methods used for analyzing ordinal data. This increase has gone hand in hand with the advancements made in these techniques and the introduction of innovative methods (18,19), such as the proportional odds model version of the cumulative logit model, and the ordered stereotype model with ordinal scores (19) from with new lines of research have developed.

There are many advantages to be gained from using ordinal methods (see a summary in **Figure 3**) instead of the standard nominal procedures (18) such as simpler description of the data, more powerful inferences about population characteristics, and greater variety of models, most of which are more parsimonious and, thus, have simpler interpretations than the standard models for nominal variables (14,15). However, ordinal variables, although extremely common, are almost exclusively analyzed with statistical models that falsely assume them to be metric (19), or nominal (23), leading to inadequate solutions. On the one hand, the nominal approach ignores the intrinsic ordering of the data and, thus, the results are less powerful than they could be. On the other hand, models for continuous variables have similarities to those for ordinal variables, (like data description measures, such as correlations, slopes, and means), but the use of them with ordinal variables has many disadvantages (see please **Figure 3**). For instance, in rating scales, ignoring that the response is ordinal and using classical regression models that assume that the response is at least on an interval scale (24) can lead to distorted effect-size estimates, inflated error rates, failures to detect effects, and loss of information, since this approach does not fully utilize the available information (25,26).

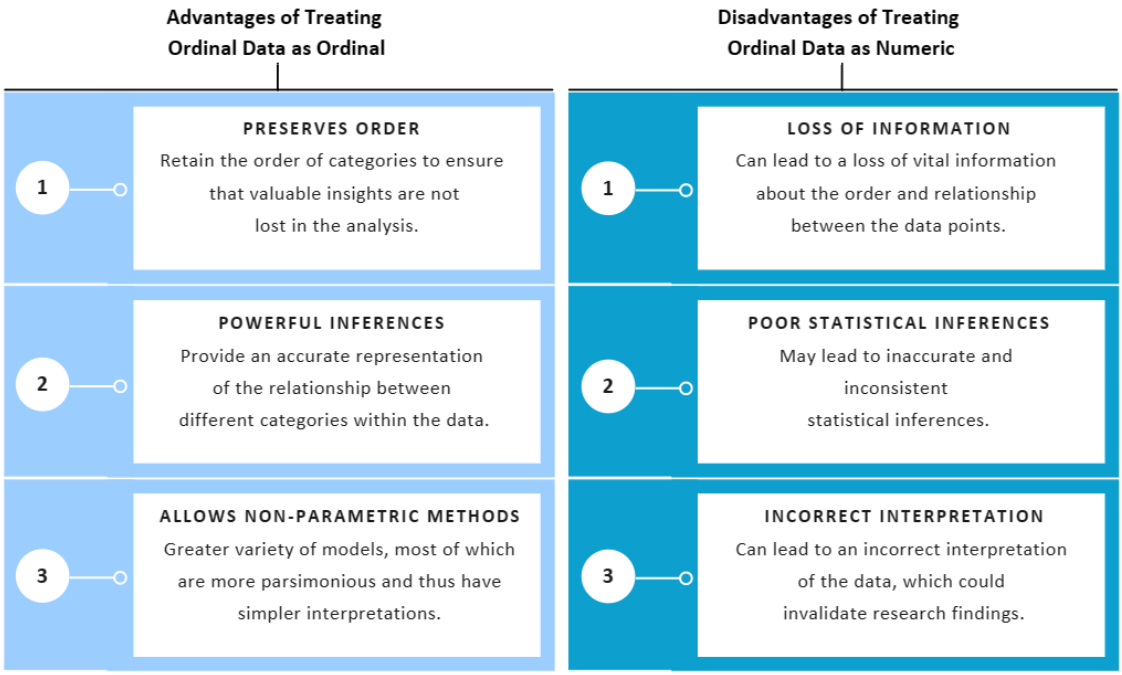


***Figure 3**: The figure displays a comparison between the advantages of treating ordinal data as ordinal and the disadvantages of treating ordinal data as numeric.*

Ordinal Response Data in the Sports Field

Nowadays, many studies in the Sports Field involve response variables that are measured on ordinal scales, and thus inherently non-normal in their distribution (27,28). In sports one is usually interested in the final rankings of teams (see **Figure 4**), players (see **Figure 5**), and final results, such as times and distances. Also, coaches or players aim to identify the crucial areas of performance that require attention, evaluate the effectiveness of tactical and technical performance, and assess the strengths and weaknesses of upcoming oppositions (29). For that, categorical variables measured on ordinal scales can be used in decision making.

***Figure 4**: Ranking The Best NBA Teams In Orlando By Tiers. Source: (29)*

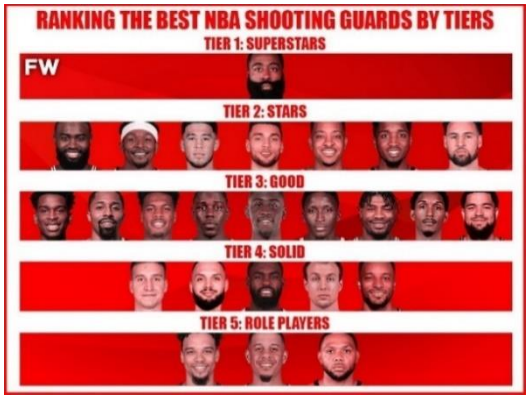


***Figure 5**: Ranking The Best NBA Shooting Guards By Tiers. Source: (30)*

In sports performance analysis, we have variables representing different locations of the playing area. For example, a soccer pitch could be divided into 18 areas (Hughes and Franks, 2004b). The 18 areas that Hughes and Franks (2004b) denoted 'A' to 'V' are not merely different values of an ordinal scale measure because some are neighboring (30).

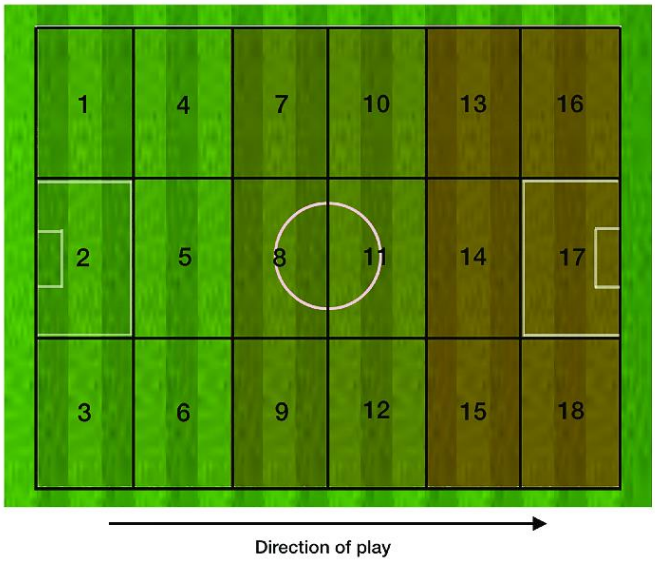


***Figure 6**: Soccer pitch divided into 18 zones. Source: (32)*

Moreover, within the field of sports medicine, ordinal responses are highly common (**Figure 7**). In sports medicine one is usually concerned with quantifying injury occurrence with respect to who is affected by injury, where and when injuries occur and what is their outcome, for the purposes of explaining why and how injuries occur and identifying strategies to control and prevent them (31).

Many studies use as response variables the Injury Severity Score (ISS), the Abbreviated Injury Scale (AIS), the Time-Loss (TL) injuries, the Oslo Sports Trauma Research Center (OSTRC) severity score, the Copenhagen Hip and Groin Outcome Score (HAGOS), or the Rating of Perceived Exertion (RPE), which have a natural ordering of severity or certainty (25,26). In addition to these measures, there have been proposed new models to categorize injuries, such as the Subsequent Injury Categorization (SIC) model presented by Finch et al. (32). These models aim to further enhance the understanding of injury severity and recurrence in sports studies.

Ranking the effectiveness of a rehabilitation program.

Assessing functional ability of an athlete after an injury.

Scoring the level of fatigue after a workout.

***Figure 7**: Showcases several examples of ordinal variables commonly found in the field of sports medicine, which could provide valuable insights for research and analysis.*

Furthermore, in the following section, a brief explanation of some of the mentioned response variables, their application in sports medicine, and their respective ordinal scales will be provided. This will allow to better understand the context and significance of ordinal data in sports medicine research. To aid in visualizing the variables and their corresponding levels, a schematic image will be included in **Figure 8**.

- Time-loss Injuries

Time-loss (TL) injuries are defined as injuries that results in a player being unable to take a full part in future training or match play, and the severity of an injury can be classified according to the number of days that have elapsed from the date of injury to the date of the player's return to full participation in team training and availability for match selection (33). However, there is a lack of common conceptual basis in TL injuries, since some studies classify a severe injury as one that results in five games being missed, whereas others classify a severe injury as one that requires five weeks out of competition to heal (31). Also, this approach underestimates the full impact of overuse injuries because athletes with an overuse injury can often continue to train and compete despite persistent injury-associated symptoms and limitations.

- Oslo Sports Trauma Research Center Severity Score

In order to record the magnitude, symptoms, and consequences of overuse injuries in sport, the Oslo Sports Trauma Research Center (OSTRC) developed an Overuse Injury Questionnaire (34). The OSRTC questionnaire, consists of four questions scored on a Likert scale that reflects the impact of health problems on participation, training volume, sports performance, and symptoms as reported by the athlete. The responses to each of the four questions are allocated to a numerical value from 0 to 25,

such that 0 represents no problems and 25 represents the maximum level for each question. These are summed in order to calculate the OSTRC severity score from 0 to 100 for each overuse problem. Is used to measure and monitor the severity of the health problem (35).

- Hip and Groin Outcome Score

Pain in the hip and groin region is a common musculoskeletal complaint in the young to middle-aged population affecting physical function and health-related quality of life (QOL). The HAGOS is a valid and responsive tool specifically designed to assess hip and groin pain and function, and it contains 6 separate subscales: Pain, Symptoms, Physical Function in Daily Living, Physical Function in Sports and Recreation and Hip- and Groin-related Quality of Life. All HAGOS subscales provide independent scores ranging from 0 to 100, where 0 indicates extreme groin pain and limitations, whilst 100 infers free of symptoms and full function (36).

- Rating of Perceived Exertion

Training load monitoring is crucial in sports to optimize performance and prevent negative outcomes (37). It involves assessing internal and external loads to understand an athlete's physiological response to training and the various modalities imposed on their body. Internal load, determined by the intensity and duration of stress, reflects the body's adaptation to training, while external load captures the overall training volume measured through GPS and accelerometers (38). To evaluate internal load, coaches can rely on RPE (Rated Perceived Exertion), which are accessible and non-invasive tools. RPE, measured on a scale of 0 to 10 (nothing at all to very, very heavy), correlates strongly with heart rate and GPS measurements, showcasing athletes' accurate self-assessment (38). Furthermore, RPE not only indicates training intensity but also reveals information about progressive and accumulated fatigue that may not be captured by conventional markers like heart rate and lactate concentrations. (39).

Common Ordinal Response Data in Sports Medicine

TL injuries	Number of training days lost to a particular injury, with longer time loss indicating greater severity. There is lack of common conceptual basis in the scale of TL injuries.
OSTRC Severity Score	Measures the severity of an injury based on the number of symptoms, functional limitations, and duration of symptoms. Scores from 0 to 100 for each overuse problem.
HAGOS	Measures the severity of symptoms and functional limitations in patients with hip and groin disorders. Scores ranging from 0 to 100.
Rating of Perceived Exertion	Scale of 1 to 10 used to provide insights into the perceived intensity of exercise and accumulated fatigue.
Injury Severity Score	Score based on the severity of the injuries sustained in different body regions. The scale ranges form 0 to 75.
Abbreviated Injury Scale	Uses a scale of 1 to 6 to rate the severity of an injury in different body regions.

***Figure 8**: Schematic visualization for some common response variables used in sports medicine research, including the Time-Loss (TL) injuries, the Oslo Sports Trauma Research Center (OSTRC) severity score, the Copenhagen Hip and Groin Outcome Score (HAGOS), the Rating of Perceived Exertion (RPE), the Injury Severity Score (ISS) and the Abbreviated Injury scale (AIS).*

As seen above, the severity of the sport injury can be expressed in many ways. However, it is common for authors investigating injury burden in athletic populations to focus on TL injuries (40). For instance, the paper from Hägglund et al. 2016 (41) aimed to compare the rate of injury recurrence in professional football players at the highest level of competition to those playing at national and amateur levels. The study used ordinal response data to classify the severity of recurrent injuries, with time-loss (TL) injuries collected from injury surveillance of professional, national, and amateur teams. Injury severity was based on the number of days lay-off and divided into four categories: slight/minimal (0–3 days), mild (4-7 days), moderate (8-28 days), and severe (>28 days) (41).

To properly analyze ordinal data in the sports field, a strong understanding of ordinal response regression models is crucial. Despite their significance and increasing use in academic and professional settings, there is currently a lack of guidelines that specifically address the appropriate reporting of these techniques. Furthermore, there have been no evaluations of the use and quality of information reported by ordinal response regression models in the field of sports field. This lack of information can hinder decision making by professionals, and the public when evaluating the worth and credibility of the information they are examining.

1.3. Types of Reviews in Academic Research

In recent years, modern world and their expansion each day faster, are contributing to increase all kind of advances, which have led to the development of a large amount of publications and research in all fields, due to the need to constantly share and update information. Also, the past two decades have seen a growing emphasis on basing decisions on the best available evidence. This flood of reported information has gradually generated the requirement to establish common minimum standards of quality resulting in the development of reporting guidelines and evidence synthesis or reviews and, with this, the recognition that not all evidence is equal in terms of its validity (42–46). In fact, there is a growing concern about the rigor and quality of statistical analysis in the field of sports sciences. As noted in the article (47), the use of inappropriate statistical methods can lead to erroneous conclusions and hinder the advancement of knowledge in the field. Therefore, it is essential to establish minimum standards of quality and rigor in sports research, as well as to increase the use of reporting guidelines. These guidelines specify minimum criteria for authors to report their research such that readers can both critically appraise and interpret study findings (48). This minimum set of items for reporting (guidelines) can improve the accuracy and transparency of publications, thus facilitating easier and more reliable appraisal of quality and relevance (49). This is particularly important in sports sciences, where the use of rigorous statistical methods is necessary to effectively evaluate athlete outcomes, make informed decisions in performance analysis, and assess the effectiveness of injury prevention measures. The expansion of evidence-based practice across sectors has led to an increasing variety of review types (see **Table 1**).

***Table 1**: Description about the main review types. Source: (42)*

Label	Description	Strengths	Weaknesses
Systematic Review	Is the best-known form of review, where research evidence is systematically searched for, assessed, and summarized, following	Aim to gather all available knowledge on a particular topic. The process is transparently reported to make it easier for others to replicate.	Limit studies to a single design. May not answer complex questions such as why an intervention is effective.
Scoping Review	Provides a preliminary assessment of the potential size and scope of available research literature. It aims to identify the nature and extent of research evidence (usually including ongoing research).	Can help policymakers decide if a full systematic review is necessary. Like systematic reviews in their attempt to be systematic, transparent, and replicable.	Cannot be considered as a final output because they have limitations in rigor and duration, which can lead to bias. Typically, do not include quality assessment, so their findings may be based on the existence of studies rather than their quality.
Rapid Review	Assessment of what is already known about a policy or practice issue, by using systematic review methods to search and critically appraise existing research.	Quicker versions of systematic reviews, to shorten the timescale. Can use techniques like simplifying the question, using less complex search strategies, and limiting the amount of information they gather.	Curtailing the duration of the review process runs the risk of introducing bias. May not cover everything in much detail.
Umbrella Review	Review compiling evidence from multiple reviews into one accessible and usable document. Focuses on a broad problem for which there are two or more competing interventions and highlights reviews that address these interventions and their results	Useful when reviewers face a dilemma of whether to do a broad review that covers multiple interventions (but at the cost of lost detail) or multiple focused reviews that address specific comparisons (but at the risk of fragmenting the overall picture).	Relies on the existence of narrower component reviews, which may not always be available. May not be practical for all areas of library and information practice.
Narrative or Literature Review	Its purpose is to identify a few studies that describe a problem of interest. Have no predetermined research question or specified search strategy, only a topic of interest. They are not systematic and follow no specified protocol	Helps consolidate knowledge and provide a summary of what is already known on a particular topic, so that researchers can build on previous work, avoid repeating what has already been done, and fill in any gaps.	Can omit significant parts of the literature and be influenced by bias. Authors may selectively choose literature that supports their hypothesis, which can undermine the validity of their conclusions.
Critical Review	Aims to demonstrate that the writer has extensively researched the literature and critically evaluated its quality. Includes a degree of analysis and conceptual innovation. Typically, results in hypothesis or model, not an answer.	Evaluate previous work and provide an opportunity to identify what is valuable, resolve competing ideas, and launch new phases of conceptual development.	There is no formal quality assessment. The resulting product serves as a starting point for further evaluation, not an endpoint.

Mixed Methods Review	Combines different research approaches to answer complex questions. It typically involves combining a systematic review of quantitative data with a qualitative review of attitudes or implementation issues.	It can address weaknesses of other approaches, such as systematic reviews of effectiveness and theory-driven reviews. Offer a more complete picture of the research landscape in a specific topic area.	Can be challenging to appraise and synthesize both quantitative and qualitative research and integrate the results. Bringing together studies with different structures, addressing different questions, and conducted within different paradigms can be complex.

To aid the interpretation and evaluation of research findings, hierarchies of evidence have been developed to rank research based on its validity. Not all research designs carry the same level of risk of error and bias in their results (44). Therefore, certain research methods are more reliable than others when trying to answer specific questions. That is, the validity of research results can vary depending on the methods used. The Evidence-Based Pyramid (see **Figure 9**) is a helpful diagram that illustrates how to weigh different levels of evidence when making health-related decisions (50). This pyramid represents the idea of hierarchy of medical evidence, where not all evidence is the same. Systematic reviews and meta-analyses are positioned at the top of the pyramid due to their rigorous approach to evidence gathering and analysis.

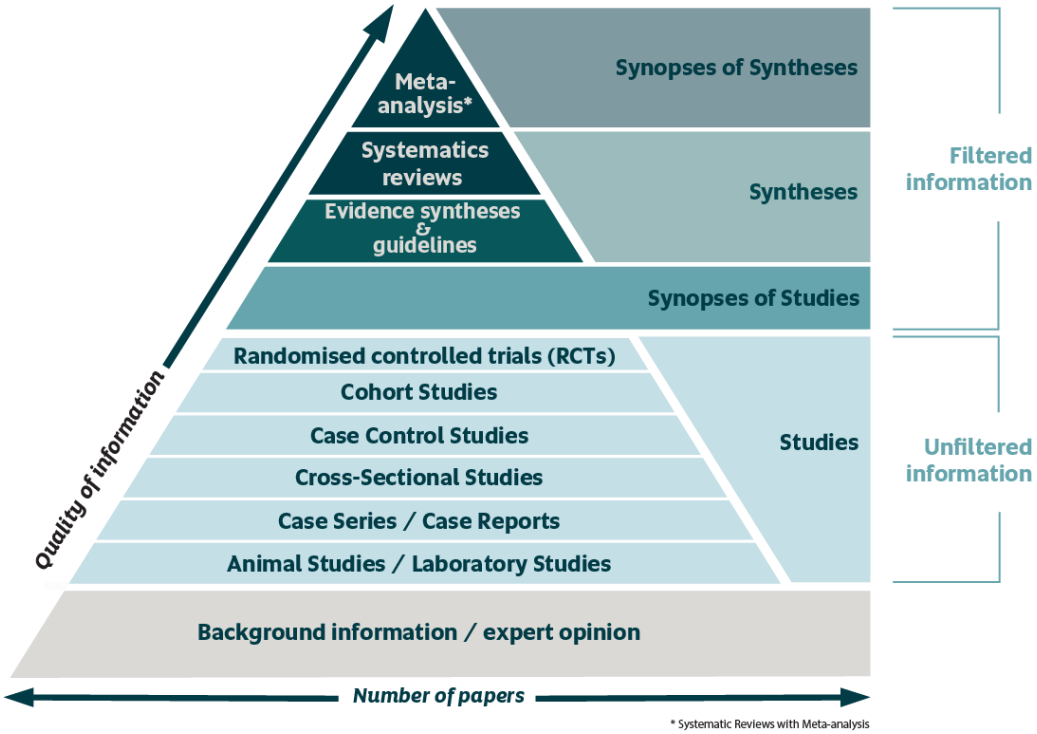


***Figure 9**: 'Hierarchy of Evidence Pyramid' adapted from EBP & the Medical Librarian training manual, Duke University 2019, and Online EBM Page Generator, Dartmouth College and Yale University 2019, under the license CC-BY-NC and 113P. Source: (49).*

To conduct a review, it is important to identify the specific needs that the study aims to address, since there are different types of scientific literature reviews, and choosing the most suitable type will depend on the specific needs of the study. Therefore, it's crucial to determine the research question or problem that the study aims to answer in order to select the most appropriate type of literature review. Also, it is important to consider several factors like the following:

- The type of data available on the topic. If the data is primarily quantitative, a systematic review or meta-analysis may be more appropriate. If the data is primarily qualitative, a narrative or scoping review may be more suitable.
- How much time you have, to complete the review. Some type of reviews, such as systematic reviews, can be time-consuming and resource intensive.
- What resources are available to conduct the review. Some type of reviews may require specialized knowledge or software.
- The methodological rigor. Some types of reviews, such as systematic reviews, require a high level of rigor in terms of study selection, data extraction, and data analysis. On the other hand, other types of reviews may not require such a rigorous approach.
- What is the purpose of the review. Some reviews are intended to identify gaps in the literature, while others aim to synthetize the existing knowledge or evaluate the effectiveness of an intervention. The purpose of the review will influence the type of review that is most appropriate.

