

## ANKLE INJURY PREVENTION IN SOCCER USING MACHINE LEARNING: BIBLIOMETRIC ANALYSIS

### PREVENCIÓN DE LESIONES DE TOBILLO EN FÚTBOL MEDIANTE APRENDIZAJE AUTOMÁTICO: ANÁLISIS BIBLIOMÉTRICO

**Marlon Felipe Burbano-Fernández<sup>1</sup>** 

**Andres Felipe Villaquiran-Hurtado<sup>2</sup>** 

**Viviana Marcela Celis-Quinayas<sup>2</sup>** 

**Jeffry Alexander Hoyos-Quisobony<sup>2</sup>** 

**Nancy Janneth Molano-Tobar<sup>3</sup>** 

<sup>1</sup> Departamento de Turismo, Facultad de Ciencias Contables, Económicas y Administrativas, Universidad del Cauca, Colombia

<sup>2</sup> Departamento de Fisioterapia, Facultad de Ciencias de la Salud, Universidad del Cauca, Colombia

<sup>3</sup> Departamento de Educación Física, Recreación y Deporte, Facultad de Ciencias Naturales, Exactas y de la Educación, Universidad del Cauca, Colombia

#### Correspondence:

Andres Felipe Villaquiran-Hurtado  
[avillaquiran@unicauca.edu.co](mailto:avillaquiran@unicauca.edu.co)

#### Short title:

Bibliometry in Ankle Injury Prevention

#### How to cite this article:

Burbano-Fernández, M.F., Villaquiran-Hurtado, A. F., Celis-Quinayas, V. M., Hoyos-Quisobony, J., & Molano-Tobar, N. J. (2024). Ankle injury prevention in soccer using machine learning: bibliometric analysis. *Cultura, Ciencia y Deporte*, 19(61), 2119. <https://doi.org/10.12800/ccd.v19i61.2119>

Received: 19 October 2023 / Accepted: 10 April 2024

### Abstract

Bibliometric analysis seeks to evaluate through statistical methods the scientific activity on the lines and trends of research, the evolution of studies, the relationships between publications, journals and collaboration between researchers. The use of these studies can guide researchers on the evolution of research processes related to injury prevention in soccer, using machine learning. The aim of this study is to analyze the scientific activity related to machine learning in the prevention of ankle injuries in soccer. The present study presents three moments: Data capture, Analysis of the information based on software (Scimat, VosViewer, Use of Text mining with R), discussion and conclusions. As for the results, the evolution of the words and networks generated shows an increase in studies relating the words "sport", "ankle", "risk factors" and "technology" (mobile applications, computational methods, wireless communication). An evolution of research in terms of the use of machine learning in injury prevention, visualization of knowledge networks and support among researchers in recent years is evident, as well as the growth of publications and the increase of networks and interaction between words.

**Keywords:** Injury sports, bibliometrics, soccer, artificial intelligence, ankle.

### Resumen

El análisis bibliométrico busca evaluar a través de métodos estadísticos la actividad científica sobre las líneas y tendencias de investigación, la evolución de los estudios, las relaciones entre publicaciones, revistas y la colaboración entre investigadores. El uso de estos estudios puede orientar a los investigadores sobre la evolución de los procesos de investigación relacionados la prevención de lesiones en el fútbol, utilizando el aprendizaje automático. El objetivo de este estudio es analizar la actividad científica relacionada con el aprendizaje automático en la prevención de lesiones de tobillo en el fútbol. El presente estudio presenta tres momentos: Captura de datos, Análisis de la información basado en software (Scimat, VosViewer, Uso de Text mining con R), discusión y conclusiones. En cuanto a los resultados, la evolución de las palabras y las redes generadas muestra un aumento de los estudios que relacionan las palabras "deporte", "tobillo", "factores de riesgo" y "tecnología" (aplicaciones móviles, métodos computacionales, comunicación inalámbrica). Se evidencia una evolución de la investigación en cuanto al uso del aprendizaje automático en prevención de lesiones, visualización de redes de conocimiento y apoyo entre investigadores en los últimos años, así como el crecimiento de publicaciones y el aumento de redes e interacción entre palabras.

**Palabras clave:** Lesiones deportivas, bibliometría, fútbol, inteligencia artificial, tobillo.



This work is licensed under a [Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License](https://creativecommons.org/licenses/by-nc-sa/4.0/).

## Introduction

The importance of recognizing and differentiating epidemiologically the most common injury in soccer is one of the steps to develop future prevention and treatment approaches by health professionals (Guldbrandsen et al., 2019; Olmedilla-Zafra, Andreu-Álvarez, Ortín-Montero, et al., 2013). Among the injuries with the highest frequency in this sport (Olmedilla-Zafra, Andreu-Álvarez, Abenza-Cano, et al., 2013), are those affecting the ankle (Kolokotsios et al., 2021), being 80% ligament sprains, and between 16 and 40% acute trauma (Halabchi & Hassabi, 2020), in addition, it is estimated that ankle and foot injuries in competition range between three to nine injuries per 1000 hours of play, producing high costs to the health system and sports clubs (Feria-Arias et al., 2018).

The public health burden of ankle injuries and long-term effects such as osteoarthritis and chronic instability has generated that primary prevention through prophylactic interventions by detecting risk factors is one of the most used tools by coaches and health professionals (Kaminski et al., 2019). Although, both intrinsic and extrinsic risk factors have been established, so far there is no clear consensus on the matter and the existing evidence is limited (Delahunt & Remus, 2019). The above may be due to the use of explanatory causality models used and the limitation of linear analyses between variables, ignoring multifactorial interrelationships (Jauhiainen et al., 2021).

Given the above, the use of artificial intelligence (AI) has shown promising results in injury risk prediction (Ramkumar et al., 2022), with machine learning (ML) being one of the computational methods used in information processing by analyzing real-world datasets to establish relationships, recognize patterns, classify, predict, exchange data, and improve interconnectivity through incremental learning as data is input (Jauhiainen et al., 2021). Despite, finding current evidence of the use of AI, the strengths, limitations, and true applications are yet to be determined (Román-Belmonte et al., 2021).

One of the bases for the initiation of new research processes is the analysis of productivity, recognition of research activity, lines, and central trends (Adán et al., 2020). Among the strategies developed by researchers are bibliometric studies, which allow establishing an overview of maturity, growth, types of research, researchers related to the area, the most prominent journals, and the mapping of publications (Müller et al., 2018).

To demonstrate interconnections and research networks, the VOSviewer software is a program that allows for the analysis and establishment of relationships of large amounts of data (Shah et al., 2020), with extensive text mining functionality through the creation, visualization, and exploration of bibliometric maps (Arruda et al., 2022).

Although there is general information on the use of machine learning in the prevention of sports injuries and the use of bibliometric analysis in sports science, so far there is no research that evidences the monitoring of scientific publications, the evolution of studies, the relationships between publications, journals, collaboration between researchers and co-occurrence specifically in relation to ankle injuries in soccer. Therefore, the aim of this article is to provide a bibliometric approach that identifies and analyzes the scientific activity related to machine learning in the prevention of ankle injuries in soccer.

## Materials and Methods

This study attempted to answer the following research questions:

What are the most critical and influential channels (authors, journals, and articles) in machine learning for ankle injury prevention in soccer?

What are the most emerging research streams in machine learning studies for ankle injury prevention in soccer?

What directions will open new avenues for future research using machine learning for ankle injury prevention in soccer?

The present bibliometric study, meso level, presents three moments: Data capture, Software-based information analysis, Discussion and Conclusions.

### Data Capture

For data capture, a search was carried out in English in the Scopus and Web of Science (WoS) databases, Pubmed, Science Direct and SportDiscus. To perform the Boolean chain, four groups of search words were selected that correspond to the topics addressed in this article. Theme one corresponds to words associated with predictive technologies through algorithms, selecting the following search words: "artificial intelligence", "Machine Intelligence", "Machine learning", "Neural network" and "learning algorithms". Topic two deals with injuries in a part of the human body to be studied, the following words are then selected: "ankle sprain", "ankle", "ankle injury", "ankle fracture" and "ankle injuries". Theme three associates the search in terms of prevention, so the following words are used in the search: "prevention and control", "preventive therapy", "preventive measures", "prevention" and "Injury prevention". Finally, a sport is contextualized through the following words, "Soccer", "Futsal", "indoor soccer", "five-a-side soccer", "street soccer", "street football".

For the creation of the search string, you considered that each topic contains complementary information, therefore, the Boolean operator used is “OR” to indicate the union in the associated words. To associate the different topics, matching information is used, so the “AND” operator is used for this purpose.

### Use of Scientometric Software

Use of Scimat: since different databases were used to find the information, it was necessary to use a tool that would allow grouping the information to pre-process the data and visualize the data clearly. Scimat allows to read different types of files, so that it is possible to consolidate a bank of information from the databases. From this, it is possible to find the documents published for each year and the most cited articles by time (Lopez-Robles et al., 2021).