Even though systematic reviews are considered the pinnacle of evidence in health care and their prominent position within the evidence hierarchy (as **Figure 9** shows), there are circumstances where systematic reviews are not practical or relevant for the review objective or knowledge user needs, like in this particular study. In these instances, one of the most common alternative approaches to knowledge synthesis are scoping reviews. Although there is clear guidance on the definition and conduct of scoping reviews, there is still a lack of clarity for authors regarding when to perform a scoping review as opposed to a systematic review. In this context, the article *Systematic review or scoping review? Guidance for authors when choosing between a systematic or scoping review approach* (51) is useful in providing guidance on when a scoping review is (and is not) appropriate. We will see in which situations fits better each review, since understanding the specific circumstances can help determine which review approach is best suited for the task.

1.3.1. Overview for Systematic Reviews

Systematic reviews are a rigorous type of research synthesis that are conducted to gather and evaluate relevant international evidence. These reviews require focused clinical questions and for this, PICO (Patient, Intervention, Comparison, Outcome) is a commonly useful tool used to develop such questions for quantitative systematic reviews. Authors of a systematic review conduct a comprehensive literature review, exclude poorly conducted studies, and provide practice recommendations based on the well-performed studies. These reviews focus solely on experimental or quantitative studies, often involving randomized controlled trials. The purpose of conducting a systematic review is to establish the quality of existing evidence, confirm or refute current practices, address any uncertainty or variation in practice, and identify gaps, deficiencies, and trends in the evidence. These reviews can also inform future research in the area under study.

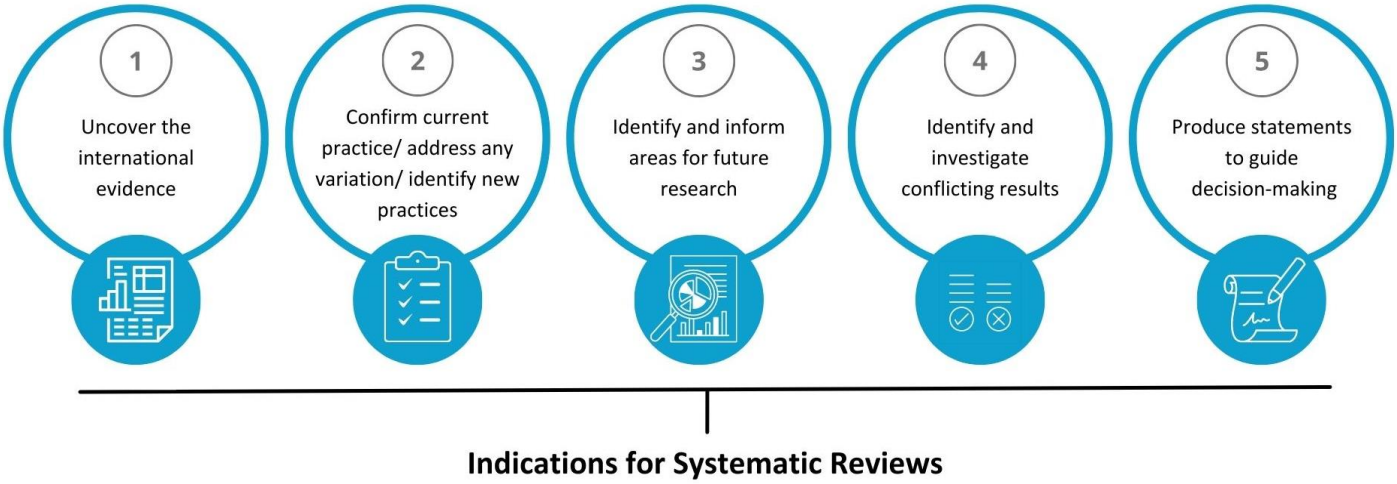


***Figure 10**: Indications for conducting a Systematic Review. Source: (50)*

1.3.2. Overview for Scoping Reviews

Scoping reviews are a valuable tool for determining the coverage and scope of literature on a particular topic. They provide an overview of the available literature, including the volume and focus of the studies. These reviews are particularly useful for examining emerging evidence and identifying gaps in knowledge before conducting more specific systematic reviews. They are also conducted when a narrow review question cannot be defined, studies use a variety of data collection and analysis techniques, no prior synthesis has been undertaken, and the quality of studies is not necessarily assessed. Overall, scoping reviews are a crucial step in understanding the available evidence and identifying areas where further research is needed. The general purpose for conducting scoping reviews is to identify and map the available evidence.

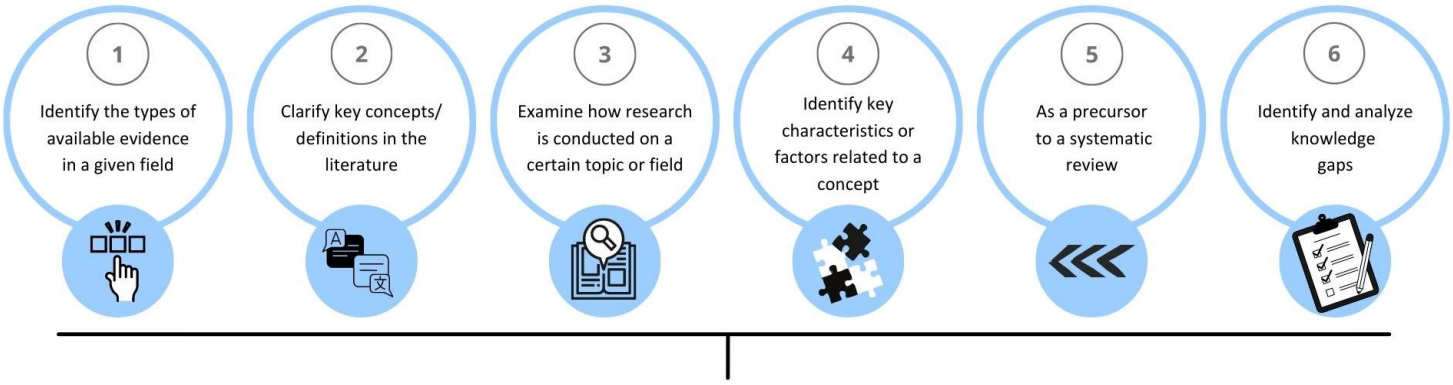


***Figure 11**: Indications for conducting a Scoping Review. Source: (50)*

This study presents a scoping review, aiming to provide a comprehensive understanding of the current state of research. Given the growing interest in the application of ordinal regression models in sports field, a scoping review is a necessary step in identifying the available evidence and evaluating the quality of studies in this area. By mapping and synthesizing the available evidence, this scoping review

seeks to contribute to the advancement of knowledge on the use of ordinal regression models in this field, ultimately helping to guide future research efforts and inform evidence-based decision making.

1.3.3. Key Considerations Systematic Review versus Scoping Review

When choosing between a systematic review and a scoping review, the most important consideration noted by Munn et al. (51) is "whether or not the authors wish to use the results of their review to answer a clinically meaningful and specific question or provide evidence to inform practice. If the authors have a specific question addressing the feasibility, appropriateness, meaningfulness or effectiveness of a certain treatment or practice, then a systematic review is likely the most valid approach". Systematic reviews are rigorous and focused, and prioritize well-conducted studies that address the effectiveness, appropriateness, and feasibility of certain treatments or practices.
However, authors do not always wish to ask such single or precise questions and may be more interested in the identification of certain characteristics or concepts in papers or studies, and in the mapping, reporting or discussion of these characteristics/concepts. In these cases, a scoping review is the better choice. They are particularly useful for examining emerging evidence when it is still unclear what other, more specific questions can be addressed through a systematic review. Ultimately, the choice between systematic and scoping reviews should be guided by the research question, the available evidence, and the goals of the review.

1.4. Objectives

For this reason, the main goal of this study is to perform for the first time a scoping review of regression models for ordinal responses in the sports field. Specifically, this study aims to review the literature to achieve the following objectives:

i. Describe the methods employed to analyze ordinal outcomes within the domain of sports field, to provide an overview of the current state of research in this area.
ii. Examining whether the articles published in this field adequately provide all necessary reporting information when applying regression models for ordinal responses, which will help in identifying the quality of reporting and the need for improvement.
iii. Facilitate the identification of gaps in the evidence base where no research has previously been conducted, to provide directions for future research and potentially contribute to the development of new approaches.
iv. Draw conclusions from the existing literature regarding ordinal response regression models.
v. Provide a case study that applies an ordinal model to a real-world sports example, to address a practical problem. The study will report all relevant information appropriately and demonstrate the effectiveness of the model in this context.

This study provides valuable insights by offering a comprehensive overview of the use of regression models for ordinal responses in the sports field, which has not been done before. Additionally, the study will identify the relevance of the need for a systematic review, potentially contributing to the advancement of research synthesis in this field.

2. Methods

2.1. Common Ordinal Regression Models

Ordinal modelling is an important aspect of statistical analysis, especially when dealing with responses that are inherently ordered. In the literature, there are numerous parametric ordinal models and approaches to modelling ordinal data. These models respect the ordinal nature of the data, without the assumption that the data are continuous (52). A brief review of some common regression models for ordinal response variables will be provided in this section. For a more detailed review refer to (18,19).

Before discussing specific models, as ordinal modeling techniques are based on logistic structures, it is important to introduce first the logistic function and the types of logits used for modeling ordinal variables, which include cumulative logits, continuation ratio logits, and adjacent-categories logits.

Logistic Function

Let Y denote a binary response variable that can assume only two values, say 0 and 1, for simplicity. Denote the expected value of Y by $E(Y) = P(Y = 1) = \pi$ and suppose that we want to model the way $\pi = \pi(\boldsymbol{x})$ depends on the values of explanatory variables $\boldsymbol{x} = (x_1, \dots, x_k)$. For ordinal variables is often more appropriate to use a model that allows a curvilinear relationship between $E(Y)$ and each quantitative x. For one predictor x, one function that has this appearance is,

$$\pi(x) = \frac{exp(\alpha + \beta \boldsymbol{x})}{[1 + exp(\alpha + \beta \boldsymbol{x})]} \quad (1)$$

- α represents the intercept term or log-odds of the outcome variable when $x = 0$.
- β represents the coefficient of the predictor variable x.

called the *logistic function.* Greater intuition can be obtained regarding this curve by noting that when $\beta > 0$, which is the distribution function of the logistic random variable having mean $-\alpha/\beta$ and standard deviation $\pi/\sqrt{3\beta}$. Then, the odds of making response 1 instead of response 0 is as follows:

$$\frac{\pi(\boldsymbol{x})}{\big(1 - \pi(\boldsymbol{x})\big)} = exp(\alpha + \boldsymbol{\beta}'\boldsymbol{x})$$

The odds increase multiplicatively by e^{β} for every unit increase in x. The log odds have the simple linear relationship.

$$log\left[\frac{\pi(\boldsymbol{x})}{\big(1 - \pi(\boldsymbol{x})\big)}\right] = \alpha + \boldsymbol{\beta}'\boldsymbol{x} \quad (2)$$

The log odds transformation is referred to as the *logit*, and a model for the log odds is called a logit (or logistic) regression model.

Logits for an Ordinal Response

If there are c categories in an ordinal response variable with $c \geq 2$, there are many ways of forming logits. For example, for a polytomous variable with c response probabilities $(\pi_1, \dots, \pi_c)$ at a certain combination of levels of explanatory variables, we could form the conditional logit

$$log\left[\frac{\left(\pi_j/(\pi_j+\pi_k)\right)}{\left(\pi_k/(\pi_j+\pi_k)\right)}\right] = log\left(\frac{\pi_j}{\pi_k}\right)$$

(3)

This is the log odds of classification in category j instead of category k, given that an observation falls in one of those two reference categories. All logits of this type can be derived from a basic set of size $c-1$. For example, if

$$L_j = log(\pi_j/\pi_c), \qquad j = 1, \dots, c-1$$

(4)

when the response variable is ordinal it makes sense to form logits in a way that takes the category order into account. In so doing, it is not necessary to use only two categories at a time in forming the logits. The logits can be formed by grouping categories that are contiguous on the ordinal scale, for example. Among the logit types that utilize category ordering are the "*cumulative"* or "*accumulate"'* logits,

$$L_j = log\left[\frac{(\pi_{j+1}+\cdots+\pi_c)}{(\pi_1+\cdots+\pi_c)}\right], \qquad j = 1, \dots, c-1$$

(5)

The "*continuation-ratio*" logits,

$$L_j = log\left[\frac{\pi_{j+1}}{\pi_1+\cdots+\pi_j}\right], \qquad j = 1, \dots, c-1$$

(6)

And the "*adjacent categories*" logits,

$$L_j = log\left[\frac{\pi_{j+1}}{\pi_j}\right], \qquad j = 1, \dots, c-1$$

(7)

The proportional odds assumption

In order to effectively analyze ordinal data in a statistical context, it is crucial to address the *proportional odds assumption (see* **Figure 12***)* for Ordinal Logistic Regression models and its significance, since this assumption must be checked before stating that the model is valid (all other assumptions passed). This assumption states that the odds ratios for an independent variable are constant across different levels of the ordinal dependent variable. One effective way to test this assumption is through the *Brant Test* (53), which assesses whether the observed deviations from the Ordinal Logistic Regression model are larger than what could be attributed to chance alone.

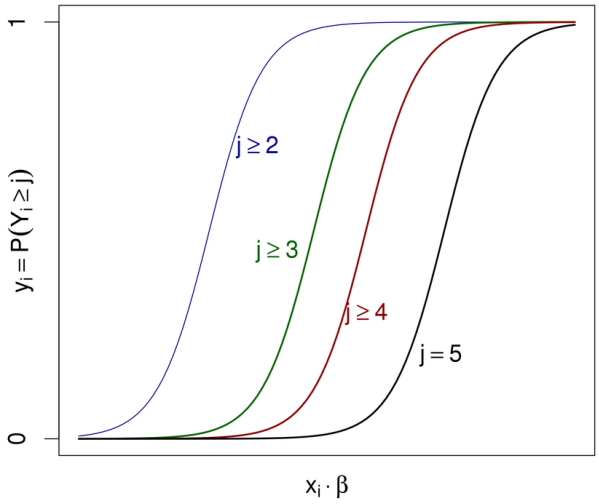


***Figure 12**: Illustration of the proportional odds assumption for 5 response categories. Source: (54)*

The *Brant Test* compares the fit of the proportional odds model to a model that allows for different coefficients across the categories of the dependent variable. Involves estimating separate models for each independent variable and produces a *p-value* indicating whether the proportional odds assumption can be rejected. If the *p-value* is less than a predetermined significance level, typically 0.05, the *proportional odds assumption* is violated, and the model may need to be adjusted.

To use the Brant Test in R, the package "*brant*" (Benjamin Schlegel and Marco Steenbergen (2020). brant: Test for Parallel Regression Assumption. R package version 0.3-0. https://CRAN.R-project.org/package=brant) can be downloaded from CRAN and used in conjunction with the *polr*() function from the *MASS* package.

Now that an introduction to the basics of ordinal modelling has been provided, including the logistic function, the types of logits used for ordinal variables and the proportional odds assumption, we can move on to a more detailed exploration of specific models and their corresponding equations. In the following sections, the models that will be described include the Proportional Odds Model Version of the Cumulative Logit Model, Mixed-Effects Proportional Odds Model, Continuation ratio Logit model, Adjacent-Categories Logit Model and Generalized Partial Credit Model. By examining deeper into each model, this study aims to provide a more comprehensive understanding of how to effectively analyze ordinal data in a statistical context.

We would like to remark that the ordinal responses Y are labeled $1, 2, \ldots, c$ throughout the following sections.

2.1.1. Proportional Odds Model Version of the Cumulative Logit Model

The proportional odds model (POM), also called cumulative odds, is a commonly used model for the analysis of ordinal categorical data and comes from the class of generalized linear models. It is a generalization of a binary logistic regression model when the response variable has more than two ordinal categories. The proportional odds model is a model for an ordinal response variable, Y that can fall in $j = 1, \ldots, J$ categories. It is used to estimate the odds of being at or below a particular category of the response variable. For example, if there are j levels of ordinal outcomes, the model

makes J − 1 predictions, each estimating the cumulative probabilities at or below the jth level of the outcome variable. Then Y follows a multinomial distribution with parameter π that denote the probability that the observation falls in response category j. We define the cumulative probabilities as:

$$P(Y \leq j) = \pi_1 + \cdots + \pi_j \tag{8}$$

In order to describe this model, we start by using the logit function, which we have previously defined as $\text{logit}(\pi) = \log\left[\frac{\pi}{1-\pi}\right]$, additionally the cumulative logits are defined as:

$$logit\big(P(Y \leq j)\big) = log\left(\frac{P(Y \leq j)}{1 - P(Y \leq j)}\right), \qquad j = 1, \dots, c-1 \tag{9}$$

so that the cumulative logits are defined for all but the last category. A cumulative logit model is a regression model for cumulative logits:

$$logit(P(Y \leq j) = \alpha_j - \boldsymbol{\beta}'\boldsymbol{x}, \qquad j = 1, \dots, c-1 \tag{10}$$

Where:

- $\boldsymbol{x}$ is a vector of explanatory variables for the observation.
- $\boldsymbol{\beta}$ is the corresponding set of regression parameters, which describes the effect of x on the log odds of the response variable.
- The $\boldsymbol{\alpha_j}$ parameters provide each cumulative logit (for each j) with its own intercept, called its *cut point*. These parameters are generally of little interest, and we might usually consider them as nuisance parameters.

A key point is that the regression part $\boldsymbol{\beta}'\boldsymbol{x}$ is independent of j, so $\boldsymbol{\beta}$ has the same effect for each of the $c-1$ cumulative logits. This independence is the proportional odds assumption. Therefore, all the cumulative logits contain the same effect of x. This has the advantage that we only have to fit one parameter instead of $c-1$ parameters. In this manner, this model assumes that all the effects β over a set of predictor variables x on the response variable Y for the defined log odds are the same regardless of any collapsing of the j-level response variable to a binary variable.

The cumulative logit model is illustrated in **Figure 12**, for data with five response categories. For small values of $\boldsymbol{\beta}'\boldsymbol{x}$ the response is likely to fall in the first category and for large values of $\boldsymbol{\beta}'\boldsymbol{x}$ the response is likely to fall in the last category. The horizontal displacements of the curves are given by the values of α_j (55). The degree of association of the response Y with the predictor variables $\boldsymbol{x}$ is determined by the value of the effects $\boldsymbol{\beta}'\boldsymbol{x}$. When the model fitted shows that $\beta = 0$, the value of $P(Y \leq j)$ as a function of x is a constant for each j and it means that Y and $\boldsymbol{x}$ are statistically independent. The negative sign preceding the effect $\boldsymbol{\beta}'\boldsymbol{x}$ in expression (10) allows a parametrization which has a natural interpretation of the effect $\boldsymbol{\beta}'\boldsymbol{x}$ regarding whether it is positive or negative. Therefore, if $\boldsymbol{\beta}'\boldsymbol{x} > 0$ then higher values of x lead to higher values of Y. This interpretation is reversed when $\boldsymbol{\beta}'\boldsymbol{x} < 0$. Meaning, the larger the value of $\boldsymbol{\beta}'\boldsymbol{x}$, the higher the probability of the response falling in a category at the upper end of the response scale. Note that $\boldsymbol{\beta}'\boldsymbol{x}$ does not contain an intercept, since the α_j act as intercepts.

Mixed-Effects Proportional Odds Model

In cases when ordinal data accommodate multiple random effects, Hedeker (56) described a ***mixed-effects proportional odds model*** (Mixed Effects POM). The mixed-effects proportional odds model is specified in the form:

$$log\left[\frac{\pi}{1-\pi}\right] = \alpha_j - \boldsymbol{\beta}'\boldsymbol{x} + \mu_{0i}, \qquad j = 1, \dots, c-1$$

(11)

The random subjects effects μ_{0i} are assumed to be normally distributed in the population of subjects with 0 mean and variance σ_u^2.
α_j are the model thresholds with $c-1$ increasing order. The marginal response probabilities in the j categories can be expressed with thresholds parameters and model intercepts. The effects of the regression coefficients are conditional on the level of the random effect μ_{0i}.