Use of VosViewer (VOSviewer - Visualizing Scientific Landscapes, n. d.): To find the topics that make up the scientific literature addressed, the bibliographic information visualization software VOSviewer was used, which, based on the density of the keywords of the documents found, makes a mapping that is translated into heat charts, which represent the frequency of repetition of the documents and the proximity to others, and graph charts, which represent the relationships between different topics as well as the creation of groupings based on the relationships and proximity of the keywords.

Use of text mining with R (R: The R Project for Statistical Computing, n. d.): Text mining was used, where a word cloud graph (Rpubs, n. d.) was constructed from the abstracts downloaded with the different articles, which represents, according to size, the words with the highest frequency of use, and according to centrality, represents how close they are in relation to the other words. A hierarchical clustering graph of the H-Clust words was also constructed (Murtagh & Legendre, 2014).

Once the results obtained from the data capture and analysis with scientometric software are considered, we proceed to examine the findings, assessing trends and emphasizing the literature in greater detail, performing a systematic mapping, to find from articles of interest, trends in the field of machine learning for the prevention of ankle injuries in soccer.

## Results

Based on the results, 144 articles were found, of which three presented duplicities, so the Scopus and Web Of Science databases were considered to continue with the analysis.

### Results With Scimat

A quick establishment of groups of words was made by joining those in singular and plural, and after a manual review, it was not found to be necessary to join other groups of terms.

From the documents and using Scimat software, the information summarizing the publications per year, the most cited articles, and the number of publications per journal was constructed. For this purpose, a total of five periods were established for analysis. Period one corresponds to publications up to the year 2010; period two, corresponds from the year 2011 to the year 2015; period three contains the publications from the year 2016 to the year 2020; period four has the publications for the year 2021 and period five covers the publications for the year 2022, the results can be seen in Figure 1.

**Figure 1**  
*Publications found by year and period*

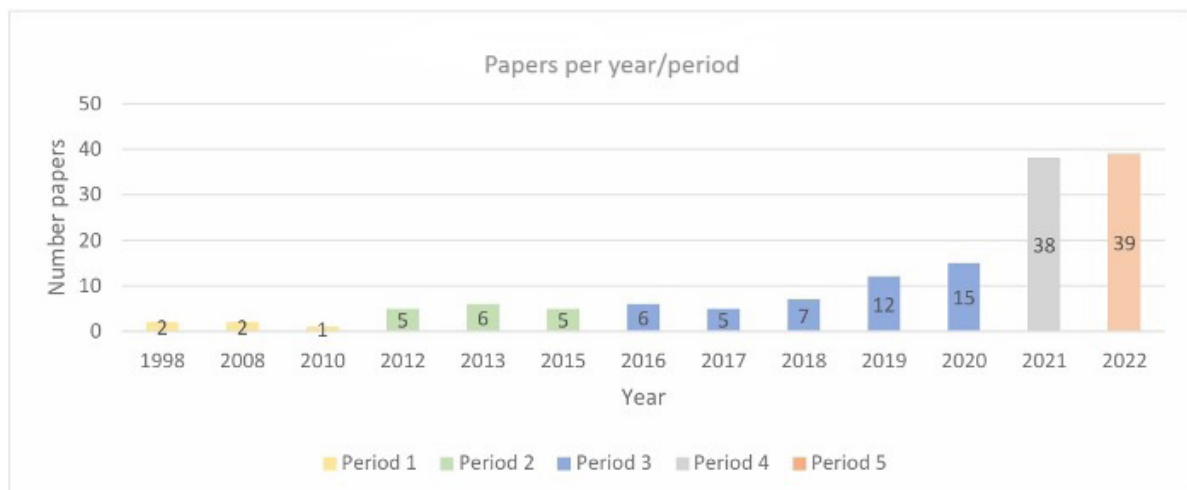


Table 1 shows the most cited articles, disaggregated by year of publication.