2.1.2. Continuation-Ratio Logit Model

The Continuation-Ratio Logit Model is another popular model used in ordinal regression. It assumes that the probability of being in a particular category is a function of the previous category. It has two different settings each focusing on a distinct aspect of the response variable. The first one is based on the log odds of each category relative to the preceding categories,

$$log(P[Y = j|Y \leq j]) = log\left(\frac{\pi_j}{\pi_1 + \cdots + \pi_j}\right), \qquad j = 1, \dots, c-1$$

(12)

while the second one is based on the log odds of transitioning to higher categories in the response variables,

$$log(P[Y = j|Y \geq j]) = log\left(\frac{\pi_j}{\pi_j + \cdots + \pi_c}\right), \qquad j = 1, \dots, c-1$$

(13)

This second setting of log odds related to higher categories is commonly used when analyzing data with ordered categories that represent states with a defined order, where the subject transitions through the different categories before reaching the final response outcome. This process is known as a *sequential process*, and it is frequently used in medical research to determine the recovery time of subjects after cancer treatment. For instance, the ordered categories in recovery time may include less than 1 year, 1 to 2 years, 2 to 3 years, 3 to 5 years, and more than 5 years. The *continuation-ratio logit model for a sequential process* (57) can be represented as follows:

$$logit(P(Y \geq j) = \alpha_j - \boldsymbol{\beta}'\boldsymbol{x}, \qquad j = 1, \dots, c-1$$

(14)

It is possible to describe this model in its partial proportional odds structure (58).

The alternative logits

$$L_j' = log\left[\frac{\pi_j}{\pi_{j+1} + \cdots + \pi_j}\right], \qquad j = 1, \dots, c-1$$

(15)

Can be interpreted as log odds for "*hazard*" probabilities, since L_j' uses the ratio of the conditional probabilities $\pi_j|(\pi_j + \cdots + \pi_c)$ and $(\pi_{j+1} + \cdots + \pi_c)|(\pi_j + \cdots + \pi_c)$. In fact, Thompson (1977)

proposed a model having these logits for the analysis of discrete survival-time data. He showed that when the lengths of the time intervals approach zero, his model converges to Cox's (1972) *proportional hazards model* for survival data.

2.1.3. Adjacent-Categories Logit Model

Adjacent category models are another set of models used for ordinal data, which do not assume proportionality of the odds ratios. These models have a separate set of regression coefficients for each adjacent pair of categories and allow for more flexibility in the relationship between the response variable and explanatory variables. The *adjacent category logit model* can be represented as follows:

$$log\left(\frac{\pi_j(x)}{\pi_{j+1}(x)}\right) = \alpha_j - \boldsymbol{\beta_j}'\boldsymbol{x}, \qquad j = 1, \dots, c-1$$

(16)

As mentioned, it considers the probability of two adjacent-categories rather than considering the probability of each category versus a baseline-category. For this reason, the effects δ_j are described with local log odds instead of the cumulative log odds.

2.1.4. Generalized Partial Credit Model

When responses to a test item are classified into two categories, dichotomous item response models can be applied. When responses to an item have more than two categories, a polytomous item response model is appropriate for the analysis of the responses. Then, if the options on a rating scale are successively ordered, the Graded Response model by Samejima (59) and its rating scale version by Muraki (60) or the Partial Credit model by Masters (61) are applicable. Therefore, the Generalized Partial Credit Model (GPCM) is a type of Item Response Theory (IRT) model that is used to estimate the latent trait of individuals based on their responses to items, useful for analyzing polytomous data, where the response categories are more than two.

For this section the responses are scored $0,1, \dots, k$ where k is the highest score category for the item. The GPCM is formulated based on the assumption that the probability of choosing the k^{th} category over the k^{1st} category is governed by the dichotomous response model. In Muraki's formulation (60), the probability of choosing category k over category $k-1$ is given by the conditional probability:

$$\pi_{jk} = exp\left[\boldsymbol{\alpha_j}\left(\theta - \boldsymbol{b_{jk}}\right)\right]\pi_{j,k-1}, \qquad k = 1,2, \dots, m_j$$

After normalizing each π_{jk}, so that $\sum \pi_{jk} = 1$, the GPCM is written as

$$\pi_j = \frac{exp\left[\sum_{v=0}^{k} D\boldsymbol{\alpha_j}\left(\theta - \boldsymbol{b_j} + \boldsymbol{d_{j,v}}\right)\right]}{\sum_{v=0}^{c} exp\left[\sum_{v=0}^{c} D\boldsymbol{\alpha_j}(\theta - \boldsymbol{b_j} + \boldsymbol{d_{j,v}}\right]}$$

where D, is a scaling constant set to 1.7 to approximate the normal ogive model, α_j is a slope parameter, b_j is an item location parameter, and $d_{j,v}$ is a category parameter. The slope parameter indicates the degree to which categorical responses vary among items as θ level changes. With m_j categories, only $m_j - 1$ category parameters can be identified. Indeterminacies in the parameters of the GPCM are resolved by setting dj,0 = 0 and setting $\sum_{k=1}^{m_j-1} d_{jk} = 0$.

2.2. Choosing and Fitting Ordinal Models

Table 2 provides a comprehensive summary of common ordinal regression models. This information is presented to help readers quickly identify the appropriate model for their specific needs. The table provides details on five commonly used models: Proportional Odds Model, Mixed-Effects Proportional Odds Model, Continuation-Ratio Model, Adjacent-Categories Model and Generalized Partial Credit Model. Each model is described in terms of its advantages and disadvantages, along with the recommended use case for each model, highlighting factors such as the flexibility it offers, complexity of interpretation, and the nature of data it is best suited for.
Additionally, for users who prefer to work in R, **Table 3** provides a list of R packages that can be used to fit these models, making it a valuable resource for those seeking to apply ordinal regression analysis in their work.

***Table 2**: Summary table of common ordinal regression models, including their advantages, disadvantages, and recommended use cases.*

Common Ordinal Regression Models			
Type of Model	***Advantages***	***Disadvantages***	***When to use?***
Proportional Odds Model	- Can handle many outcomes. - Flexible with variable scales. - Faster to estimate. - Simple and easy to interpret. - Less computationally intensive than other models.	- Assumes a linear relationship between logit and Y. - Requires the proportional odds assumption to hold. - Does not account for any correlation that may exist between categories. - Can be unstable with small sample sizes.	- When the odds ratios of each predictor are proportional - When the response variable can be interpreted as a continuous latent variable that is being categorized into discrete categories. - When unequal variances may exist. *Examples*: Likert-item scale, Health status, and patient-reported outcomes.
Mixed-Effects Proportional Odds Model	- Can simultaneously model parameters and the variability between and within subjects - Allows for the analysis of correlated observations within subjects. - Can accommodate sparse sampling	- Assumes that the odds ratios of each predictor are proportional across the different response categories. - Can be more complex to specify, estimate and interpret. - Can be computationally intensive.	- When the data exhibits hierarchical or clustered structure, such as individuals nested within groups or repeated measurements on the same subjects. - When the dataset has missing values or unbalanced designs. *Examples:* longitudinal study investigating the progression of a disease over time.
Continuation-Ratio Model	- Easy to understand. - Can handle large datasets. - Allows for different effects for each adjacent level.	- Assumes variable scales are proportional. - Requires estimation of multiple models, one for each category.	- When the ordered outcomes are related. - When category-specific effects are of interest.

	- Account for correlations that may exist between categories.	- Assumes a common intercept across all categories. - Can become computationally intensive when the number of categories is large.	*Examples*: recurrence and drug addiction studies.
Adjacent-Categories Model	- Creates more fine-grained distinctions between categories. - Allows the modeling of different effects for adjacent categories. - Robust when assumptions are violated. - Account for correlations that may exist between categories.	- Valid only for the adjacent category and not for mixed designs. - Assumes a common intercept across all categories. - Ignores the possibility of effects that may operate across multiple categories.	- When neighboring categories have more closely related meanings. - When category-specific effects are of interest, and it is important to account for correlations between adjacent categories. *Examples*: evaluating academic performance.
Generalized Partial Credit Model	- Can handle asymmetrical effects and missing data. - Accounts for complex relationships. - Minimizes loss of information by avoiding category collapse.	- May be more difficult to interpret, requiring specialized knowledge in statistical analysis. - Can be more complex to implement.	- When category-specific effects, estimate item parameters, and measure latent traits are of interest. - When analyzing data obtained from surveys or opinion polls that employ Likert or rating scales. *Examples*: assessing employees' proficiency in performing specific tasks in the workplace.

Table 3 presents an implementation of ordinal models using different packages in the R statistical computing environment. The rationale for using R software is that it is free and open source, making it accessible to everyone without any licensing fees. Additionally, many computational and statistical procedures are implemented in R before they are available in other commercial software packages (52). However, users may prefer to implement ordinal models in their preferred statistical software. Relevant resources for implementing ordinal models include Heck et al. (62) for IBM SPSS; Bender et al. (63) for SAS and S-Plus; and Long et al. (64) for Stata.

Table 3: Summary table providing a list of R packages that can be used to fit ordinal models.

R packages for fitting Ordinal models	*Citation* [1]
Package 'ordinal': ◦ *clm():*Fit cumulative link models. New and improved implementations in *clm2()* ◦ *clmm():*Fit cumulative link mixed models. New and improved implementations in *clmm2()*	Christensen RHB (2022). "ordinal—Regression Models for Ordinal Data ." R package version 2022.11-16. https://CRAN.R-project.org/package=ordinal.
Package 'brms': ◦ *brm():* Fit Bayesian Generalized (Non-)Linear Multivariate Multilevel Models.	Bürkner P (2017). "brms: An R Package for Bayesian Multilevel Models Using Stan." *Journal of Statistical Software*, 80(1), 1–28. doi: 10.18637/jss.v080.i01.
Package 'rms': ◦ *lrm():* Fit binary and proportional odds ordinal logistic regression models using maximum likelihood estimation or penalized maximum likelihood estimation. ◦ *orm():* Fits ordinal cumulative probability models for continuous or ordinal response variables, efficiently allowing for a large number of intercepts. Five different distribution functions are implemented, with the default being the logistic (i.e., the proportional odds model).	Frank E Harrell Jr (2022). rms: Regression Modeling Strategies. R package version 6.3-0. https://CRAN.R-project.org/package=rms
Package 'MASS': ◦ *polr():* Fit a logistic or probit regression model to an ordered factor response. The default logistic case is proportional odds logistic regression, after which the function is named.	Venables, W. N. & Ripley, B. D. (2002) Modern Applied Statistics with S. Fourth Edition. Springer, New York. ISBN 0-387-95457-0
Package 'VGAM': ◦ *cumulative():* Fits a cumulative link regression model to a ordered factor response. ◦ *acat():* Fits an adjacent categories regression model to an ordered factor response. ◦ *propodds():* Fits the proportional odds model to an ordered factor response. ◦ *cratio():* Fits a continuation ratio logit/probit/cloglog/cauchit/... regression model to an ordered factor response.	Thomas W. Yee (2023). VGAM: Vector Generalized Linear and Additive Models. R package version 1.1-8. URL https://CRAN.R-project.org/package=VGAM
Package 'MXM': ◦ *ordinal.reg():* Generalised ordinal regression.	Lagani V, Athineou G, Farcomeni A, Tsagris M, Tsamardinos I (2017). "Feature Selection with the R Package MXM: Discovering Statistically Equivalent Feature Subsets." _Journal of Statistical Software_, *80*(7). doi: 10.18637/jss.v080.i07 (URL: https://doi.org/10.18637/jss.v080.i07).
Package 'ordinalNet':	Wurm MJ, Rathouz PJ, Hanlon BM (2021). "Regularized Ordinal Regression and the ordinalNet R Package." _Journal of Statistical Software_, *99*(6), 1-42. doi: 10.18637/jss.v

[1] Packages are cited in accordance with the APA procedure.

◦ *ordinalNet():*Fits ordinal regression models with elastic net penalty by coordinate descent. Supported model families include cumulative probability, stopping ratio, continuation ratio, and adjacent category.	099.i06 (URL: https://doi.org/10.18637/jss.v099.i06).
Package 'mvord': ◦ *mvord():* Fits multivariate ordinal regression models.	Hirk R, Hornik K, Vana L (2020). "mvord: An R Package for Fitting Multivariate Ordinal Regression Models." _Journal of Statistical Software_, *93*(4), 1-41. doi: 10.18637/jss.v093.i04 (URL: https://doi.org/10.18637/jss.v093.i04).
Package 'lordif': ◦ *lordif():* Performs iterative hybrid ordinal logistic regression/IRT DIF. ◦ *runorl():*Runs ordinal logistic regression models and produces DIF statistics and effect size measures	Seung W. Choi, with contributions from Laura E. Gibbons and Paul K. Crane (2016). lordif: Logistic Ordinal Regression Differential Item Functioning using IRT. R package version 0.3-3. https://CRAN.R-project.org/package=lordif
Package 'gnlm': ◦ *nordr(): F*its arbitrary nonlinear regression functions (with logistic link) to ordinal response data by proportional odds, continuation ratio, or adjacent categories. ◦ *ordglm(): F*its linear regression functions with logistic or probit link to ordinal response data by proportional odds.	Bruce Swihart and Jim Lindsey (2019). gnlm: Generalized Nonlinear Regression Models. R package version 1.1.1. https://CRAN.R-project.org/package=gnlm
Package 'lme4': ◦ *glmer():*Fit mixed-effects models for ordinal response variables.	Douglas Bates, Martin Maechler, Ben Bolker, Steve Walker (2015). Fitting Linear Mixed-Effects Models Using lme4. Journal of Statistical Software, 67(1), 1-48. doi:10.18637/jss.v067.i01.
Package 'GLMMadaptive': ◦ *mixed_model():* Fits generalized linear mixed effects models under maximum likelihood using adaptive Gaussian quadrature.	Dimitris Rizopoulos (2023). GLMMadaptive: Generalized Linear Mixed Models using Adaptive Gaussian Quadrature. R package version 0.8-8. https://CRAN.R-project.org/package=GLMMadaptive
Package 'psychotools': ◦ *gpcmodel():* Fitting function for generalized partial credit models providing a wrapper around *mirt* and *multipleGroup* relying on marginal maximum likelihood (MML) estimation via the standard EM algorithm.	Zeileis A, Strobl C, Wickelmaier F, Komboz B, Kopf J, Schneider L, Debelak R (2022). _psychotools: Infrastructure for Psychometric Modeling_. R package version 0.7-2, <URL: https://CRAN.R-project.org/package=psychotools>.
Package 'jrt': ◦ *jrt():* Fit ordinal IRT models on judgment data and return factor scores and statistics.	Myszkowski, N. (2021). Development of the R library jrt: Automated item response theory procedures for judgment data and their application with the consensual assessment technique. Psychology of Aesthetics, Creativity and the Arts, 15(3), 426-438.

3. Methodological Procedures in the Scoping Review

3.1. Criteria for the Scoping Review

3.1.1. Design

A scoping review has been carried out on the application of ordinal response regression models in the sports field, following the guidelines of the PRISMA Extension for Scoping Reviews (PRISMA-ScR) statement (65).

3.1.2. Sources of Information and Strategy

The scoping review search was performed in different databases such as PubMed, Web of Science (WoS), and specialized journals of sports statistics (not included in previous databases) such as Journal of Quantitative Analysis in Sports (JQAS) and Journal of Sports Analytics (JSA). In all databases, a search was conducted using the same Boolean terms and key words: ("ordinal regression" OR "ordinal models" OR "ordinal forest" OR "ordered response" OR "ordered categories" OR "ordinal" OR "ordered category" OR "proportional odds model" OR "stereotype model" OR "ordinal variable" OR "ordinal response") AND (“sport*”). The last search was performed on March 4, 2023.

3.1.3. Eligibility Criteria and Selection of Studies

To determine the eligibility of the studies, the following inclusion and exclusion criteria were defined:

- Inclusion Criteria

The articles included in the scoping review had to be original articles, written in English, and applying ordinal response regression models. In addition, these studies had to be within the sports field.

- Exclusion Criteria

Conference proceedings, theses, dissertations, and non-original articles (such as special issues, opinions, reviews, etc.) were excluded from the study selection process. Moreover, all articles written in a language other than English and those not related to the sports field were also excluded. It is important to note that activities denoted as *physical activity*, *exercises*, and *recreational sports* were not considered as part of the sports field in this scoping review.

The study selection process consisted of two stages. Firstly, studies were selected based on their title and abstract. Subsequently, the full texts of the articles were reviewed to evaluate whether they met the eligibility criteria in situations where a decision could not be made based on the information provided by the title and abstract.

3.1.4. Identification of the Studies

Figure 13 shows the PRISMA-SCR flowchart to summarize all the stages of the selection process. In the first phase, a total of 1.147 articles were collected using the Boolean terms and keywords mentioned above. These articles belonged to different databases: PubMed (n = 346), Web of Science (n = 808), Journal of Quantitative Analysis in Sports (n = 44). No articles were found in the Journal of Sports Analytics. Following the removal of 310 duplicates, 888 articles remained for further evaluation.

During the first phase of the review, 822 articles were excluded after reading their title and abstract. Specifically, 6 articles written in a language other than English, 9 non-original articles, 753 articles not related to the sports field, and 54 articles not related to ordinal response models were excluded. This left 66 articles for the second phase of the review, where the full texts were reviewed to determine if they used any ordinal response regression models in their studies. Out of these 66 articles, it was found that 31 articles used the word 'ordinal' only to refer to the method used for analysis or comparison without using an actual ordinal regression model. Additionally, 1 article written in German that could not be filtered in the first phase was excluded. Ultimately, a total of 34 articles met the established inclusion and exclusion criteria and were included for a more detailed review. **Figure 13** provides a summary of the number of identified articles and the reasons for exclusion at each phase of the review process.

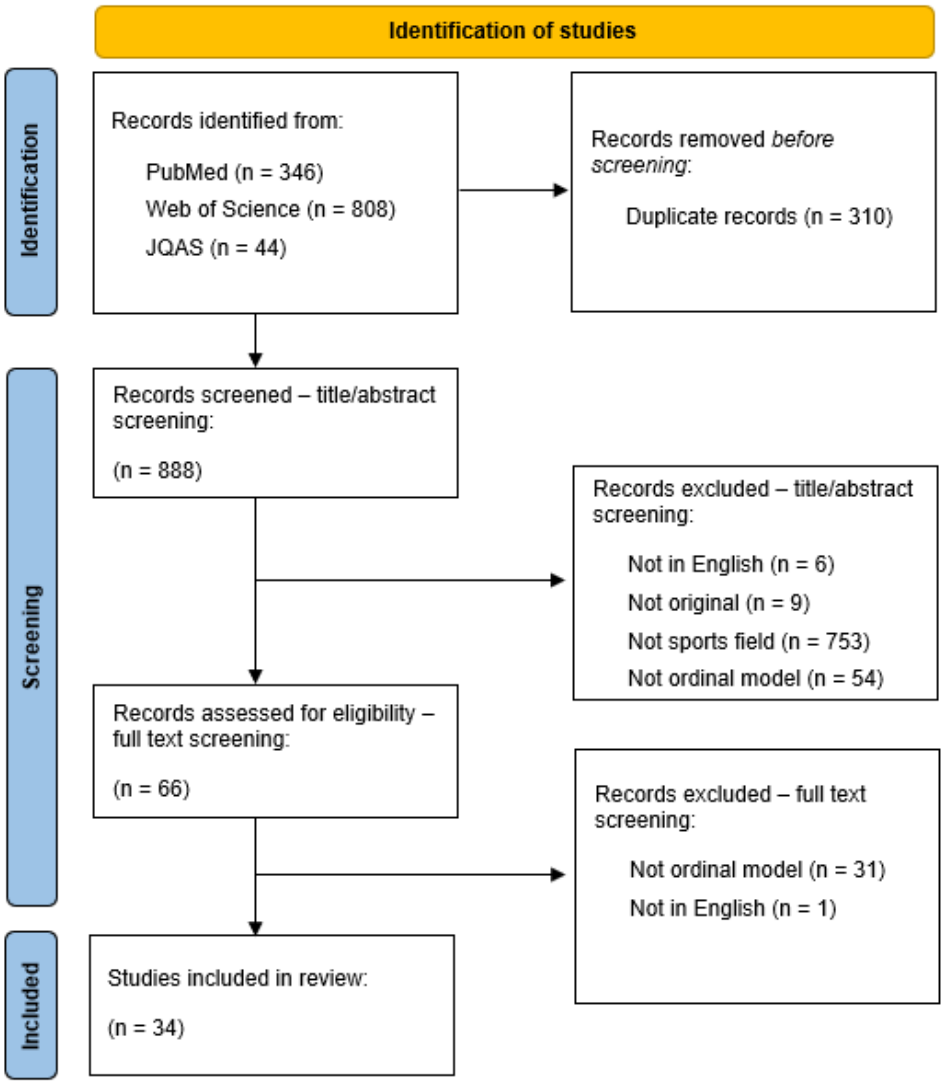


***Figure 13**: PRISMA-SCR: Flow diagram of the scoping review of the use of ordinal regression models in original articles in the field of sports.*

3.1.5. Data Extraction

The information collected in the selected studies were grouped into three categories: general characteristics of the articles, general characteristics of the sport itself, and general characteristics of the ordinal regression models. The general characteristics of the articles collected the author, the country of origin of the data, the year of publication, the name and type of journal where the article was published, the number of participants and their age, and the purpose of the study (**Table 4**). The general characteristics of the sport provided variables such as the type of sport focused on in the article, the gender of the participants (men, women, or both), the category of participants (professional, amateur, or both), the source of the data, and the field of the article within sports science (Sports performance analysis, Sports technology, Movement Integration and Health) (**Table 5**). For the last category, the general characteristics of the ordinal regression models, provided variables such as the type of ordinal model, the type of paradigm (Bayesian, or frequentist), the validation used, the type of objective of the model (Descriptive, predictive, or causal), the software and package used, and whether the article shares the code or data used (**Table 6**).

The data was collected and stored in a database. The data was then checked for discrepancies among the authors. Discrepancies were resolved by consensus after reviewing the conflicting articles again. Specifically, 11 articles where discussed, all of them excluded for not being ordinal regression models.

***Table 4**: Descriptive table of the general characteristics for the selected articles*

Variable	Description
General characteristics and design of the selected articles	
Author (citation)	APA format citation of the authors
Country	Country of the data of the article
Publication Year	Year of publication of the article
Journal Name	Name of the Journal where is published the article
Type of Journal	Type of the Journal topics (*Sport, Statistics, Medicine, Multidisciplinary, …)*
Number (participants)	Number of participants of the study
Age (participants)	Average or age interval of the participants
Type of design	Type of the study design (*Cross-sectional, Longitudinal, Mixed*)
Principal Aim	Description of the principal aim of the study

***Table 5**: Descriptive table of the general characteristics of the sport*

General characteristics of the sport	
Author (citation)	APA format citation of the authors
Type of Sport	Sport or sports in which the article focuses
Gender	Sex of the participants in which the article focuses (*Male, Female, Both*)
Category participants	Category of the participants (*Professional, Amateur, Both*)
Name of source data (League, Association, Organization, Federation ...)	Name of the source of the data of the study
Category classification	Four categories: *1) Sports Performance Analysis; 2) Health; 3) Sports Technology; 4) Movement Integration*

***Table 6**: Descriptive table of the general characteristics of the ordinal regression models*

General characteristics of the Ordinal regression models	
Author (citation)	APA format citation of the authors
Type of model	Type of Ordinal regression model used in the article (*Proportional Odds Model, Adjacent-Category Model, Continuation-ratio Model, Stereotype Model, Mixed-effect ordinal model, ...*)
Type of paradigm	The overall approach or theoretical framework used in the analysis (F*requentist, Bayesian*)
Validation used	Method used to assess the performance of the model (*cross-validation, bootstrap, hold-out validation, ...*)
Type of objective of model	Research question or objective that the model aims to answer (*Descriptive, Predictive, Causal*)
Statistical software	Software package used for the statistical analysis (*R, SAS, STATA, SPSS, ...*)
Statistical package, function or macro:	Specific function or macro used within a statistical package to perform a specific analysis
Data shared	Whether the article shares the data used in the study for other researchers to access and use
Code shared	Whether the article shares the code used in the study for other researchers to access and use
Repository of Data or Code shared	Whether the article shares a repository or database for the data or code used in the study

.

3.1.6. Statistical Analysis

The statistical analysis was conducted using R version 4.1.3 (2022-03-10) R Core Team (2022). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/. The *ggplot2* package was used for graphical representation.

The reproducible code used in this study is available on a publicly accessible GitHub repository (https://github.com/Mireiaestopanan/TFG_R.code.git), allowing for the transparency and replicability of the statistical analysis. This ensures that other researchers can reproduce the findings of the study and build upon it for further research.

4. Results

4.1. Descriptive Analysis

This section presents the results of the descriptive analysis of the variables described in the previous section's tables regarding the different characteristics of the 34 articles included in the scoping review. First, a frequency table is provided with the main results for each characteristic.

It should be noted that for some variables, the frequency of possible values is greater than the number of selected articles (n = 34). This occurs because some variables have multiple responses, given that certain articles report more than one type of value for each record. To clarify these situations, each case is detailed below.

One of the 34 included articles applied an ordinal response regression model to two distinct sets of data, with different countries and sports analyzed for each of the two data sets. Additionally, one article used more than one country to conduct the study. For this case, the sum of frequencies for the "*Country*" variable in the **Table 7** is (n = 36), and the sum of frequencies for the "*Sport*" variable in the **Table 8** is (n = 35), which exceeds the number of articles (n = 34).
Table 9, for the "*Validation*" and "*Software used*" variables, there are articles that use more than one type of validation method or different software, resulting in multiple responses. Therefore, the total sum of the number of validation methods (n = 43) and software (n = 36) also exceeds the number of articles (n = 34).

In addition, it is important to note that the values of each variable in the tables are sorted by frequency, from highest to lowest, except for the "*Publication Year*" variable, which is sorted from the earliest to the latest year. However, any values reported as "others" or "not reported" are always placed at the end of the list, regardless of their frequency. This is done to ensure consistency in the presentation of the results and facilitate the interpretation of the tables.

***Table 7**: Frequency distribution of general characteristics of the selected articles by group.*

Variable (N = 34*)	Category	n	%
General characteristics and design of the selected articles			
Country *(N = 36*)*	United States	7	19.4%
	Australia	5	13.9%
	Italy	3	8.3%
	Brazil	2	5.6%
	Germany	2	5.6%
	Others [2]	9	25.0%
	Not Reported	8	22.2%
Publication Year	2000	1	2.9%
	2005	1	2.9%
	2006	1	2.9%
	2008	1	2.9%
	2012	1	2.9%
	2014	1	2.9%
	2015	1	2.9%

[2] The variable Country, "Others" reports the following countries with a frequency of 1: Belgium; Finland; France; Greece; Netherland; New Zealand; Spain; Switzerland; 44 countries; England.

	2016	2	5.9%
	2017	5	14.7%
	2018	2	5.9%
	2019	2	5.9%
	2021	7	20.6%
	2022	4	11.8%
	2023	5	14.7%
Type of Design	Cross-Sectional	7	20.6%
	Longitudinal	5	14.7%
	Not Reported	22	64.7%
Type of Journal	Sports	9	26.5%
	Sports Medicine	7	20.6%
	Statistics	5	14.7%
	Sports Statistics	4	11.8%
	Psychology	2	5.9%
	Medicine	2	5.9%
	Multidisciplinary	2	5.9%
	Others [3]	3	8.8%

***Table 8**: Frequency distribution of general characteristics of the sport in which focuses the articles by group.*

General characteristics of the sport			
Sports *(N = 35*)*	Football [4]	10	28.6%
	Basketball	6	17.1%
	Australian Football	3	8.6%
	Multidisciplinary	3	8.6%
	Volleyball	3	8.6%
	Cricket	2	5.7%
	Shooting	2	5.7%
	Others [5]	6	17.1%
Gender	Both	12	35.3%
	Male	11	32.4%
	Female	1	2.9%
	Not Reported	10	29.4%
Category Participants	Professional	27	79.4%
	Amateur	5	14.7%
	Both	1	2.9%
	Not Reported	1	2.9%
Category Classification	Sports Performance Analysis	26	76.5%
	Health	5	14.7%
	Movement Integration	2	5.9%
	Sports Technology	1	2.9%

[3] The variable Type of Journal, "Others" reports the following: Mathematics; Sports and Psychology.
[4] The variable Sports, "Football" refers specifically to the soccer (n=9) and futsal (n=1) type of football.
[5] The variable Sports, "Others" reports the following sports with a frequency of 1: Cycling; Muay Thai; Skeet Shooting; Swimming; Tennis.

***Table 9**: Frequency distribution of general characteristics of the ordinal regression model used in the selected articles by group.*

General characteristics of the Ordinal regression model			
Type of Model	POM	11	32.3%
	Mixed Effects POM	4	11.8%
	GPCM	1	2.9%
	Proposed Model* [6]	8	23.5%
	Not Reported	10	29.4%
Proposed Model*	No	26	76.5%
	Yes	8	23.5%
POM validation* [7]	No	8	53.3%
	Yes	7	46.7%
Validation *(N = 43*)*	Information Criterion	5	11.6%
	R-Squared	5	11.6%
	Likelihood Ratio Test	4	9.3%
	Deviance	3	7.0%
	Residual Metric	3	7.0%
	Cross-Validation	2	4.6%
	Others [8]	3	7.0%
	Not Reported	18	41.9%
Type of Paradigm	Frequentist	28	82.4%
	Bayesian	6	17.6%
Type of Objective of the Model	Descriptive	13	38.2%
	Predictive	11	32.4%
	Causal	0	0%
	Not Reported	10	29.4%
Software Used *(N = 36)*	R	14	38.8%
	SPSS	7	19.4%
	SAS	2	5.6%
	STATA	2	5.6%
	S-Plus 2000	1	2.8%
	Under request	1	2.8%
	Not Reported	9	25%
Package Used	Reported	13	38.2%
	Not Reported	21	61.8%
Data Shared	No	25	73.5%
	Yes	6	17.6%

[6] The variable Type of Model "Proposed Model*" reports whether the authors proposed modifications, additions, generalizations, or a new model for the study, or used an existing model. The proposed models are the following with a frequency of 1: Dynamic Cumulative model; Functional Ordinal Logistic Regression model; General Latent Trait model; Generalization to CUB model; Generalized Partial Credit model; Isotonic Least Squares Pairwise Comparison model; Latent Factor Distributed Lag model. And with a frequency of 2: Ordinal Logistic Regression and Expectation.

[7] The variable "POM Validation*" reports whether the POM model conducted a test to validate the assumption of proportionality.

[8] The variable Validation "Others" reports the following methods with a frequency of 1: Posterior Intervals; Posterior Predictive Cheks; Log-Loss plot.

	Under request	3	8.8%
Code Shared	No	30	88.2%
	Yes	3	8.8%
	Under request	1	3.0%
Repository of Data or Code Shared	No	30	88.2%
	Yes	4	11.8%

4.1.1. General Characteristics of the Articles

In the course of the study, 34 relevant articles were analyzed, leading to the identification of 36 countries from which data was sourced. The data sets analyzed originated from various parts of the world, including 9 countries in the American continent. Among these, the United States featured as the most frequently represented country, accounting for 19.4% of the data. A total of 13 data sets were obtained from countries located in Europe, where Italy was the most frequently represented country, featuring in 8.3% of the data. Similarly, 6 data sets originated from Oceania, with Australia emerging as the most frequently represented country, featuring in 13.9% of the data. Finally, 8 data sets could not be attributed to a specific continent of origin.

Regarding the publication year of the articles, **Table 7** reveals a significant trend of increased utilization of ordinal regression models in the field of sports in recent years. Specifically, the data indicates that there has been a yearly variation in the number of original articles published within the past 9 years, ranging from 1 to 7. Notably, the year 2021 stands out as the year with the highest number of publications, accounting for 20% of the overall number of articles. In contrast, for the other 14 years, with sporadic absences of publications, the number of articles published remained constant at one per year, representing only 2.9% of the total. **Figure 14** displays the temporal evolution of the number of publications utilizing ordinal regression models in sports field. The first publication was recorded in 2000, followed by a gap of four years until 2004 without any further publications in the field. Since then, a growing trend in the number of publications has been observed.

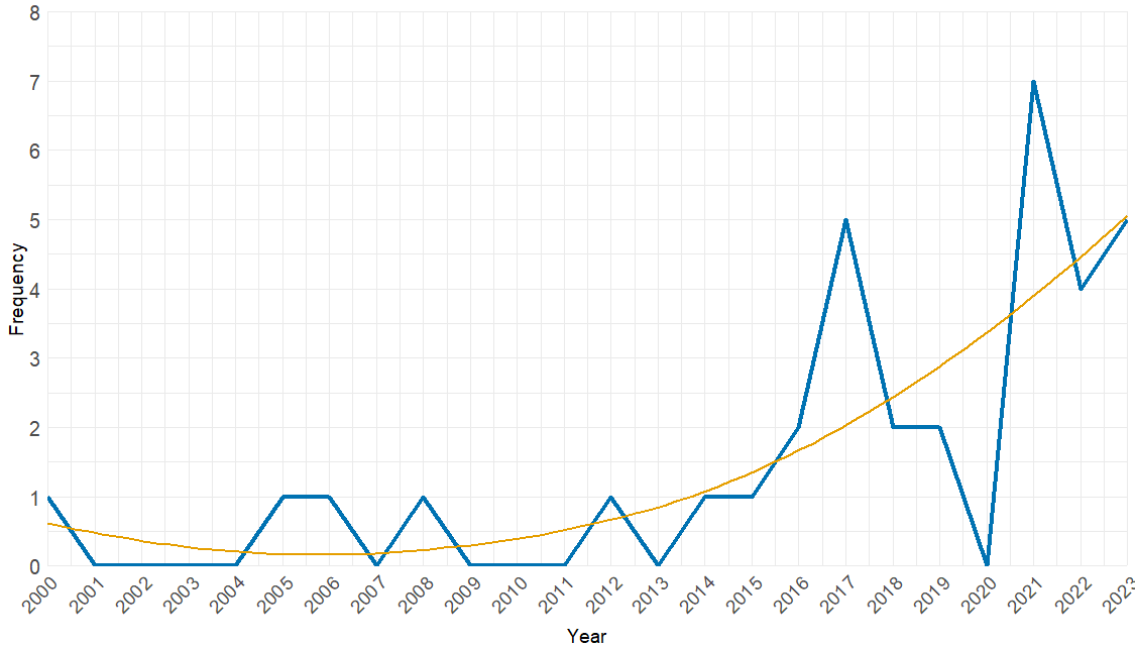


***Figure 14*: Evolution of Articles Published per Year**. *This graph presents the temporal trend of published articles in the sports field that utilize ordinal regression models. The y-axis represents the number of articles published and the x-axis shows the years. The graph provides an overview of the popularity of using ordinal regression models in the sports field over time.*

According to the data presented in the **Table 7**, the selected articles were classified into three types of study designs. Among the 34 articles, 7 (20.6%) employed a cross-sectional design, 5 (14.7%) used a longitudinal design, and 22 (64.7%) did not specify their study design.

Among the journals where the articles were published, sports journals accounted for the majority (26.5%), followed closely by sports medicine journals (20.6%). Statistics journals also had a notable presence, representing 14.7% of the total. Additionally, there were 11.8% of articles that were published in statistics and sports journals. The remaining articles were distributed among medicine, multidisciplinary, psychology, and others, each representing less than 10% of the total.

4.1.2. General Characteristics of the Sport

The analysis included data from 35 studies, each focusing on a different sport. The distribution of the number of studies per sport is presented in **Figure 15**. Of the total of 35 studies, 32 focused on a single sport. Among these, football (soccer) and basketball were the most commonly studied sports, appearing in 10 (28.6%) and 6 (17.1%) studies, respectively. 3 out of the 35 studies (8.6%) examined multiple sports, with participants from different sporting disciplines. These studies were classified as multidisciplinary. In addition, Australian football (rugby) and volleyball were the focus of 3 studies (8.6%) each and cricket and shooting accounted for 5.7% each, of the total studies. The remaining studies (17.1%) fell into the "Others" category, covering a range of sports with frequency of 1 that did not fit into the previous categories. This category included cycling, Muay Thai, rowing, skeet shooting, swimming, and tennis.

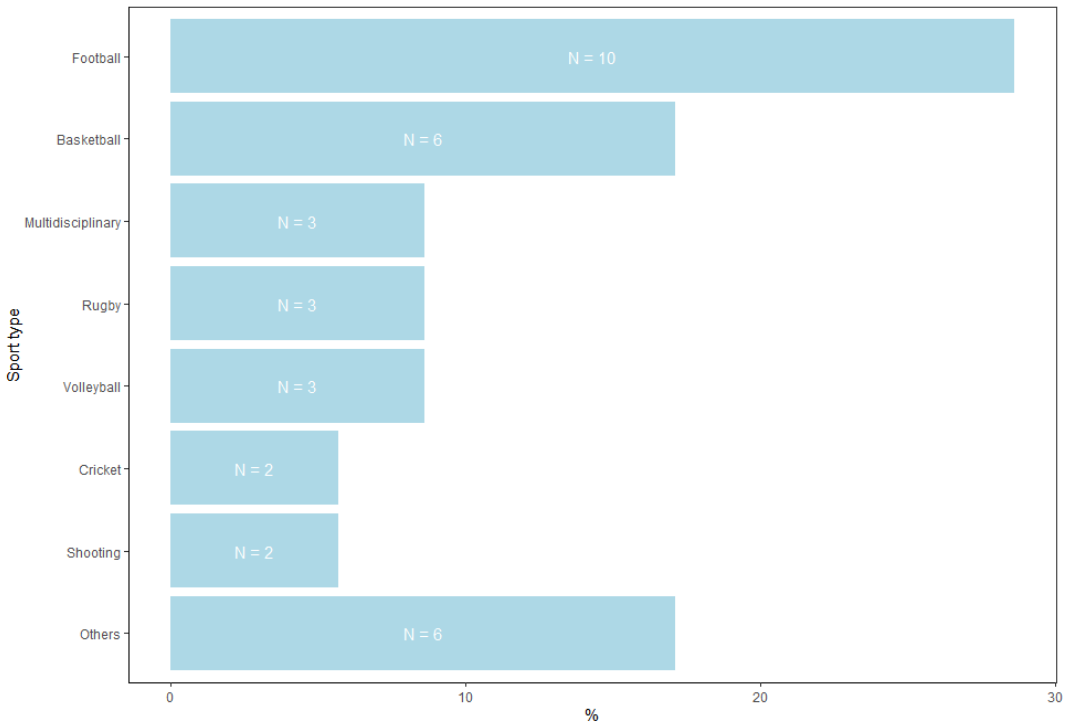


***Figure 15*: Frequency of Sport Type in which the Study Focuses**. *This bar chart displays the percentage and frequency of studies focusing on different types of sports used in the analysis. The length of each bar represents the number of studies in which the corresponding type of sport was analyzed. (Football refers to soccer and futsal, Rugby refers to Australian football).*

As shown in **Table 8**, out of the 34 articles analyzed, 11 (32.4%) focused solely on male athletes, whereas only one (2.9%) was centered around female athletes. The majority of the articles (12 or 35.3%) included both male and female athletes. Also, 10 articles (29.4%) did not report the gender of the participants.

As for the variable "Category participants", aims to identify the type of participants involved in the research (amateur, professional or both). The results, displayed in **Table 8**, reveal that 79.4% of the studies focused on professional participants, while only 14.7% involved amateurs. One article (2.9%) included both professional and amateur participants, and one article did not report the category of the participants (2.9%). However, it is worth mentioning that the exclusion criteria eliminated articles related to physical activity, recreational sports, and exercises and therefore many amateurs may have been excluded in this review.

Regarding the variable "Category classification", four different categories were defined: Sports performance analysis, Health, Sports Technology and Movement Integration. The results are presented in **Table 8**. Out of the total, 26 studies (76.5%) were focused on the sports performance analysis, making it the most studied category. A total of 14.7% of the studies were classified under the “Health” category, while the categories "Movement Integration" and "Sports Technology" were represented in 2 (5.9%) and 1 (2.9%) studies respectively.

4.1.3. General Characteristics of the Ordinal Regression Models

Table 9 presents the results of the characteristics of Ordinal regression models utilized in the selected 34 articles. The analysis reveals that the Proportional Odds model is the most used type of ordinal regression model, utilized in 11 articles (32.4%). On the other hand, the Mixed Effects Proportional Odds model is another type that is less frequently used, found in 11.8% of the articles. The newly

proposed ordinal regression models are grouped under the "Proposed Model" category, representing 23.5% of the articles. There is one instance where the Generalized Partial Credit model is used, representing 2.9% of the articles. Also, the type of ordinal regression model used is not reported in 10 articles (29.4%).

It is also worth noting that the proposed models used in the articles include the Dynamic Cumulative model, Functional Ordinal Logistic Regression (FOLR) model, General Latent Trait model, Generalization CUB models, Isotonic Least Squares Pairwise Comparison model, Latent Factor Distributed Lag model, and Ordinal Logistic Regression and Expectation (OLRE) model.