**Table 1**  
*Most cited articles from 1998 to 2022*

Year	Title	Citations
2012	Foot plantar pressure measurement system: A review	492
2008	Head and neck position sense	125
2008	Haptic rendering: Foundations, algorithms, and	117
2015	Associations Between Measures of Balance and Lower-	103
2015	MEMS sensor technologies for human centred	99
2018	Photobiomodulation therapy for the improvement of	95
2013	Correlation of clinical and magnetic resonance imaging	78
2012	Virtual reality in medicine	71
2016	A comparative review of footwear-based wearable	69
2015	A systematic review to determine reliability and	64
2019	Principles of Motor Learning to Support Neuroplasticity	59
2018	Wearable monitoring devices for biomechanical risk	56
2020	Measuring biomechanical loads in team sports“from	45
1998	Factors contributing to function of the knee joint after	43
2013	Postural sway in volleyball players	38
2017	Identifying poses of safe and productive masons using	34
2021	Multidimensional Ground Reaction Forces and	32
2017	Effect of bike-fit in the perception of comfort, fatigue	29
2013	Human Motion Simulation: Predictive Dynamics	26
2012	Motor imagery and 'placebo-racket effects' in tennis	26
2018	A systematic review of game technologies for pediatric	25
2020	Gait analysis “ Available platforms for outcome	24
2018	Allopurinol partially prevents disuse muscle atrophy in	22
2013	Design of human surrogates for the study of	21
2019	Wearables, biomechanical feedback, and human motor-	20
2021	Rehabilitation and return to sport assessment after	18
2012	Functional Movement Development Across the Life	18
2019	Sleep deprivation exacerbates concussive head injury	15
2016	Improved determination of dynamic balance using the	15
2020	Multifactorial individualised programme for hamstring	14
2019	Training rugby athletes with an external attentional	14
2019	Use of machine learning and wearable sensors to	13
2021	A machine-learning approach to measure the anterior	12
2021	New Machine Learning Approach for Detection of	12
2019	Rehabilitation Engineering: A perspective on the past	12
2019	Is a low Functional Movement Screen score	12
2019	Using Machine Learning to Predict Lower-Extremity	11
2021	Eccentric Strength Assessment of Hamstring Muscles	9
2021	Applications of pose estimation in human health and	9
2021	Biomechanical Response of the Lower Extremity to	9
2017	Automatic measurement of anthropometric dimensions	9
2021	A Data Science approach analysing the Impact of	8
2021	The future of in-field sports biomechanics: wearables	8
2021	Injury risk prediction of aerobics athletes based on big	8

Table 2 shows the journals with the number of published papers.



**Table 2**  
*Journals with the most published articles*

Journal	Documents	
Sensors	8	
Sports Medicine	5	
Applied Sciences (Switzerland)	5	
Frontiers in Physiology	4	
Sports Biomechanics	4	
Sensors (Switzerland)	4	
Sports Medicine - Open	3	
Orthopaedic Journal of Sports Medicine	3	
Scientific Reports	3	
International Journal of Environmental	3	
Thermology International	3	
Bioengineering	2	
International Journal of Sports Medicine	2	
Frontiers in Sports and Active Living	2	
Revista Brasileira de Medicina do Esporte	2	
Science and Medicine in Football	2	
Journal of Physical Education and Sport	2	
Arthroscopy - Journal of Arthroscopic and	2	
Frontiers in Psychology	2	
BMJ Open Sport and Exercise Medicine	2	
Critical Reviews in Biomedical Engineering	2	
Journal of Sports Sciences	2	
Journal of NeuroEngineering and	1	
Diagnostics	1	
Frontiers in Bioengineering and Biotechnology	1	
Life	1	
Clinical Journal of Sport Medicine	1	
Yiyong Shengwu Lixue/Journal of Medical	1	
Frontiers of Information Technology and	1	
IEEE Journal of Biomedical and Health	1	
Current Physical Medicine and Rehabilitation	1	
Biology of Sport	1	
Journal of Sensors	1	
Journal of Intelligent and Fuzzy Systems	1	
Contrast Media and Molecular Imaging	1	
Russian Journal of Biomechanics	1	
European Journal of Sport Science	1	
Mobile Information Systems	1	
2022 IEEE/CVF Winter Conference on	1	
Methods in Molecular Biology	1	
Journal of Clinical Orthopaedics and Trauma	1	
International Journal of Computer Science in	1	
International Journal of Sports Science and	1	
Frontiers in Bioscience - Landmark	1	
Journal of Orthopaedic Research	1	
Revista Pesquisa em Fisioterapia	1	
Physiological Measurement	1	
Information Systems	1	
2021 International Conference on Artificial	1	
Sports Orthopaedics and Traumatology	1	
Journal of Sport and Health Science	1	