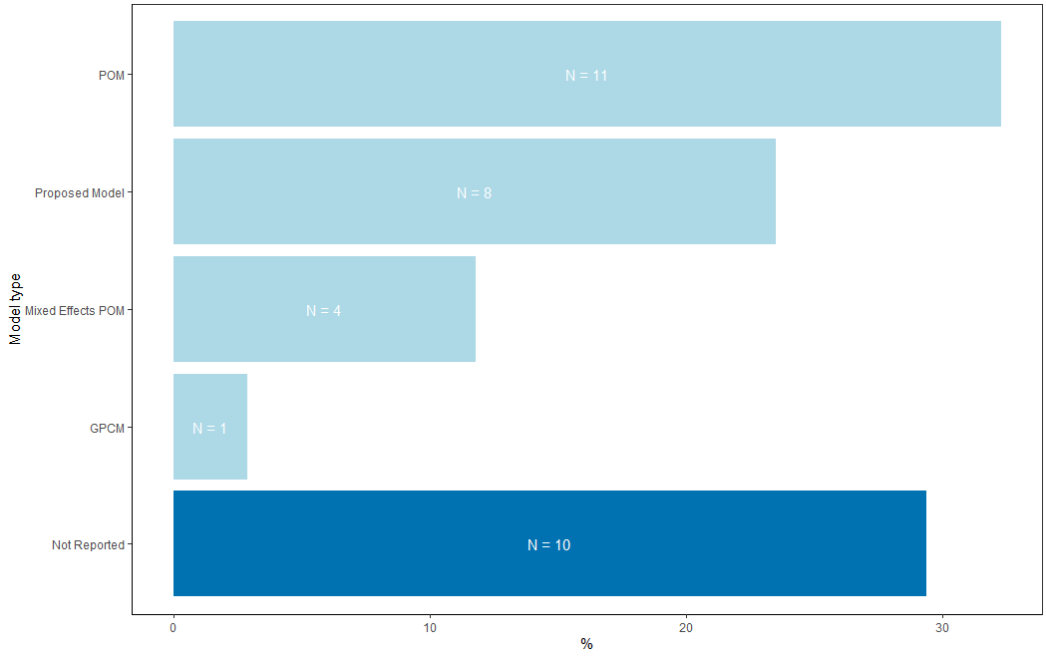


***Figure 16**:* **Frequency of Type of Ordinal Regression Model in which the Article Focuses**. *This chart displays the percentage and frequency of different types of ordinal regression model used in the selected articles to perform the analysis. The Proposed Models are with a frequency of 1: Dynamic Cumulative model; Functional Ordinal Logistic Regression model; General Latent Trait model; Generalization to CUB model; Generalized Partial Credit model; Isotonic Least Squares Pairwise Comparison model; Latent Factor Distributed Lag model. And with a frequency of 2: Ordinal Logistic Regression and Expectation.*

The following section presents the frequency of validation methods used in articles that employ different types of ordinal regression models. Both, Information Criterion and R-Squared are the most commonly used methods, found in at least 5 out of the 43 methods (11.6%) each. The Likelihood Ratio test is used in 4 cases, representing 9.3%, while Deviance and Residual metrics are used in 3 cases each, representing 7.0%. In the "Others" category, there are methods such as Posterior Intervals, Posterior Predictive Checks, and Log-Loss plot, each with a frequency of 1 and representing 7.0% of the methods. The least used method is Cross-Validation, which appears in 2 cases, representing 4.6%. Finally, 41.9% of articles do not report the validation method used.

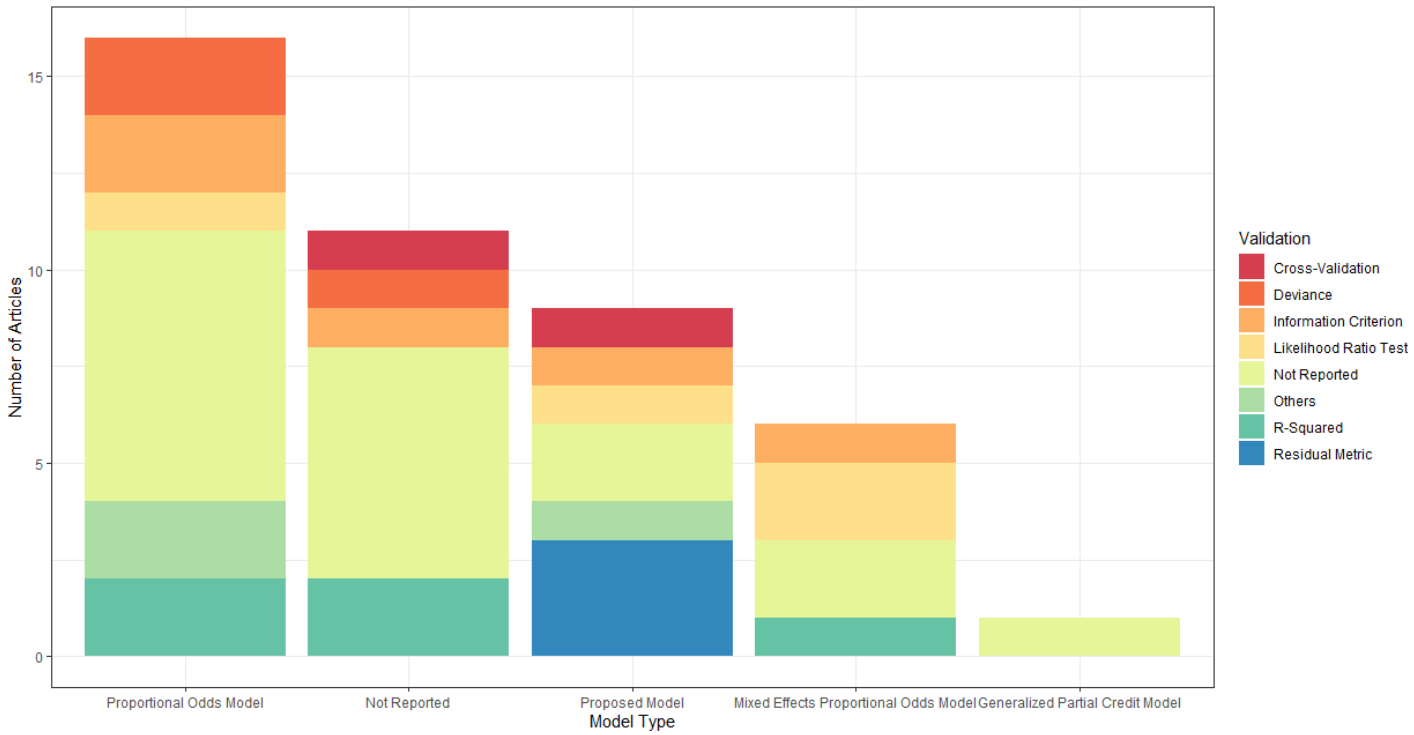


Figure 17: **Type of Validation Method Used per Type of Ordinal Regression Model.** *The chart displays the distribution of validation methods across different ordinal regression model types. For this stacked bar plot, each bar represents a specific model type and is segmented based on the types of validation methods used. The height of each segment indicates the frequency of the validation method used for the corresponding model type.*

Additionally, it was found that most of the reported ordinal models were based on the frequentist paradigm (82.4%), with a minority of models based on the Bayesian paradigm (17.7%). In terms of the objective of the reported models, the most common type was descriptive (38.2%), followed by predictive (32.4%). It is worth noting that in 29.4% of the articles, the objective of the model was not reported.

Based on the analysis, 36 software records were obtained due to multiple responses. The findings revealed that R (38.8%) and SPSS (19.4%) were the most frequently reported software. SAS/STAT and Stata were each reported twice (5.6%), whereas S-Plus 2000 was the least used software, representing only 2.8% of the reported software. In one case, the software availability was indicated to be upon request from the author, constituting 2.8% of the cases. It is noteworthy that a significant proportion of articles (25%) did not report the software used. Moreover, it was observed that 21 articles (61.8%) did not report the use of any specific statistical package, while 13 articles (38.2%) reported using one or more statistical packages. Among the articles reporting the use of R software, the 'ordinal' package was the most used, being reported 3 times (7.0%).

Finally, it was found that data sharing was not very common, with only 6 articles (17.6%) reporting data availability. In contrast, 25 articles (73.5%) did not report data sharing, and 3 articles (8.8%) indicated that data availability was available upon request. Similarly, code sharing was also not very common, with only 3 articles (8.8%) reporting code availability. In contrast, 30 articles (88.2%) did not report code sharing, and 1 article (3.0%) indicated that code availability was available upon request. Regarding repository sharing, it was found that only 4 articles (11.8%) reported sharing a repository with the data or code used in their studies. In contrast, 30 articles (88.2%) did not report repository sharing.

4.1.4. Bivariate Analysis

This section examines the bivariate analysis performed to explore the relationship between some variables in the selected studies. The primary focus is on showing the validation of the assumption of odds proportionality in Proportional Odds Models (POMs) and the reporting of statistical software used in relation to the type of model employed. Figures and tables are utilized to effectively present the findings and provide insights into the prevalence and patterns observed in the analyzed data.

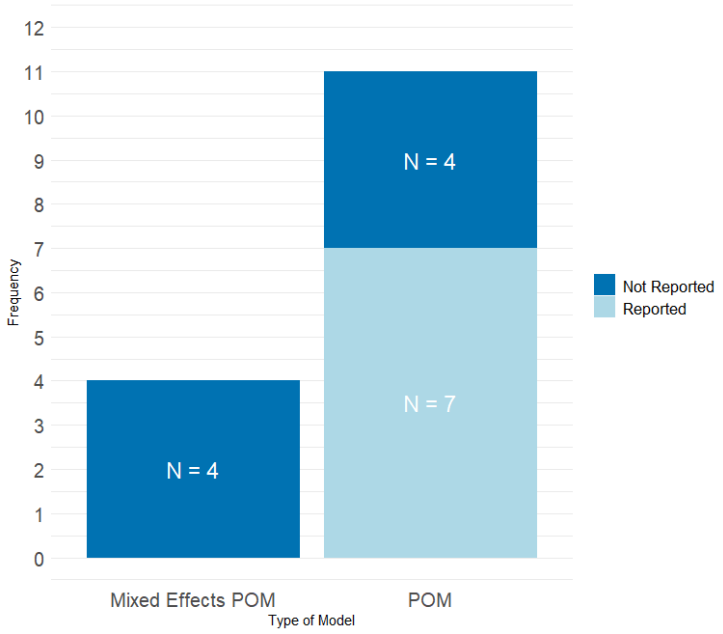


***Figure 18**:* **Frequency of Not Reported/Reported validations in POMs**. *This chart displays the frequency of POMs that did not report any use of validation method for the proportional odds assumption.*

Figure 18 illustrates the findings regarding the use of POMs in the reviewed studies. Among the 15 POMs, it was observed that 8 of them did not conduct tests to validate the assumption of odds proportionality. However, it is noteworthy that 7 articles did perform the necessary test to validate the assumption of odds proportionality.

***Table 10**: Cross Table of Statistical Software Not Reported in the selected articles per Type of Model Not Reported*

	Software Reported	***Software Not Reported***	*Total*
Type of Model Reported	20 (58.8%)	4 (11.8%)	24 (70.6%)
Type of Model Not Reported	5 (14.7%)	5 (14.7%)	10 (29.4%)
Total	25 (73.5%)	9 (26.5%)	**34**

The types of models that did not report the statistical software used where the following: 1 Mixed-Effects POM; 1 POM; 2 Proposed Models (Latent Factor Distributed Lag Model; Isotonic Least Squares Pairwise Comparison Model

Regarding the software used and the type of ordinal model employed, **Table 10** provides valuable insights. Out of the 9 articles that did not report any software used, 5 of them (14.7%) also failed to specify the type of model utilized. On the other hand, among the 4 models that did not report the software used, they did provide information about their specific model type. These models consisted of one Mixed-Effects POM, one standard POM, and two Proposed Models, the Latent Factor Distributed Lag model, and the Isotonic Least Squares Pairwise Comparison model. Most of the articles (58.8%) reported both the software and the type of model employed, while a small portion (14.7%) did not specify the model used but did mention the software.

5. Discussion

To the best of current knowledge, there has not been a review focusing specifically on how ordinal responses are reported in the sports field, being a relatively unexplored area of sports research. This study aims to fill this gap by providing insights into the appropriate analysis of such outcomes, ultimately improving understanding in this domain. Additionally, through exploring the existing literature, this study highlights the actual state of practice for ordinal regression models in the sports field. Moreover, the review of existing literature in the sports field using regression models for ordinal response has uncovered significant key findings and identified research gaps.

Getting deeper into findings, the articles selected in this review showed a significant increase in the number of articles that use ordinal regression models in sports field per year (see **Figure 14**), particularly in 2021, which coincided with the global COVID-19 pandemic. This increase may be attributed to the effects of the pandemic, including lockdowns and stay-at-home measures, which potentially influenced the focus on sports-related research (66–69). This trend aligns with a study by Selman et al. (22), which reported a rise in the use of ordinal outcomes in trials during the same period. However, it is worth noting that compared to other models such as generalized linear mixed models (GLMMs), and in the field of clinical medicine the number of articles selected, is slightly lower (70). This indicates the need for further exploration and experimentation specifically addressing ordinal regressions in sports field.

One research gap is the unequal representation of male (32.4%) and female (2.9%) athletes in the existing literature, with a prevailing focus on male athletes, along with a significant gap in the representation of amateur sports, appearing in only 14.7% of the articles examined. This observation aligns with recent research on R packages in sport (17), suggesting that these disparities extend beyond just ordinal regression models and encompass broader aspects such as gender representation and professional categorization. To address these dual disparities, it is crucial to promote greater inclusion of women's sports and allocate research efforts towards studying amateur sports in the published literature. Initiatives like Women in Sports Data (71) are actively working towards promoting these necessary changes. Additionally, the analysis highlighted that football (soccer), and basketball are the most extensively studied sports, emphasizing the importance of allocating more attention to underrepresented sports, particularly those popular in specific regions or countries. The prevalence of football-related articles can be attributed to the nature of the sport, where matches frequently result in ties, unlike basketball. The existence of tie outcomes in football introduces a three-level ordering in the outcome variable, offering researchers increased possibilities to investigate ordinal phenomena.

Examining the software used in the analyzed articles, the findings indicate that R (38.8%) and SPSS (19.4%) were the most frequently reported software packages. This observation suggests the popularity and effectiveness of these software's among the sports field. However, it is noteworthy that a significant proportion of articles (25%) did not report the software used, which may stem from various reasons, such as authors using proprietary software or neglecting to provide comprehensive details. Furthermore, a substantial number of articles (61.8%) did not report the use of any specific statistical package. It is crucial to consider the implications of these findings and emphasize the principles of FAIR (Findable, Accessible, Interoperable, and Reusable) and transparency in research (72). Adhering to FAIR principles in research software ensures similar benefits of enabling transparency, reproducibility, and reusability of research, making it easier for industry, science, education, and society to have effective access to software-based knowledge. Without proper reporting and reproducibility, the reliability of the studies may be compromised.

In terms of journal publications, it was found that many articles (47.1%) were published in sports and medicine journals, which aligns with the common use of ordinal regression models in studies involving

injury scores or treatment outcomes. Additionally, 26.5% of the articles were published in statistics and sports statistics journals, suggesting an increasing presence of statisticians in sports research. Nevertheless, it is crucial to emphasize the importance of fostering collaboration between statisticians and sports professionals, as advocated by Sainani et al. (1), who stressed the need for increased statistical support in sports science, sports medicine, and sports physiotherapy. Furthermore, most studies focused on sports performance analysis and health, with a relatively low presence of studies focused on movement integration and sports technology.

The type of objective pursued by the reported ordinal models showed an association with the type of paradigm, with Bayesian models primarily used for predictive purposes and frequentist models for descriptive purposes. However, some articles did not report (29.4%) the type of objective (e.g., describe, predict) pursued. A notable observation arising from this review is the absence of studies with a causal objective. This suggests that the inclusion of causal analysis in sports research is still relatively new and developing (73–77). The majority of studies continue to focus on descriptive and predictive approaches. This finding emphasizes the potential for future research to contribute significantly to exploring and understanding the causal aspects within the field of sports using ordinal models.

Regarding the type of models employed, the analysis reveals that the Proportional Odds model is the most commonly utilized ordinal regression model (32.4%), followed by the Mixed Effects Proportional Odds model, which was found in 11.8% of the articles. Notably, a considerable proportion of articles (23.5%) proposed novel ordinal regression models. Additionally, the specific type of ordinal regression model used was not reported in 10 articles, accounting for 29.4% of the total. This finding is consistent with the statement made by O'Connell et al. (78) who mentioned that the proportional odds model is widely used due to its simplicity and natural correspondence to ordinary logistic regression.

This study also revealed two significant limitations in reporting practices. Firstly, there is a lack of validation of the proportional odds assumption in studies utilizing POMs. The analysis, as shown in **Figure 18**, revealed that among the selected studies, a concerning 8 out of the 15 POMs did not conduct tests to validate the assumption of odds proportionality. The importance of testing this assumption cannot be overstated, as its failure to hold can lead to incorrect interpretations of results (79). This lack of reporting (53.3%) compromises the credibility of their findings, it is imperative for researchers to incorporate appropriate validation tests to ensure the validity of the proportional odds assumption.

Secondly, another crucial aspect that emerged from the analysis is the lack of reporting on software usage and the type of ordinal model employed. Interestingly, when the software is not reported, it is also common for the type of model to be omitted. This observation is illustrated in **Table 10**, where out of the 9 articles that did not report any software used, 5 of them also failed to mention the type of model utilized. This lack of reporting obscures essential information that is vital for understanding and replicating the study. Additionally, it is worth noting that among the reviewed studies, there were 4 models that did not report the software used yet provided information regarding their specific type. These models included one Mixed-Effects POM, one standard POM, and two Proposed Models, namely the Latent Factor Distributed Lag model and the Isotonic Least Squares Pairwise Comparison model. Interestingly, this lack of reporting aligns with findings from several studies. For instance, in a study by Plonsky et al. (80), it was observed that authors frequently fail to indicate the type of model being reported in approximately 29% of the cases. Furthermore, another study by Clatworthy et al. (81) highlighted that nearly half of the studies reviewed did not report the software used.

Therefore, based on these results, it is evident that there is a significant need for improved reporting practices in sports research for ordinal regression outcomes. Clear reporting of the type of model used,

along with its validation methods, statistical software employed, and the availability of data and code repositories, is crucial for enhancing transparency, reproducibility, and the overall quality of research in this domain. Moreover, this study identified limitations in reporting study design, a lack of representation in terms of sports and gender.

As mentioned earlier, the findings highlight the importance of increased statistical support in sports research. Collaboration between statisticians and sports professionals are necessary to improve the quality of statistical analysis in the sports field (82–84). Additionally, there is a need to prioritize statistical education and the development of analytical skills within the sports research community (85). These actions will play a crucial role in overcoming the current lack of reproducibility in scientific findings.

Limitations

One limitation of this study could be the relatively low number of articles identified, despite the review of a considerable number of articles. Although efforts were made to include widely recognized and relevant databases such as WoS, PubMed, JQAS, and JSA, there might still be articles in other existing sources that were not captured in this review. Therefore, findings of this review may not fully represent the entire landscape of research utilizing ordinal regression models in the sports field. This limitation highlights the need for future studies to explore additional databases and sources to ensure a more comprehensive coverage of the literature.

Another potential limitation is the coverage bias inherent in the selected databases and journals. While reputable, they may not include all relevant literature. The search terms and criteria used may introduce subjectivity and potentially exclude relevant studies. Although efforts were made to ensure a comprehensive search strategy, it is possible that some relevant studies were inadvertently missed.

On a positive note, this study is the first known scoping review of articles applying ordinal regression models in the field of sports. It provides a foundation for future research and serves as a valuable resource for promoting the use of statistics in sports-related domains. The review highlights the versatility of ordinal regression models in understanding and predicting various aspects of sports. The findings can inspire researchers to explore the use of these models in their own investigations, thereby advancing sports research. Future studies can address the identified limitations and contribute to a more comprehensive understanding of the application of ordinal regression models in the sports field.

Conclusions

During recent years, there has been a significant rise in publication of articles that use regression models for analyzing ordinal responses in the sports field. These models have primarily been employed in sports performance analysis and health-related studies, with a particular focus on sports like football and basketball. The majority of these articles have been published in journals from disciplines outside of statistics.

The findings of this review highlight the need for improvement in the quality of reporting when utilizing ordinal regression models. Particular emphasis should be placed on enhancing the analytical characteristics, validation procedures, meticulous reporting of the selection of suitable models, and data sharing. A stronger collaboration between statisticians and professionals from other fields is necessary to improve the overall quality of studies conducted. Additionally, the review reveals several

gaps that should be addressed. There is a clear need for greater representation of women's sports and amateur sports within the existing body of literature.

Based on the analysis and review of the included articles, it is highly recommended to advocate for the adoption of minimal rules and standardized guidelines of reporting when presenting the results of ordinal regression models. By adhering to these guidelines, researchers can ensure transparent and accurate reporting, ultimately enhancing the credibility and reproducibility of their findings.

While conducting this study, certain difficulties and limitations were encountered. One notable challenge was the time constraints associated with the TFG that posed a limitation on the extent to which the research could be conducted and the scope of the literature review. A more extensive investigation and inclusion of a larger number of studies could have provided a more comprehensive understanding of the topic.