Journal	Documents	
Sport Mont	1	
Journal of Military Medicine	1	
Journal of Strength and Conditioning Research	1	
Scientific Programming	1	
Footwear Science	1	
IEEE Transactions on Biomedical Engineering	1	
Complementary Therapies in Clinical Practice	1	
15th IEEE International Symposium on Medical	1	
Gait and Posture	1	
Journal of Science and Medicine in Sport	1	
Injury	1	
2020 IEEE International Conference on	1	
Team Sports Training: The Complexity Model	1	
International Journal of Sports Physiology and	1	
JMIR mHealth and uHealth	1	
Biocybernetics and Biomedical Engineering	1	
Journal of Experimental Orthopaedics	1	
Medical Engineering and Physics	1	
Journal of Coastal Research	1	
Medicine and Science in Sports and Exercise	1	
Progress in Brain Research	1	
Computers in Biology and Medicine	1	
Motor Control, Learning and Development:	1	
9th International Conference on Digital	1	
Lasers in Medical Science	1	
Automation in Construction	1	
PLoS ONE	1	
IET Computer Vision	1	
The Science of Gymnastics: Advanced	1	
Electronics (Switzerland)	1	
Journal of Biomechanical Engineering	1	
SpringerBriefs in Applied Sciences and	1	
10th International Conference on Ubiquitous	1	
Medicina dello Sport	1	
International Journal of Occupational	1	
Journal of Mobile Multimedia	1	
Kinesiology	1	
Human Movement Science	1	
Routledge Handbook of Motor Control and	1	
Human Motion Simulation: Predictive	1	
Psychology of Sport and Exercise	1	
Virtual Reality in Medicine	1	
The Science of Footwear	1	
Functional Movement Development Across	1	
Haptic Rendering: Foundations, Algorithms,	1	
Deutsche Zeitschrift fur Sportmedizin	1	
Journal of Bone and Joint Surgery	1	

In the analysis of the five periods established from 1998 to 2022, the study yielded the following results. Analysis Configuration:

Unit of analysis: Words (authorRole = true, sourceRole = true, addedRole = true)

Kind of network: Co-occurrence

Normalization measure: Equivalence index

Cluster algorithm: centers simples

Max cluster size: 6

Min cluster size: 1

Evolution measure: Jaccard index

Overlapping measure: Inclusion index

The overlapping map generated with Scimat shows that for period one there were three relevant words, but in the transition to period two, two words leave the map and only one continues, i.e. 33% of the words remain on the map. In period two, there are 17 relevant words, of which 16 are new, in the transition between period two and period three, four words leave the map and 13 remain in the map, while 76% of the words move to period three (Figure 2).

Meanwhile, in period three, there are 47 relevant words, of which 34 new words entered the map, and in the transition between period three and period four, 21 words are lost and 26, i.e. 79%, pass. In period four, 33 relevant words are observed, of which 7 new words enter, nine words leave in the transition to period five, while between period four and period five, 24 words are lost, i.e. 73%. Finally, in period five, there are 43 relevant words, 19 of which are new. Figure 1 shows the evolution of the words over the five periods, corroborating what was shown in the overlapping map.