Given more time, several aspects would have been further explored and expanded upon. Firstly, a more extensive analysis of different sports disciplines could have been undertaken to determine the specific patterns and trends in the use of regression models for ordinal responses. Furthermore, it would have been particularly interesting to delve deeper into the specific types of ordinal variables/outcomes used in the selected articles for the review. Analyzing the nature and characteristics of these ordinal variables could have provided valuable insights into the underlying factors influencing sports performance and health-related outcomes. Understanding the range and diversity of the ordinal responses employed, such as pain scales, injury severity, or performance rankings, would have allowed for a more nuanced examination of the applicability and effectiveness of regression models in capturing and analyzing these responses.

6. Study Case Example

6.1. Abstract

This section introduces a case study showcasing the utilization of two ordinal logistic regression models in the field of sports. Through the scoping review, it was identified that the Proportional Odds Model (POM) is the most commonly employed model in this field, followed by the Mixed-Effects Proportional Odds Model. Building upon these findings, this case study offers a practical demonstration of implementing the POM and the Mixed-Effects POM, aiming to provide a clear and organized example of how to properly implement these models.

This case example is based on the article published by Esteve et al. (36), where a linear mixed-effects model was employed to examine the association between the HAGOS Sports/Rec subscale, and 1) players that reported groin problems and those who didn't report groin problems, and 2) players reporting groin problems with and without time loss, considering the variability across weeks and individual players. For this study case, the same dataset and objective were addressed but using an ordinal regression model.

The study investigates the extent of groin problems in a full Spanish football season, involving 17 amateur male teams followed over 39 consecutive weeks. The prevalence of groin problems, both with and without time loss, is calculated by combining data on groin-injury time loss and self-reported groin pain. The Copenhagen Hip and Groin Outcome Score questionnaire's subscale assessing hip- and groin-related sporting function (HAGOS, Sport/Rec) is recorded every 4 weeks.

6.2. Data Source and Outcomes

Categorization of the response variable (HAGOS)

In this study, the HAGOS score (ranging from 0 to 100) is categorized into three levels. This categorization is based on previous research by Thorborg et al. (86), who aimed to assess the validity of the Copenhagen five-second squeeze test [9] by comparing it to the HAGOS questionnaire. The squeeze test measures pain during adductor activation, which aligns with the HAGOS (Sport) subscale focused on hip and groin function during sports. To establish validity, the researchers expected higher correlations between the squeeze test and the HAGOS (Sport) subscale. They also divided the pain rating scores into three groups: go (or 'safe'), attention (or 'acceptable'), and stop (or 'high risk'), following Thomee's pain monitoring model (87). The study revealed significant correlations between increased groin pain during the squeeze test and lower HAGOS (Sport) scores, supporting the validity of the test. There was also significant score differences among the three pain level groups.

Therefore, based on these findings, the HAGOS score in this study is categorized into the following three levels:

- **Stop (representing red)**: HAGOS scores from 0 to 61.
- **Attention (representing yellow)**: HAGOS scores from 62 to 84.
- **Go (representing green)**: HAGOS scores from 85 to 100.

[9] The Copenhagen five-second squeeze test is performed by placing one arm between the ankles of the player, instructing the player to isometrically and continuously squeeze as hard as possible for 5 seconds and subsequently rate the pain experienced in the groin during this test on a 0–10 Numerical Rating Scale (NRS). This NRS ratings are divided into three groups. Group 1: (NRS=0–2), which is considered safe, group 2, (NRS=3–5) considered acceptable, and group 3 (NRS=6–10), considered high risk.

This categorization aligns with the colors of a traffic light, representing different levels of pain or injury. The red light indicates the player should 'stop' current football activities and should seek clinical assessment by a health professional. Yellow light indicates 'attention', with players requiring clinical review by a health professional before determining their participation in football. Lastly, the green light represents a 'go' for football participation, but players still need a clinical review by a health professional with regards to their return to football decisions.

Outcomes

The primary outcome of this study is the prevalence of all groin problems. All groin problems included groin problems with and without time loss. As stated in the article (36), a player was considered as having a groin problem with time loss when he missed at least one training session or a match due to a groin injury. When a player answered "yes" to the question "Did you have groin pain (inguinal region including adductors, lower abdomen, and pubic region) in the past week?", but no time loss was reported due to a groin problem in the week in question, the player was considered as having a groin problem without time loss. From this information, players were classified, as 1) experiencing a groin problem with time loss, 2) experiencing a groin problem without time loss, or 3) not experiencing a groin problem.

The secondary outcome of this study is the categorization of HAGOS, Sports/Rec subscale score into three levels of pain monitoring (stop/attention/go) according to the previously explained methodology.

6.3. Analysis and Interpretation

6.3.1. Applying Proportional Odds Model in Sports Research

Before applying the POM, it is worth noting that the dataset used comprises longitudinal data with a temporal component. However, to adhere to the cross-sectional nature required for the POM, a targeted time window is selected, allowing for the extraction of data pertaining to a specific week. By considering a single week from the study, individuals are effectively represented only once in the dataset. This approach ensures independence among the observations, complying with the assumptions of the POM.

As the primary goal of this study case is to provide a clear example of how to properly implement this model, only modelling outputs will be presented and all details of the estimates of the POM will be shown in the Statistical Methods section.

Data Analysis

To assess differences in pain monitoring (categorized HAGOS, Sports/Rec subscale) 1) between players reporting groin problems and players not reporting groin problems, and 2) between players reporting groin problems with and without time loss, respectively, two proportional odds model (POM) were performed. Thus, before the modelling, the proportional odds assumption was assessed using the *Brant test*. Furthermore, the Likelihood ratio test was performed to evaluate the model.

For the first model, the independent variable "Groin problem" (yes/no) was included as fixed effect, with the base category being "Groin Problem = YES".
For the second model, the independent variable "Time loss groin problem" (yes/no) was included as fixed effect, with the base category being "Time loss groin problem = NO".

The pain monitoring (stop/attention/go) was set as the dependent variable in both models.

It is important to note that in this context, the base categories of the independent variables represent the higher level. As for the dependent variable, pain monitoring, the "stop" category is considered the higher level, while the "go" category is considered as the lower level.

Analyses were performed using statistical software RStudio (R Core Team (2022). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.), with the package ***MASS*** (Venables, W. N. & Ripley, B. D. (2002) Modern Applied Statistics with S. Fourth Edition. Springer, New York. ISBN 0-387-95457-0) and ***brant*** (Benjamin Schlegel and Marco Steenbergen (2020). brant: Test for Parallel Regression Assumption. R package version 0.3-0. https://CRAN.R-project.org/package=brant). The ***effects*** package (John Fox, Jangman Hong (2009). Effect Displays in R for Multinomial and Proportional-Odds Logit Models: Extensions to the effects Package. Journal of Statistical Software, 32(1), 1-24. doi 10.18637/jss.v032.i01) was used for graphical representation. A p-value below 0.05 was considered significant in all analyses. All players from week 10 participating in the study were included in the analyses as long as they were part of one of the included teams.

Statistical Methods

Model 1: Proportional odds model fit to assess differences in pain monitoring between players reporting groin problems and players not reporting groin problems.

Previously to fitting the model, it is crucial to evaluate whether the proportional odds assumption is tenable. For this, the *Brant test* from the *brant* package was used. Results showed that this model satisfies this proportional odds assumption since the probability value (0.07) is greater than alpha (0.05).

To fit this model the statistical function (*polr*), available in the R package *MASS was used*.

Table 11: *Proportional Odds Model estimates of pain monitoring between players reporting and player not reporting groin problems.*

Coefficients	*Estimate*	*SE*	*OR*	*95% CI*	*Pr. LR test*
Groin Problem (Yes)	-2.65	0.33	0.07	(0.036 ; 0.14)	< 0.001
Intercepts	*Estimate*	*SE*	*t-value*	*p-value*	
Stop|Attention	-3.81	0.31	-12.42	< 0.001	
Attention|Go	-1.97	0.180	-10.93	< 0.001	

SE: Standard Error; OR: Odds Ratio; LR test: Likelihood Ratio test

The intercept 'Stop|Attention' represents the logarithm of the odds of requiring the football practice to be stopped compared to needing attention or being able to continue playing ($logit[P(Y \leq 1)]$. In other words, it can be interpreted as the log of odds of the situation where the player must stop playing versus situations where attention is needed, or they can continue playing.

Likewise, the intercept 'Attention|Go' corresponds to the logarithm of the odds of requiring either a stop or attention compared to being able to continue playing ($logit[P(Y \leq 2)]$). It can be interpreted as the log of odds of the situations where the player needs to stop playing or requires attention versus situations where they can continue playing without any issues.

With these estimates, we can calculate the cumulative probabilities for both Groin Problem outcomes (Yes or No).

Table 12: Cumulative probabilities for players reporting groin problem prevalence and players not reporting groin problems.

Cumulative probabilities for Groin Problem Prevalence = NO		Cumulative probabilities for Groin Problem Prevalence = YES	
Stop|Attention	**Attention|Go**	**Stop|Attention**	**Attention|Go**
0.022	0.12	0.24	0.66

The cumulative probabilities indicate the likelihood of experiencing a specific outcome (Stop/Attention/Go) regarding the groin problem prevalence. For players reporting no groin problems the cumulative probability of needing to stop compared to needing attention or continue football practice is only 2%, and the cumulative probability of needing to stop or attention compared to continue playing is 12%.
On the other hand, for players reporting yes to groin problems, the cumulative probability of needing to stop compared to needing attention or continue football practice is 24%, while the cumulative probability of needing to stop or attention compared to continue playing is 66%.
Hence, in cases where no groin problem is reported, it is more likely that the player will continue their current football practice without any major issues. On the other hand, when a groin problem is reported, there is a greater probability that the player will either need to stop their practice or require attention, rather than being able to continue playing without any significant concerns.

Then it is possible to calculate the odds by comparing the probability of success to the probability of failure. By calculating the ratio of these probabilities, we obtain the odds. We can perform this calculation for all levels to analyze both outcomes.

Table 13: Odds for players reporting groin problem prevalence and players not reporting groin problems.

Odds for Groin Problem Prevalence = NO		Odds for Groin Problem Prevalence = YES	
Stop|Attention	**Attention|Go**	**Stop|Attention**	**Attention|Go**
0.022	0.14	0.31	1.96

For players not reporting groin problems, the odds of needing attention or continue football practice are 45.45 (1/0.022) times the odds of needing to stop current practice, and the odds of continue football practice are 7.14 times the odds of needing to stop or attention.
On the other hand, for players reporting groin problems, the odds of needing attention or continue playing are 3.23 times the odds of needing to stop, while the odds of needing to stop or attention are 1.96 times the odds to continue playing.
Therefore, once again, when no groin problem is reported, it is more likely that the player will continue playing without problem. However, when a groin problem is reported, there is a greater probability that the player will either need to stop or require attention, rather than being able to continue playing.

Additionally, we can compare the odds of players reporting groin problems to the odds of players not reporting groin problems. The ratio between these two odds is 14.09. This means that, regardless of the level of groin problem prevalence, players who report groin problems have approximately 14.09 times higher odds of needing to stop playing compared to players who don't report groin problems.

Based on the POM results, we can calculate the predicted probabilities for each pain level category.

***Table 14**: Predicted probabilities for players reporting groin problem prevalence and players not reporting groin problems.*

Predicted probabilities for Groin Problem Prevalence = NO			Predicted probabilities for Groin Problem Prevalence = YES		
Stop	**Attention**	**Go**	**Stop**	**Attention**	**Go**
0.022	0.10	0.88	0.24	0.43	0.34

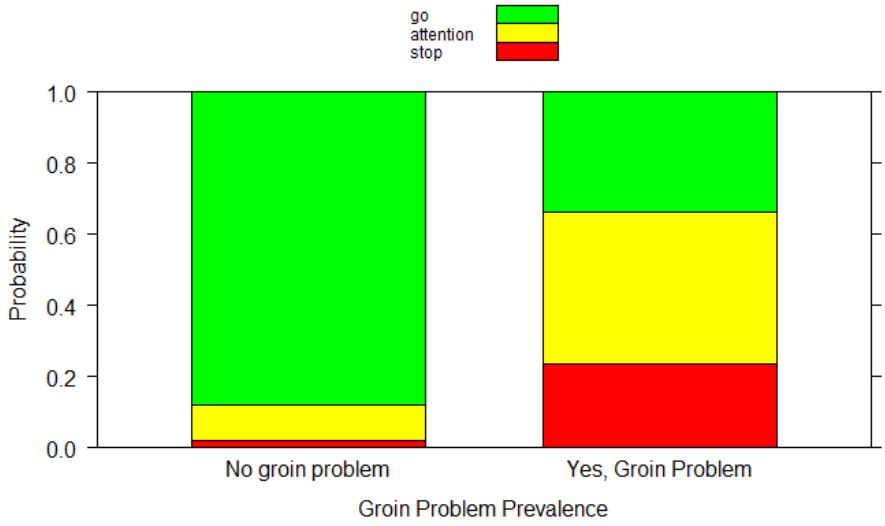


***Figure 19:* Predicted Probabilities of each outcome.** *This chart displays the predicted probabilities of each outcome category based on the fitted Proportional Odds Model (POM). The x-axis represents the groin problem prevalence variable, while the y-axis represents the predicted probabilities ranging from 0 to 1. Each block in the chart corresponds to a different outcome category, representing different levels of severity. Red light indicates the player should 'stop' current football activity. Yellow light indicates 'attention', the player should be clinically reviewed by a health professional before football participation is decided on, and green light indicates 'go' for football participation.*

These probabilities provide estimates of the likelihood of an individual being categorized into different levels of pain based on their groin problem prevalence. For instance, when individuals do not report groin problems, there is an 88% probability that they can continue playing without any significant issues. Only 10% of these individuals would require attention, indicating a potential need for monitoring or additional care. Furthermore, less than 5% of individuals without reported groin problems would need to stop playing due to injury concerns. In contrast, among players who have reported groin problems, the probability of being able to continue playing without any concerns decreases to 34%. A higher proportion, 43%, would require attention, suggesting a need for closer monitoring or medical intervention. Additionally, 24% of individuals with reported groin problems would need to stop playing due to the severity of their injuries (**Figure 19**).

Model 2: Proportional odds model fit to assess differences in pain monitoring between players reporting groin problems with and without time loss.

This model also satisfies this proportional odds assumption since the probability value (0.96) is greater than alpha (0.05)

***Table 15**: Proportional Odds Model estimates of pain monitoring between players reporting groin problems with and without time-loss.*

Coefficients	*Estimate*	*SE*	*OR*	*95% CI*	*Pr. LR test*
Time-loss Groin Problem (No)	0.24	0.24	1.27	(0.26 ; 6.37)	0.76
Intercepts	*Estimate*	*SE*	*t-value*	*p-value*	
Stop|Attention	-0.72	0.77	-0.93	0.35	
Attention|Go	0.72	0.77	0.93	0.35	

SE: Standard Error; OR: Odds Ratio; LR test: Likelihood Ratio test

We can calculate the cumulative probabilities for both Time-loss groin problem outcomes (Yes or No).

***Table 16**: Cumulative probabilities for players reporting groin problems with and without time loss.*

Cumulative probabilities for Time-loss Groin Problem = NO		Cumulative probabilities for Time-loss Groin Problem = YES	
Stop|Attention	**Attention|Go**	**Stop|Attention**	**Attention|Go**
0.28	0.62	0.33	0.67

For players reporting groin problems without time-loss the cumulative probability of needing to stop compared to needing attention or continue football practice is 28%, and the cumulative probability of needing to stop or attention compared to continue playing is 62%.
On the other hand, for players reporting groin problems with time-loss, the cumulative probability of needing to stop compared to needing attention or continue football practice is 33%, while the cumulative probability of needing to stop or attention compared to continue playing is 67%.
Hence, in this case, there is no great difference between players reporting groin problem with or without time loss. In both cases, there are higher probabilities that player will continue playing or requiring attention, rather than needing to stop their practice.

Then we calculate the odds.

***Table 17**: Odds for players reporting groin problems with and without time loss.*

Odds for Time-loss Groin Problem Prevalence = NO		Odds for Time-loss Groin Problem Prevalence = YES	
Stop|Attention	**Attention|Go**	**Stop|Attention**	**Attention|Go**
0.38	1.61	0.49	2.05

For players reporting groin problems without time loss, the odds of needing attention or continue football practice are 2.63 times the odds of needing to stop current practice, and the odds of needing to stop or attention are 1.61 times the odds of continue playing.
On the other hand, for players reporting groin problems with time loss, the odds of needing attention or continue playing are 2.04 times the odds of needing to stop, while the odds of needing to stop or attention are 2.05 times the odds to continue playing.

The ratio between the odds of players reporting groin problems without time loss to the odds of players reporting groin problems with time-loss is 0.79. This means that, regardless of the level of groin problem prevalence, players who report groin problems without time-loss have approximately 21%

times lower odds of needing to stop playing compared to players who report groin problems with time-loss.

We can calculate the predicted probabilities for each pain level category.

***Table 18**: Predicted probabilities for players reporting groin problems with and without time loss.*

Predicted probabilities for Time-loss Groin Problem = NO			Predicted probabilities for Time-loss Groin Problem = YES		
Stop	**Attention**	**Go**	**Stop**	**Attention**	**Go**
0.28	0.34	0.38	0.33	0.34	0.33

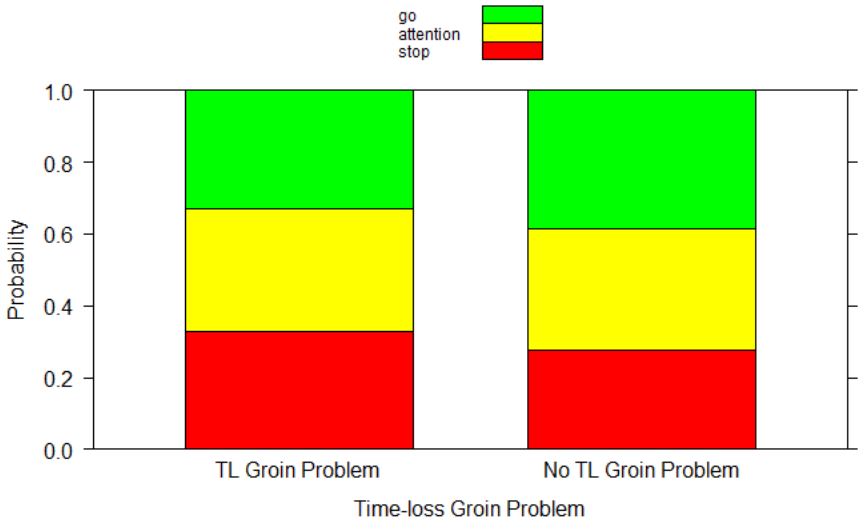


***Figure 20:* Predicted Probabilities of each outcome.** *This chart displays the predicted probabilities of each outcome category based on the fitted Proportional Odds Model (POM). The x-axis represents the time-loss groin problem variable, while the y-axis represents the predicted probabilities ranging from 0 to 1. Each block in the chart corresponds to a different outcome category, representing different levels of severity. Red light indicates the player should 'stop' current football activity. Yellow light indicates 'attention', the player should be clinically reviewed by a health professional before football participation is decided on, and green light indicates 'go' for football participation.*

After calculating the predicted probabilities from this model, we can observe that there are no differences among these probabilities for the three levels of the response variable indicating pain. For each level of the time loss groin problem variable, whether with or without time loss, the probabilities of belonging to any level of the response variable for pain hover around 30%.

Results

Model 1 aimed to assess the differences in pain monitoring between players reporting groin problems and players not reporting groin problems. The estimated coefficients from the model are presented in **Table 11**. Players reporting groin problems (Groin Problem = Yes) had a coefficient of -2.65, indicating a significant negative association with pain monitoring (OR = 0.07, 95% CI: 0.036; 0.14, $p < 0.001$), meaning that for the higher category of Groin Problem (Yes), there was a tendency for a lower category of pain monitoring (stop or attention). The 'Stop|Attention' intercept had a coefficient of -3.81 (t-value = -12.42, $p < 0.001$), while the 'Attention|Go' intercept had a coefficient of -1.97 (t-value = -10.93, $p <$

0.001). The odds ratios (OR) showed that for players reporting groin problems there were 14.09 higher odds of needing to stop playing compared to players not reporting groin problems.
Finally, the predicted probabilities presented in **Figure 19** indicated that for players not reporting groin problems, there was higher probability of being able to continue playing (88%). However, among players reporting groin problems, the probability of being able to continue playing decreased significantly (34%), to 34%, while the probabilities of needing attention and requiring a stop were 43% and 24%, respectively.

Model 2 model aimed to assess the differences in pain monitoring between players reporting groin problems with and without time loss. **Table 15** presents the estimated coefficients for this model. The coefficient for players reporting groin problems with time loss (Time-loss Groin Problem = No) was 0.24, suggesting direct association with pain monitoring. However, no significant difference in pain compared to players reporting groin problems without time loss (OR = 1.27, 95% CI: 0.26; 6.37, p = 0.76).