**Figure 2**  
Map of overlapping and evolution of words generated with Scimat

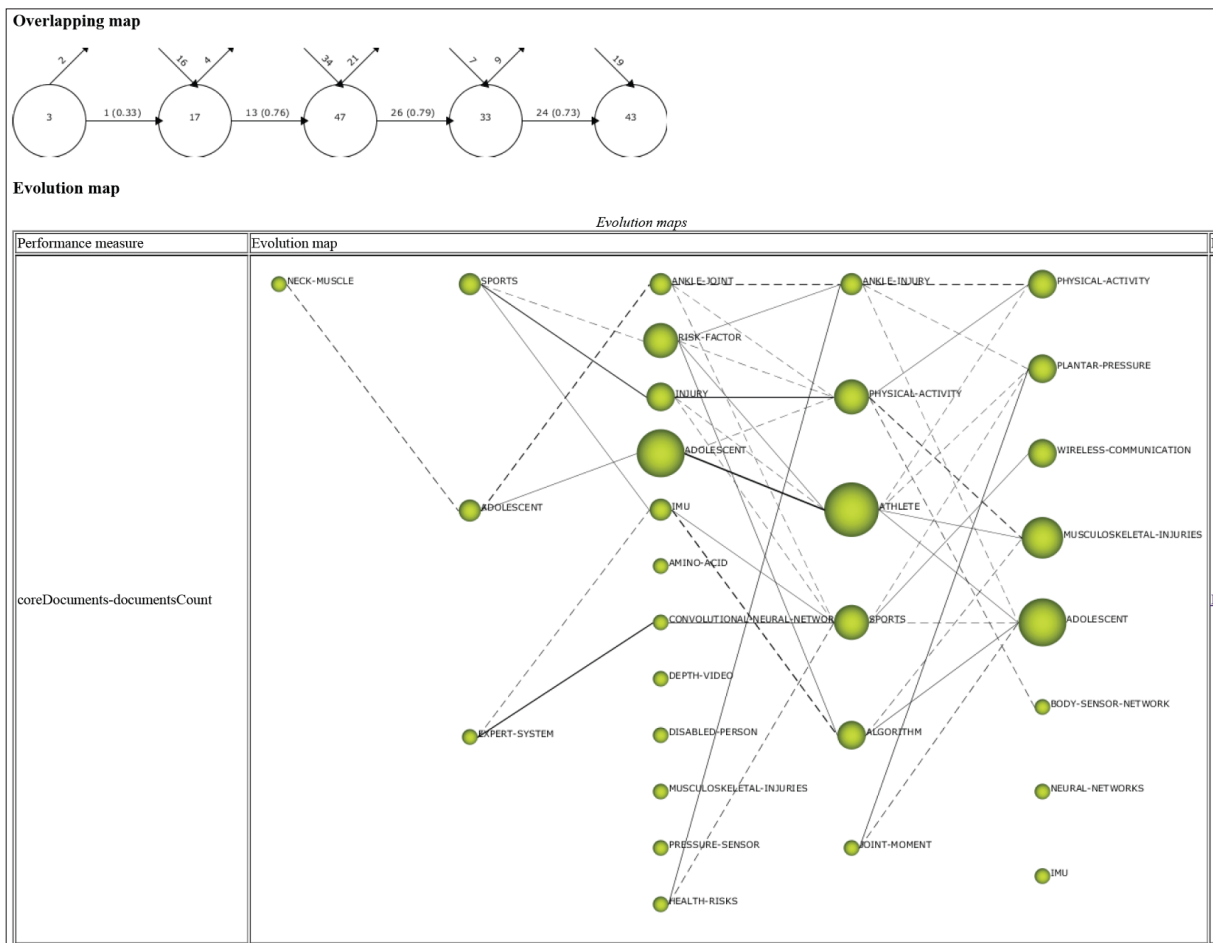


Figure 3 shows the strategic diagram (upper left figure) and the clustering networks associating the words (upper right figure and lower right and left figures) for period three (2016-2020), in this case only the clustering networks with more than two associated words are observed.

**Figure 3**  
Strategic diagram (upper left figure) and clustering networks for the period 2016-2022

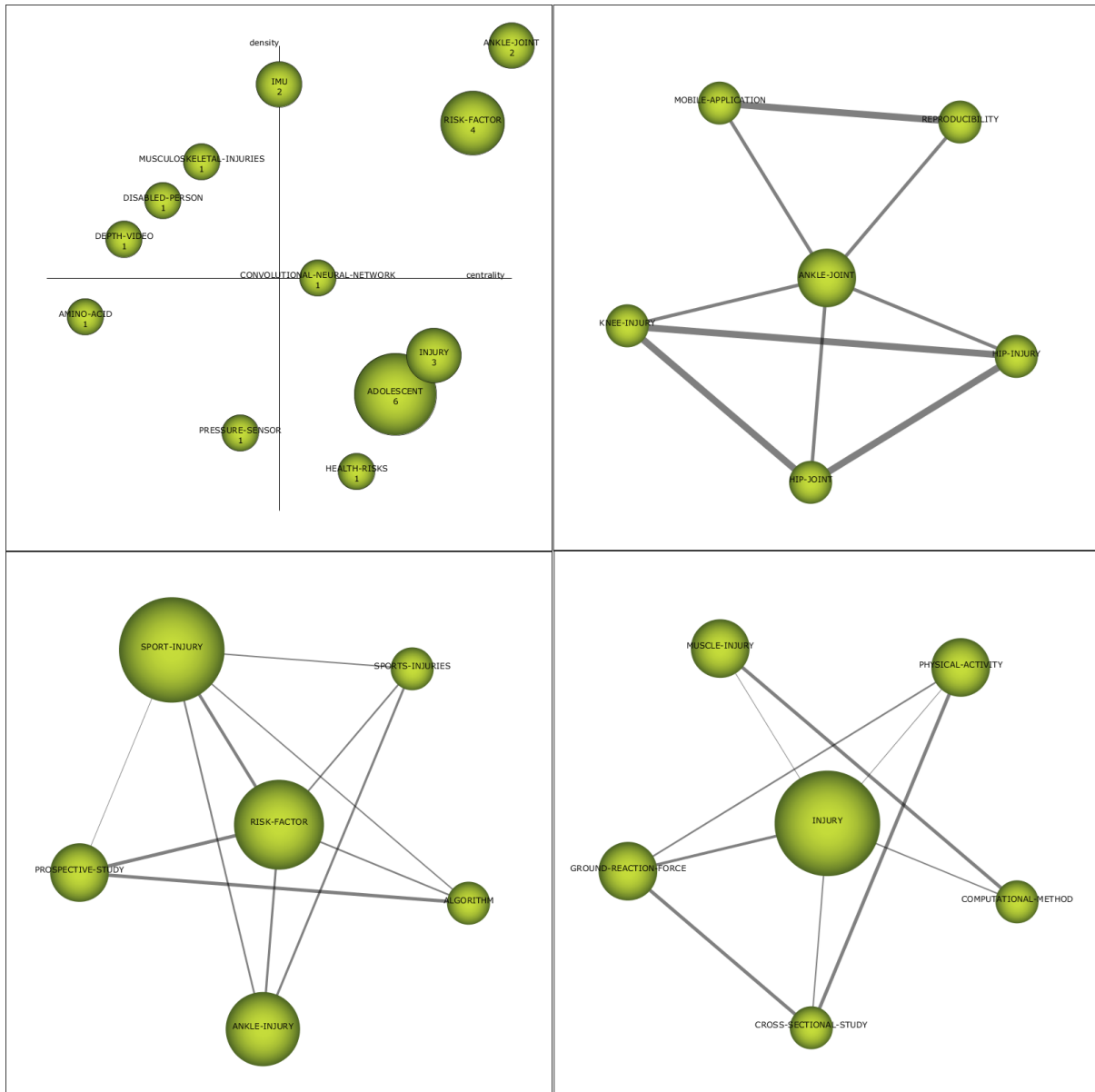
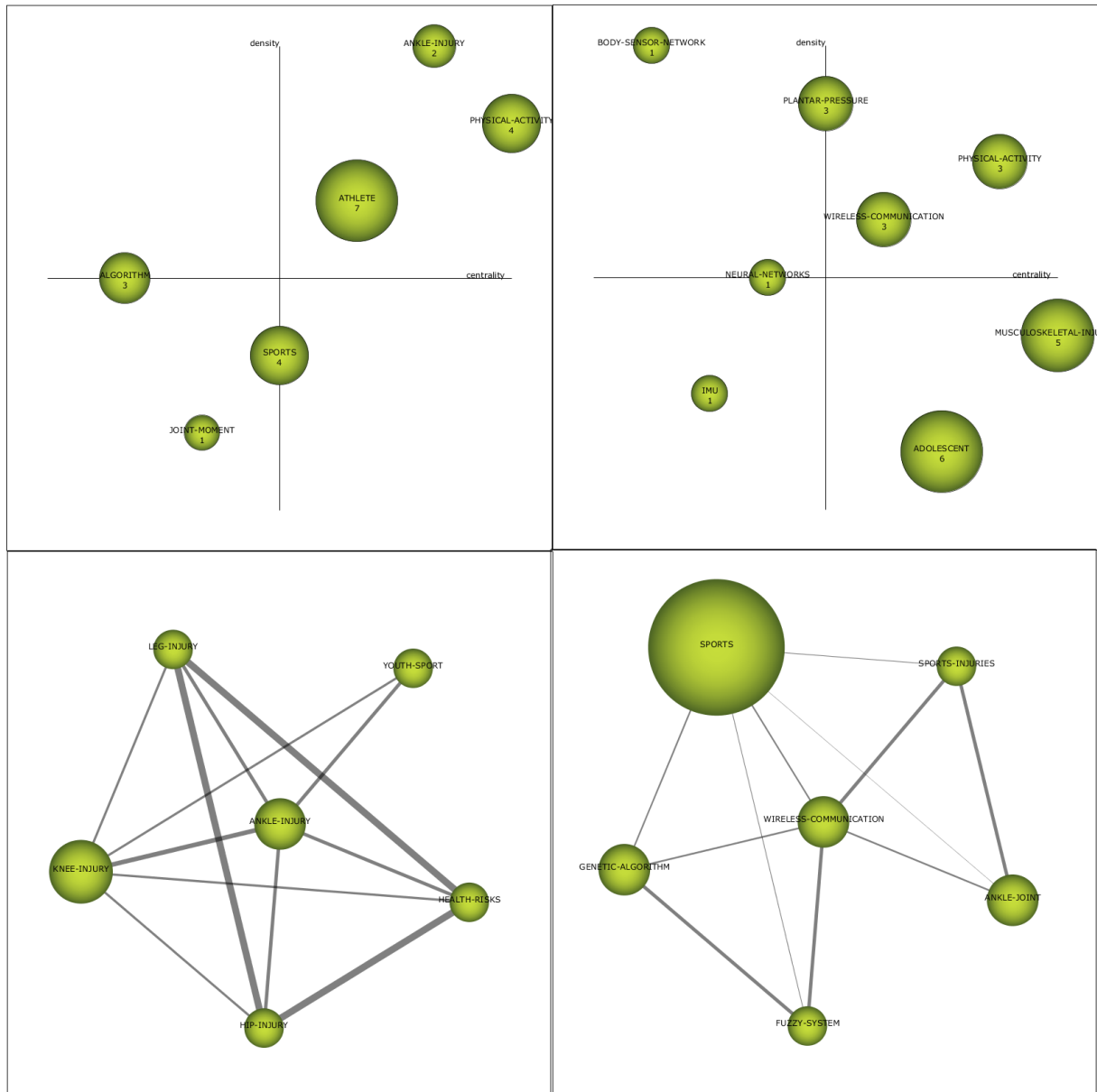


Figure 4 shows the strategic diagram (upper left figure) and the clustering networks that associate the words (upper right figure and lower right and left figures) for period four and five (2021- 2022), in this case only the clustering networks with more than two associated words are shown.