6.3.2. Applying Mixed-Effects POM in Sports Research

Before proceeding with the model application, it is important to consider that the dataset utilized in this case study contains the complete longitudinal data with a temporal component. Additionally, it is worth mentioning that while the study conducted by Esteve et al. (36), which employed a linear mixed model, utilized nested effects (i.e., an intercept and slope for each player's progression over weeks), in this case we will model the effects of week and player as crossed due to convergence issues.

Like before, the main objective of this case study is to provide a clear and illustrative example of the proper implementation of the Mixed-Effects POM. Therefore, in the Statistical Methods section, the modeling outputs will be presented and highlight essential details concerning the estimates of the Mixed-Effects POM.

Data Analysis

To assess differences in pain monitoring 1) between players reporting groin problems and players not reporting groin problems, and 2) between players reporting groin problems with and without time loss, respectively, two mixed effects proportional odds model (Mixed-Effects POM) were performed. The mixed effects proportional odds assumption was assessed using the nominal test.

For the first model, the independent variable “Groin problem” (yes/no), with the base category being “Groin Problem = YES”, and the variable week were included as fixed effects, while players and week number were considered as crossed effects.
For the second model, the independent variable “Time loss groin problem” (yes/no), with the base category being “Time loss groin problem = NO”, and the variable week were included as fixed effects, while players and week number were considered as crossed effects.
The pain monitoring (stop/attention/go) was set as the dependent variable in both models.

It is important to note that like in POM, the base categories of the independent variables represent the higher level. As for the dependent variable, pain monitoring, the “stop” category is considered the higher level, while the “go” category is considered as the lower level.

Analyses were performed using statistical software RStudio, with the package *ordinal* (Christensen, R. H. B. (2022). ordinal - Regression Models for Ordinal Data. R package version 2022.11-16. https://CRAN.R-project.org/package=ordinal.), *lsmeans* (Russell V. Lenth (2016). Least-Squares Means:

The R Package lsmeans. Journal of Statistical Software, 69(1), 1-33. doi:10.18637/jss.v069.i01) and *sjPlot* (Lüdecke D (2023). _sjPlot: Data Visualization for Statistics in Social Science_. R package version 2.8.14, <URL:https://CRAN.R-project.org/package=sjPlot>.). The *sjPlot* package was also used for graphical representation. A p-value below 0.05 was considered significant in all analyses. All players participating in the study were included in the analyses as long as they were part of one of the included teams.

Statistical Methods

It is important to note that in mixed effects POM, unlike in POM, independence is not an assumption. Therefore, the utilization of a repeated-measure design is applicable and valid.

Model 3: Mixed effects proportional odds model fit to assess differences in pain monitoring between players reporting groin problems and players not reporting groin problems.

To assess the appropriateness of this model, it is crucial to evaluate whether the proportional odds assumption is tenable. For this case, there is no function available within the ordinal package to do this directly. However, we can employ an alternative approach by conducting separate logistic regressions for each pairwise comparison and obtaining two sets of odds ratios (ORs) and intercepts. Additionally, we can calculate the ORs for different collapsing levels of the ordinal response variable. The procedure involves:

- Fitting a mixed-effects logistic regression model, collapsing the first and second categories of the ordinal response variable, and calculate the odds ratio and its confidence interval for the covariate.
- Repeat the process, collapsing the second and third categories of the response variable.

If the confidence intervals of the two odds ratios overlap, it indicates that the proportional odds assumption holds for that particular covariate.

Reference articles (88,89) were utilized to guide the approach and methodology for evaluating this assumption.

***Table 19**: Test result for assessing the proportional odds assumption.*

Variable	*OR*	*95% CI*
OR for Groin Problem and pain >= "attention"	0.018	(-0.38 ; 0.42)
OR for Groin Problem and pain >= "stop"	0.028	(-0.497 ; 0.55)

SE: Standard Error

This model satisfies the proportional odds assumption for Groin Problem since the confidence intervals of the two odds ratios overlaps.

To fit this model a statistical function (*clmm*) is available in R package *Ordinal*.

***Table 20**: Mixed Effects Proportional Odds Model estimates of pain monitoring between players reporting and player not reporting groin problems.*

Variable	*Estimate*	*SE*	*OR*	*95% CI*	*p-value*
Groin Problem (Yes)	-3.57	0.33	0.03	(0.02 – 0.04)	< 0.001
Week	0.009	0.006	1.01	(1.00 – 1.02)	0.16
SD of random effect (player)	1.99				
Conditional R^2	0.632				
Threshold coefficients	*Estimate*	*SE*			
Stop|Attention	-5.63	0.28			< 0.001
Attention|Go	-3.14	0.23			< 0.001

SD Standard Deviation; SE: Standard Error; OR: Odds Ratio

Based on the results, we can calculate the predicted probabilities for each pain level category.

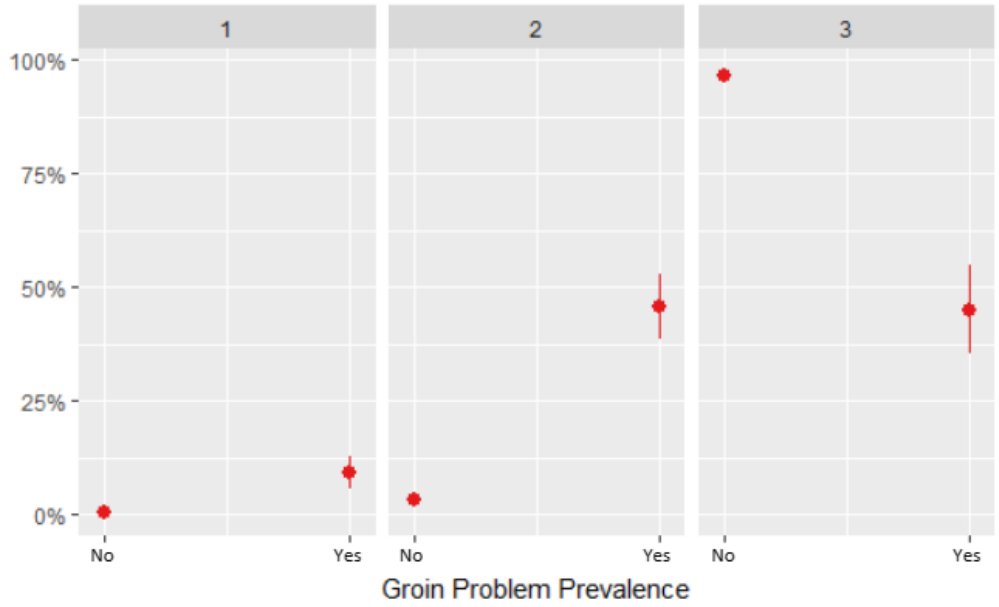


***Figure 21**: This chart displays the predicted probabilities of each outcome category based on the fitted Mixed-Effects Proportional Odds Model. The x-axis represents the groin problem variable, while the y-axis represents the predicted probabilities ranging from 0 to 1. Each box and number correspond to a different outcome category, representing different levels of severity. 1 indicates the player should 'stop' current football activity. 2 indicates 'attention', the player should be clinically reviewed by a health professional before football participation is decided on, and 3 indicates 'go' for football participation.*

Figure 21 show that when individuals do not report groin problems (red line on the left of every box), there is approximately a 90% probability that they can continue playing without any significant issues. Only 5% approximate of these individuals would require attention, indicating a potential need for monitoring or additional care. Furthermore, less than 5% of individuals without reported groin problems would need to stop playing due to injury concerns. In contrast, among players who have reported groin problems (red line on the right of every box), the probability of being able to continue playing without any concerns decreases to 40% approximately. A similar proportion would require attention. Additionally, 10% of individuals with reported groin problems would need to stop playing.

Model 4: Mixed effects proportional odds model fit to assess differences in pain monitoring between players reporting groin problems with and without time loss.

Once again, it is crucial to evaluate the proportional odds assumption.

Table 21: *Test result for assessing the proportional odds assumption.*

Variable	*OR*	*95% CI*
OR for Time-loss Groin Problem and pain >= "attention"	0.51	(-0.49 ; 1.51)
OR for Time-loss Groin Problem and pain >= "stop"	3.599	(2.55 ; 4.65)

SE: Standard Error

The current model applied to the Time-loss Groin Problem does not meet the proportional odds assumption. This is evident from the non-overlapping confidence intervals of the two odds ratios (see **Table 21**). Consequently, the reliability of the model's results is compromised. In this situation, it may be necessary to explore alternative modeling approaches, such as a partial proportional odds model or a generalized ordered logit model. These models accommodate varying effects of certain covariates across the levels of the ordinal response variable. However, given the context of this thesis and having as a goal to illustrate the ordinal models, it will continue to proceed with the current model even without fulfilling the assumption. Nevertheless, it is essential to acknowledge and demonstrate that the assumption has been evaluated.

Table 22*: Mixed-Effects POM estimates of pain monitoring between players reporting groin problems with and without time-loss.*

Variable	*Estimate*	*SE*	*OR*	*95% CI*	*p-value*
Time-loss Groin Problem (No)	0.26	0.43	1.29	(0.56 – 2.98)	0.54
Week	-0.03	0.009	0.97	(0.96 – 0.99)	0.005
SD of random effect (player)	2.22				
Conditional R^2	NA				
Threshold coefficients	*Estimate*	*SE*			
Stop|Attention	-2.25	0.28			< 0.001
Attention|Go	0.91	0.23			0.056

SD Standard Deviation; SE: Standard Error; OR: Odds Ratio

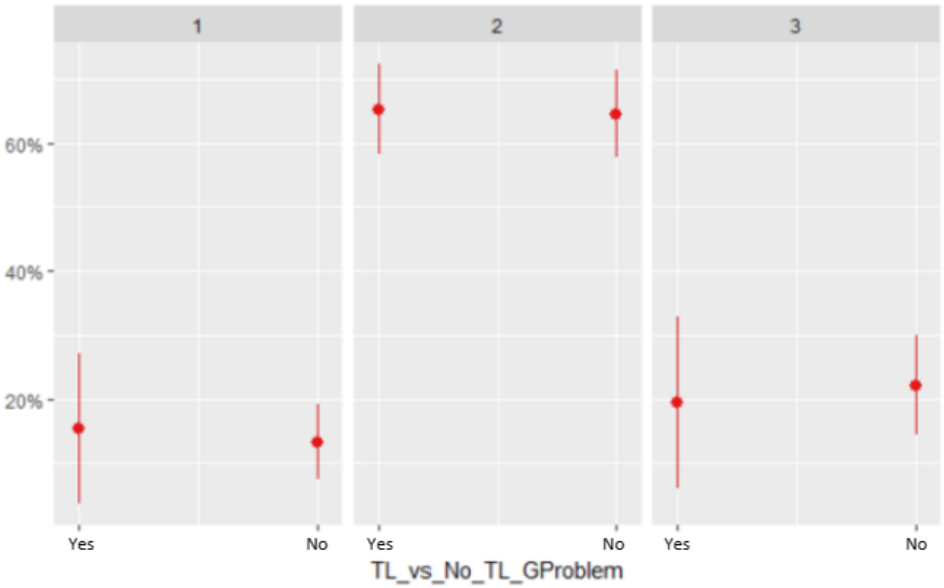


Figure 22*: This chart displays the predicted probabilities of each outcome category based on the fitted Mixed-Effects Proportional Odds Model. The x-axis represents the time-loss groin problem variable, while the y-axis represents the predicted probabilities ranging from 0 to 1. Each box and number correspond to a different outcome category, representing different levels of severity. 1 indicates the player should 'stop' current football activity. 2 indicates 'attention', the player should be clinically reviewed by a health professional before football participation is decided on, and 3 indicates 'go' for football participation.*

From **Figure 22** we can observe that there are no differences among these probabilities within the three levels of the response variable. However, there are differences between the three levels. For levels 1 and 3 ('stop' and 'go') the probabilities of belonging to these levels, whether with or without time loss, are around 20%. And for level 2 ('attention') the probability of belonging to this level, whether with or without time loss, is around 65%

For more comprehensive information on an alternative approach to test the proportional odds assumption, along with additional details on assessing the significance of differences in pain monitoring for the mixed-effects proportional odds model, please refer to the Supplementary file (Appendix).

Results

In concordance with the results obtained from the POM analysis, Model 3, which aimed to assess differences in pain monitoring between players reporting groin problems and those not reporting groin problems, showed significant findings. Players reporting groin problems (Groin Problem = Yes) had a coefficient of -3.57, indicating a significant negative association with pain monitoring (OR = 0.03, 95% CI: 0.02; 0.04, $p < 0.001$), meaning that for the higher category of Groin Problem (Yes), there was a tendency for a lower category of pain monitoring (stop or attention). The odds of the injury being less severe (i.e., go or attention versus stop) for players with groin problems were 97% lower than athletes without groin problems. Additionally, the effect of the number of weeks did not show a statistically significant association (OR: 1.01, 95% CI: 1.00; 1.02, $p = 0.16$). The odds of needing to stop did not significantly change as the weeks progressed.

Model 4 aimed to assess differences in pain monitoring between players reporting groin problems with and without time loss. However, this model did not meet the proportional odds assumption. Consequently, the following results are not reliable. However, if we disregard this fact, the analysis still indicates no significant difference in pain monitoring between players reporting groin problems with or without time loss (OR: 1.29, 95% CI: 0.56; 2.98, $p = 0.54$). This suggests that the presence of time loss did not substantially impact the severity of the pain in terms of sport activities.

Comparison of Interpretive Insights between Mixed Effects POM and LMM

To gain a comprehensive understanding of the associations and interpretative benefits between the mixed effects proportional odds models and the previously conducted linear mixed model (LMM), we compared the results obtained from both analyses.

The LMM revealed that players reporting any groin problem exhibited lower HAGOS, Sports/Rec subscale scores compared to those not reporting groin problems. The mean difference between the two groups was statistically significant (-19.5, 95% CI: -20.7 – -18.4, $p < 0.001$). This finding aligns with the results from the mixed effects POM, indicating a consistent association between groin problems and injury being more severe (and consequently lower HAGOS score).

However, when examining the comparison between players reporting groin problems with and without time loss, the results differed between the two types of models. The mixed effects POM, due to violation of the proportional odds assumption, did not provide reliable results. Nevertheless, when disregarding this limitation, the model indicated no significant difference in pain monitoring between the two groups. On the other hand, the LMM also indicated that there was no difference in Sports/Rec HAGOS subscale scores between players reporting groin problem with or without time loss (mean difference: 4.0 (95% CI: -1.1 – 9.1), $p = 0.12$).

In terms of interpretative benefits, the mixed effects POM offers a more nuanced understanding of the severity of groin problems by analyzing the odds of injury being less severe (go or attention) compared to players without groin problems. This additional information enhances the clinical interpretation of the results and helps healthcare professionals gain insights into the potential impact of groin problems on athletes' performance. The LMM, on the other hand, provide a broader perspective by examining response rates, overall prevalence of groin problems, and descriptive subscale scores.

While the results obtained from the two types of models may differ in terms of statistical significance and specific numerical values, both approaches ultimately highlight the negative impact of groin problems on football practice.

7. References

1. Sainani KL, Borg DN, Caldwell AR, Butson ML, Tenan MS, Vickers AJ, et al. Call to increase statistical collaboration in sports science, sport and exercise medicine and sports physiotherapy. Br J Sports Med. 2021 Jan;55(2):118–22.

2. Borg DN, Bon J, Sainani KL, Baguley BJ, Tierney N, Drovandi C. Sharing data and code: a comment on the call for the adoption of more transparent research practices in sport and exercise science. 2020;

3. Borg DN, Lohse KR, Sainani KL. Ten common statistical errors from all phases of research, and their fixes. PM&R. 2020;12(6):610–4.

4. Casals M, Finch CF. Sports Biostatistician: a critical member of all sports science and medicine teams for injury prevention. Inj Prev. 2017;23(6):423–7.

5. Lopez MJ. Bigger data, better questions, and a return to fourth down behavior: an introduction to a special issue on tracking datain the National football League. J Quant Anal Sports. 2020;16(2):73–9.

6. Jim Albert MEG, Tim B. Swartz RHK. Handbook of Statistical Methods and Analyses in Sports [Internet]. Available from: https://www.routledge.com/Handbook-of-Statistical-Methods-and-Analyses-in-Sports/Albert-Glickman-Swartz-Koning/p/book/9780367331016

7. Department of Biostatistics [Internet]. 2016 [cited 2023 Feb 19]. The Sports Analytics Laboratory at Harvard. Available from: https://www.hsph.harvard.edu/biostatistics/2016/08/the-sports-analytics-laboratory-at-harvard/

8. Sports Analytics Group at Berkeley [Internet]. [cited 2023 Feb 19]. Available from: https://sportsanalytics.berkeley.edu/

9. Sport Analytics Major - Syracuse.edu [Internet]. [cited 2023 Feb 19]. Available from: https://www.syracuse.edu/academics/undergraduate-majors-minors/sport-analytics/

10. Stats in Sports [Internet]. [cited 2023 Feb 19]. Available from: https://www.stat.cmu.edu/statsinsports/

11. Moneyball. Sony Pictures Releasing; 2011.

12. Concussion. Sony Pictures Releasing; 2015.

13. Sports Science Applications [Internet]. BEST Performance Group. 2014 [cited 2023 Feb 20]. Available from: http://bestperformancegroup.com/?page_id=1288

14. Baumer BS, Matthews GJ, Nguyen Q. Big Ideas in Sports Analytics and Statistical Tools for their Investigation [Internet]. arXiv; 2023 [cited 2023 Feb 19]. Available from: http://arxiv.org/abs/2301.04001

15. R Core Team (2022). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

16. Baumer BS, Nguyen Q, Matthews GJ. CRAN Task View: Sports Analytics [Internet]. Comprehensive R Archive Network (CRAN); 2023 [cited 2023 Feb 19]. Available from: https://CRAN.R-project.org/view=SportsAnalytics

17. Casals M, Fernández J, Martínez V, Lopez M, Langohr K, Cortés J. A systematic review of sport-related packages within the R CRAN repository. Int J Sports Sci Coach. 2022 Nov 11;17479541221136238.

18. Agresti A, John Wiley & Sons. Analysis of Ordinal Categorical Data. New York; 2002.

19. Agresti A, John Wiley & Sons. Analysis of Ordinal Categorical Data. Second Edition. Florida; 2010.

20. Ferrari PA, Barbiero A. Simulating ordinal data. Multivar Behav Res. 2012;47(4):566–89.

21. Lall R, Campbell MJ, Walters SJ, Morgan K, Co-operative MC. A review of ordinal regression models applied on health-related quality of life assessments. Stat Methods Med Res. 2002;11(1):49–67.

22. Selman CJ, Lee KJ, Mahar RK. Statistical analyses of ordinal outcomes in randomised controlled trials: protocol for a scoping review [Internet]. arXiv; 2022 [cited 2023 Mar 19]. Available from: http://arxiv.org/abs/2208.06154

23. Hoffman DL, Franke GR. Correspondence analysis: graphical representation of categorical data in marketing research. J Mark Res. 1986;23(3):213–27.

24. Tutz G. Regression for Categorical Data [Internet]. Cambridge. Cambridge Series in Statistical and Pobabilistic Mathematics; Available from: https://www.cambridge.org/core/books/abs/regression-for-categorical-data/ordinal-response-models/5ECEEA8DAC1730F9CF2DA20B7C3350C3

25. Van Mechelen W, Hlobil H, Kemper HC. Incidence, severity, aetiology and prevention of sports injuries: a review of concepts. Sports Med. 1992;14:82–99.

26. Ananth CV, Kleinbaum DG. Regression models for ordinal responses: a review of methods and applications. Int J Epidemiol. 1997;26(6):1323–33.

27. O'Connell AA. An illustration of multilevel models for ordinal response data. In: Data and context in statistics education: Towards an evidence-based society Proceedings of the Eighth International Conference on Teaching Statistics (ICOTS8). 2010.

28. Zhang MY, Mlynash M, Sainani KL, Albers GW, Lansberg MG. Ordinal prediction model of 90-day modified rankin scale in ischemic stroke. Front Neurol. 2021;12:727171.

29. Martinez G. Sport Performance Analysis. 2020 [cited 2023 Feb 22]. What is Performance Analysis in Sport? Available from: https://www.sportperformanceanalysis.com/article/what-is-performance-analysis-in-sport

30. O'Donoghue P, Holmes L. Data Analysis in Sport. London: Routledge; 2014. 278 p.

31. Phillips LH. Sports injury incidence. Br J Sports Med. 2000 Apr 1;34(2):133–6.

32. Finch C, Cook J. Categorising sports injuries in epidemiological studies: The subsequent injury categorisation (SIC) model to address multiple, recurrent and exacerbation of injuries. Br J Sports Med. 2013 Mar 16;48.

33. Fuller CW, Ekstrand J, Junge A, Andersen TE, Bahr R, Dvorak J, et al. Consensus statement on injury definitions and data collection procedures in studies of football (soccer) injuries. Br J Sports Med. 2006 Mar 1;40(3):193–201.

34. Clarsen B, Myklebust G, Bahr R. Development and validation of a new method for the registration of overuse injuries in sports injury epidemiology: the Oslo Sports Trauma Research Centre (OSTRC) Overuse Injury Questionnaire. Br J Sports Med. 2013 May;47(8):495–502.

35. Franke TPC, de Vet HCW, Huisstede BMA. Minimally important change and smallest detectable change of the OSTRC questionnaire in half- and full-marathon runners. Scand J Med Sci Sports. 2021;31(5):1048–58.

36. Esteve E, Clausen MB, Rathleff MS, Vicens-Bordas J, Casals M, Palahí-Alcàcer A, et al. Prevalence and severity of groin problems in Spanish football: A prospective study beyond the time-loss approach. Scand J Med Sci Sports. 2020 May;30(5):914–21.

37. Inoue A, dos Santos Bunn P, do Carmo EC, Lattari E, da Silva EB. Internal Training Load Perceived by Athletes and Planned by Coaches: A Systematic Review and Meta-Analysis. Sports Med - Open. 2022 Mar 4;8(1):35.

38. Internal Load; Judging the Stress of your Sport | Sparta Science [Internet]. [cited 2023 May 23]. Available from: https://spartascience.com/blog/internal-load-judging-the-stress-of-your-sport

39. Falk Neto JH, Tibana RA, de Sousa NMF, Prestes J, Voltarelli FA, Kennedy MD. Session Rating of Perceived Exertion Is a Superior Method to Monitor Internal Training Loads of Functional Fitness Training Sessions Performed at Different Intensities When Compared to Training Impulse. Front Physiol [Internet]. 2020 [cited 2023 May 23];11. Available from: https://www.frontiersin.org/articles/10.3389/fphys.2020.00919

40. Kerr ZY, Lynall RC, Roos KG, Dalton SL, Djoko A, Dompier TP. Descriptive Epidemiology of Non–Time-Loss Injuries in Collegiate and High School Student-Athletes. J Athl Train. 2017 May 1;52(5):446–56.

41. Hägglund M, Waldén M, Ekstrand J. Injury recurrence is lower at the highest professional football level than at national and amateur levels: does sports medicine and sports physiotherapy deliver? Br J Sports Med. 2016 Jun;50(12):751–8.

42. Grant MJ, Booth A. A typology of reviews: an analysis of 14 review types and associated methodologies. Health Inf Libr J. 2009;26(2):91–108.

43. Alqahtany A, Al-Ghamdi H. Systematic Literature Review: Some Examples. 2019.

44. Evans D. Hierarchy of evidence: a framework for ranking evidence evaluating healthcare interventions. J Clin Nurs. 2003;12(1):77–84.

45. Peters MDJ, Marnie C, Tricco AC, Pollock D, Munn Z, Alexander L, et al. Updated methodological guidance for the conduct of scoping reviews. JBI Evid Synth. 2020 Oct;18(10):2119.

46. Cobo E, Cortes J, Ribera JM, Cardellach F, Selva-O'Callaghan A, Kostov B, et al. Effect of using reporting guidelines during peer review on quality of final manuscripts submitted to a biomedical journal: masked randomised trial. BMJ. 2011 Nov 22;343(nov22 2):d6783–d6783.

47. Borg DN, Barnett AG, Caldwell AR, White NM, Stewart IB. The bias for statistical significance in sport and exercise medicine. J Sci Med Sport [Internet]. 2023 Mar 7 [cited 2023 Mar 28]; Available from: https://www.sciencedirect.com/science/article/pii/S1440244023000403

48. Bennett C, Manuel DG. Reporting guidelines for modelling studies. BMC Med Res Methodol. 2012 Dec;12(1):168.

49. Altman DG, Simera I, Hoey J, Moher D, Schulz K. EQUATOR: reporting guidelines for health research. Open Med. 2008 Apr 28;2(2):e49–50.

50. Minkow D. The Evidence-Based Medicine Pyramid! [Internet]. Students 4 Best Evidence. 2014 [cited 2023 Mar 26]. Available from: https://s4be.cochrane.org/blog/2014/04/29/the-evidence-based-medicine-pyramid/

51. Munn Z, Peters MD, Stern C, Tufanaru C, McArthur A, Aromataris E. Systematic review or scoping review? Guidance for authors when choosing between a systematic or scoping review approach. BMC Med Res Methodol. 2018;18:1–7.

52. Bürkner PC, Vuorre M. Ordinal regression models in psychology: A tutorial. Adv Methods Pract Psychol Sci. 2019;2(1):77–101.

53. Lee E. Brant Test in R [Internet]. evangelinelee. 2019 [cited 2023 Apr 27]. Available from: https://medium.com/evangelinelee/brant-test-for-proportional-odds-in-r-b0b373a93aa2

54. Schlarmann J, Galatsch M. Regressionsmodelle für ordinale Zielvariablen. GMS Med Inform Biom Epidemiol. 2014 Mar 1;10.

55. Christensen RHB. Analysis of ordinal data with cumulative link models — estimation with the R-package ordinal.

56. Hedeker D. A mixed-effects multinomial logistic regression model. Stat Med. 2003 May 15;22(9):1433–46.

57. McCullagh P, Nelder JA. Binary data. In: Generalized linear models. Springer; 1989. p. 98–148.

58. Cole SR, Ananth CV. Regression models for unconstrained, partially or fully constrained continuation odds ratios. Int J Epidemiol. 2001;30(6):1379–82.

59. Samejima F. Graded Response Model. In: van der Linden WJ, Hambleton RK, editors. Handbook of Modern Item Response Theory [Internet]. New York, NY: Springer; 1997 [cited 2023 Jun 26]. p. 85–100. Available from: https://doi.org/10.1007/978-1-4757-2691-6_5

60. Muraki E. A Generalized Partial Credit Model: Application of an Em Algorithm. ETS Res Rep Ser. 1992;1992(1):i–30.

61. Masters GN. Partial Credit Model. In: Handbook of Item Response Theory. Chapman and Hall/CRC; 2016.

62. Heck RH, Thomas SL, Tabata LN. Multilevel and longitudinal modeling with IBM SPSS. Routledge; 2013.

63. Bender R, Benner A. Calculating ordinal regression models in SAS and S-Plus. Biom J J Math Methods Biosci. 2000;42(6):677–99.

64. Long JS, Freese J. Regression models for categorical dependent variables using Stata. Vol. 7. Stata press; 2006.

65. Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, et al. PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checklist and Explanation. Ann Intern Med. 2018 Oct 2;169(7):467–73.

66. Casado-Aranda LA, Sánchez-Fernández J, Viedma-del-Jesús MI. Analysis of the scientific production of the effect of COVID-19 on the environment: A bibliometric study. Environ Res. 2021 Feb;193:110416.

67. Evans AB, Blackwell J, Dolan P, Fahlén J, Hoekman R, Lenneis V, et al. Sport in the face of the COVID-19 pandemic: towards an agenda for research in the sociology of sport. Eur J Sport Soc. 2020 Apr 2;17(2):85–95.

68. Horbach SPJM. Pandemic publishing: Medical journals strongly speed up their publication process for COVID-19. Quant Sci Stud. 2020 Aug;1(3):1056–67.

69. Vancini RL, Andrade MS, Viana RB, Nikolaidis PT, Knechtle B, Campanharo CRV, et al. Physical exercise and COVID-19 pandemic in PubMed: Two months of dynamics and one year of original scientific production. Sports Med Health Sci. 2021 Jun 1;3(2):80–92.

70. Casals M, Girabent-Farrés M, Carrasco JL. Methodological Quality and Reporting of Generalized Linear Mixed Models in Clinical Medicine (2000–2012): A Systematic Review. PLOS ONE. 2014 Nov 18;9(11):e112653.

71. Women in Sports Data [Internet]. [cited 2023 Jun 13]. Available from: https://www.womeninsportsdata.org/

72. Lamprecht AL, Garcia L, Kuzak M, Martinez C, Arcila R, Martin Del Pico E, et al. Towards FAIR principles for research software. Data Sci. 2020 Jan 1;3(1):37–59.

73. Nielsen RO, Simonsen NS, Casals M, Stamatakis E, Mansournia MA. Methods matter and the "too much, too soon" theory (part 2): what is the goal of your sports injury research? Are you describing, predicting or drawing a causal inference? Br J Sports Med. 2020 Nov;54(22):1307–9.

74. Losciale JM, Bullock GS, Collins GS, Arundale AJH, Hughes T, Arden NK, et al. Description, Prediction, and Causation in Sport and Exercise Medicine Research: Resolving the Confusion to Improve Research Quality and Patient Outcomes. J Orthop Sports Phys Ther. 2023 Jul;0(7):1–7.

75. Yam DR, Lopez MJ. What was lost? A causal estimate of fourth down behavior in the National Football League. J Sports Anal. 2019 Jan 1;5(3):153–67.

76. Yam D, Lopez M. Quantifying the Causal Effects of Conservative Fourth Down Decision Making in the National Football League [Internet]. Rochester, NY; 2018 [cited 2023 Jun 9]. Available from: https://papers.ssrn.com/abstract=3114242

77. Benz LS, Lopez MJ. Estimating the change in soccer's home advantage during the Covid-19 pandemic using bivariate Poisson regression. AStA Adv Stat Anal. 2023 Mar 1;107(1):205–32.

78. O'Connell AA, Liu X. Model Diagnostics for Proportional and Partial Proportional Odds Models. J Mod Appl Stat Methods. 2011 May 1;10(1):139–75.

79. Arı E. PARALEL LINES ASSUMPTION IN ORDINAL LOGISIC REGRESSION AND ANALYSES APPROACHES. Int Interdiscip J Sci Res. 2014 Jan 1;1:8–23.

80. Plonsky L, Ghanbar H. Multiple Regression in L2 Research: A Methodological Synthesis and Guide to Interpreting R^2 Values. Mod Lang J. 2018 Dec;102(4):713–31.

81. Clatworthy Jane, Buick Deanna, Hankins Matthew, Weinman John, Horne Robert. The use and reporting of cluster analysis in health psychology: A review. Br J Health Psychol. 2005 Sep;10(3):329–58.

82. (PDF) THE RATIONALES OF STATISTICAL MODELLING IN EDUCATION RESEARCH FROM A MATHEMATICAL MODELLING PERSPECTIVE [Internet]. [cited 2023 Jun 6]. Available from: https://www.researchgate.net/publication/353212471_THE_RATIONALES_OF_STATISTICAL_MODELLING_IN_EDUCATION_RESEARCH_FROM_A_MATHEMATICAL_MODELLING_PERSPECTIVE

83. Zapata-Cardona L, Martínez-Castro CA. Statistical modeling in teacher education. Math Think Learn. 2023 Jan 2;25(1):64–78.

84. Henley SS, Golden RM, Kashner TM. Statistical modeling methods: challenges and strategies. Biostat Epidemiol. 2020 Jan 1;4(1):105–39.

85. SCORE Sports Data Repository [Internet]. [cited 2023 Jun 9]. Available from: https://data.scorenetwork.org/

86. Thorborg K, Branci S, Nielsen MP, Langelund MT, Hölmich P. Copenhagen five-second squeeze: a valid indicator of sports-related hip and groin function. Br J Sports Med. 2017 Apr 1;51(7):594–9.

87. Thomeé R. A Comprehensive Treatment Approach for Patellofemoral Pain Syndrome in Young Women. Phys Ther. 1997 Dec 1;77(12):1690–703.

88. Drubay D, Collette L, Paoletti X. Proportional odds assumption for modeling longitudinal ordinal multiple toxicity outcomes in dose finding studies of targeted agents: A pooled analysis of 54 studies. Contemp Clin Trials Commun. 2020 Mar 1;17:100529.

89. Ordinal Logistic Regression | R Data Analysis Examples [Internet]. [cited 2023 Jun 9]. Available from: https://stats.oarc.ucla.edu/r/dae/ordinal-logistic-regression/

Appendix

***Table 23**: Extended summary table for model 3 (Mixed Effects Proportional Odds Model of HAGOS, Sport/Rec subscale scores between players reporting and player not reporting groin problems) and model 4 (Mixed-Effects POM of HAGOS, Sport/Rec subscale scores between players reporting groin problems with and without time-loss).*

Summary table for the model 3

	NRS		
Predictors	*Odds Ratios*	*CI*	*p*
stop|attention	0.00	0.00 – 0.00	**<0.001**
attention|go	0.03	0.02 – 0.06	**<0.001**
Groin Problem Prevalence: Yes, Groin Problem	0.02	0.02 – 0.04	**<0.001**
factor(week)10	0.88	0.53 – 1.47	0.626
factor(week)14	0.71	0.41 – 1.21	0.204
factor(week)18	1.12	0.61 – 2.04	0.714
factor(week)22	2.01	1.01 – 4.00	**0.046**
factor(week)26	1.83	0.89 – 3.74	0.099
factor(week)30	2.13	0.98 – 4.63	0.055
factor(week)34	2.16	0.94 – 4.99	0.070
factor(week)38	1.40	0.58 – 3.39	0.457
factor(week)42	2.41	0.91 – 6.37	0.076
factor(week)44	4.64	1.57 – 13.77	**0.006**
Random Effects			
σ^2	3.29		
$\tau_{00\ id.player}$	6.25		
$\tau_{11\ id.player.week}$	0.01		
$\rho_{01\ id.player}$	-0.53		
ICC	0.66		
$N_{id.player}$	416		
Observations	3440		
Marginal R^2 / Conditional R^2	0.178 / 0.716		

Summary table for the model 4

	NRS		
Predictors	*Odds Ratios*	*CI*	*p*
stop|attention	0.11	0.04 – 0.28	**<0.001**
attention|go	2.49	0.98 – 6.33	0.056
TL_vs_No_TL_GProblem: No TL Groin Problem	1.29	0.56 – 2.98	0.545
week	0.97	0.96 – 0.99	**0.005**
Random Effects			
σ^2	3.29		
$\tau_{00\ id.player}$	4.91		
$\tau_{00\ week}$	0.00		
N_{week}	11		
$N_{id.player}$	183		
Observations	518		
Marginal R^2 / Conditional R^2	0.035 / NA		

In the Statistical Methods section of the Mixed-Effects POM, we have presented one plausible alternative for testing the assumption of proportional odds. However, another plausible alternative would be to visually inspect the proportionality of the odds ratios on a variable level. This is done by plotting the observed mean of the covariate against the expected mean implied by the proportional odds model. If the proportional odds assumption holds, the symbols representing each set of categories of the dependent variable should remain similarly distant for each predictor variable. Deviations from this pattern would indicate a violation of the assumption.

Table 24: *Chart visualizations of the alternative testing for the proportional odds assumption, plotting the observed mean of the covariate against the expected mean implied by model 3 (Mixed Effects Proportional Odds Model of HAGOS, Sport/Rec subscale scores between players reporting and player not reporting groin problems) and model 4 (Mixed-Effects POM of HAGOS, Sport/Rec subscale scores between players reporting groin problems with and without time-loss).*

Chart for the alternative testing of the proportional odds assumption for model 3

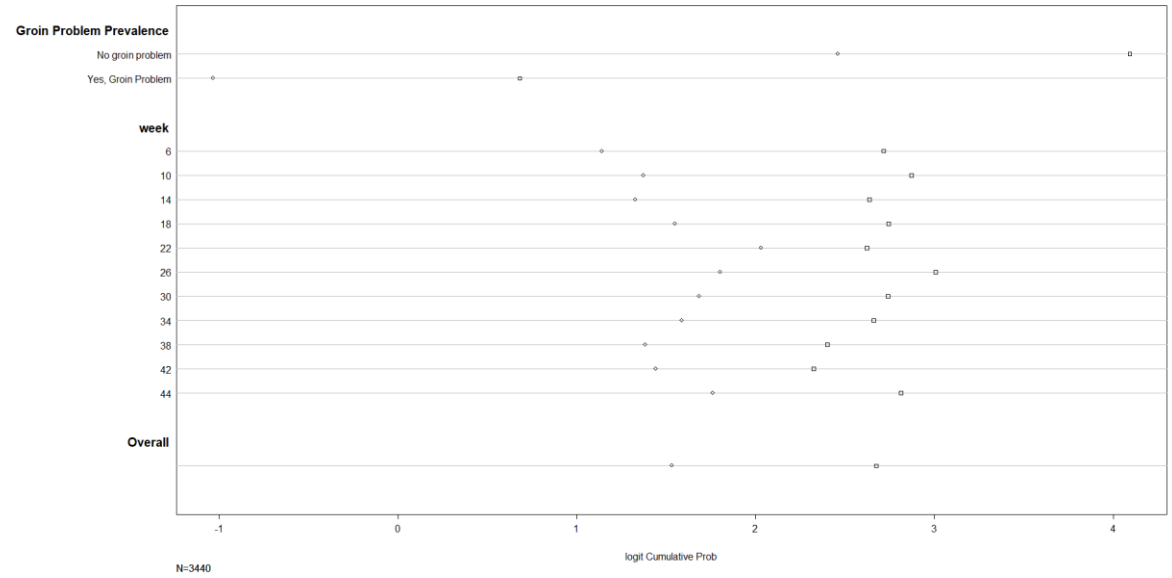
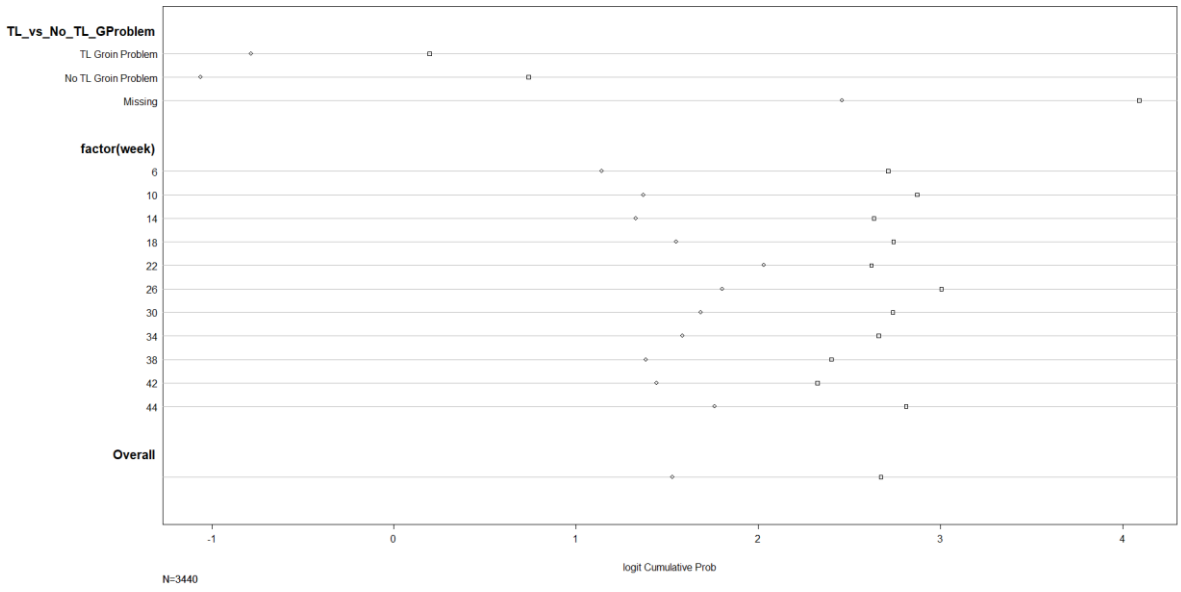


Chart for the alternative testing of the proportional odds assumption for model 4



The mixed effects proportional odds model offers a valuable approach for analyzing the significance of differences in pain monitoring. In this context, we present below the details of how to assess these differences. Specifically, we focus on the following two comparisons: 1) between players who report groin problems and those who do not, and 2) between players who report groin problems with and without time loss. To test if differences are significant, the package *lsmeans* is used.

Model 3 Output function *lsmeans*:

```
 groin.problem.prev lsmean    SE  df asymp.LCL asymp.UCL
 No groin problem     4.61 0.198 Inf     4.226      5.00
 Yes, Groin Problem   1.05 0.200 Inf     0.653      1.44

Confidence level used: 0.95

$contrasts
 contrast                              estimate    SE  df z.ratio p.value
 No groin problem - Yes, Groin Problem     3.57 0.168 Inf  21.258  <.0001
```

From this output we can observe that there are significant differences in pain monitoring level between players who report groin problems, and those who don't report groin problems. This is evident from the p-value that is lower to the significance value (0.05).

Model 4 Output function *lsmeans*:

```
 tl.vs.no.tl.gproblem lsmean    SE  df asymp.LCL asymp.UCL
 TL Groin Problem     0.0207 0.435 Inf    -0.833     0.874
 No TL Groin Problem  0.2783 0.215 Inf    -0.143     0.699

Confidence level used: 0.95

$contrasts
 contrast                              estimate    SE  df z.ratio p.value
 TL Groin Problem - No TL Groin Problem   -0.258 0.425 Inf  -0.606  0.5446
```

From this output, however, we can observe that there are no significant differences in pain monitoring level between players reporting groin problems with time loss, and those without time loss. This is evident from the p-value that is greater to the significance value (0.05).