**Figure 4**  
 Strategic diagram (upper left figure) and clustering networks for the period 2021-2022

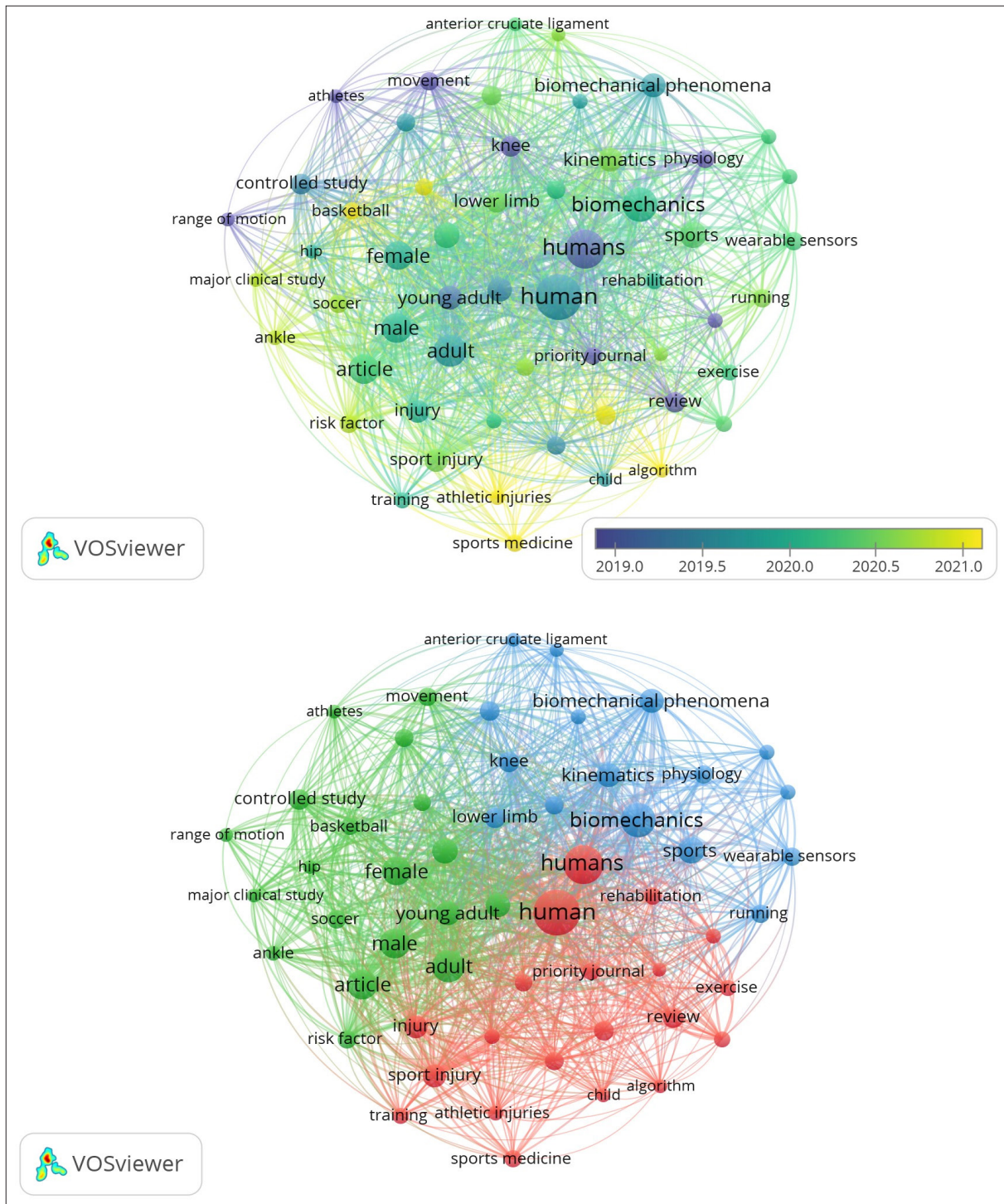


### Results With VosViewer

#### Results With VosViewer Based on the Keywords

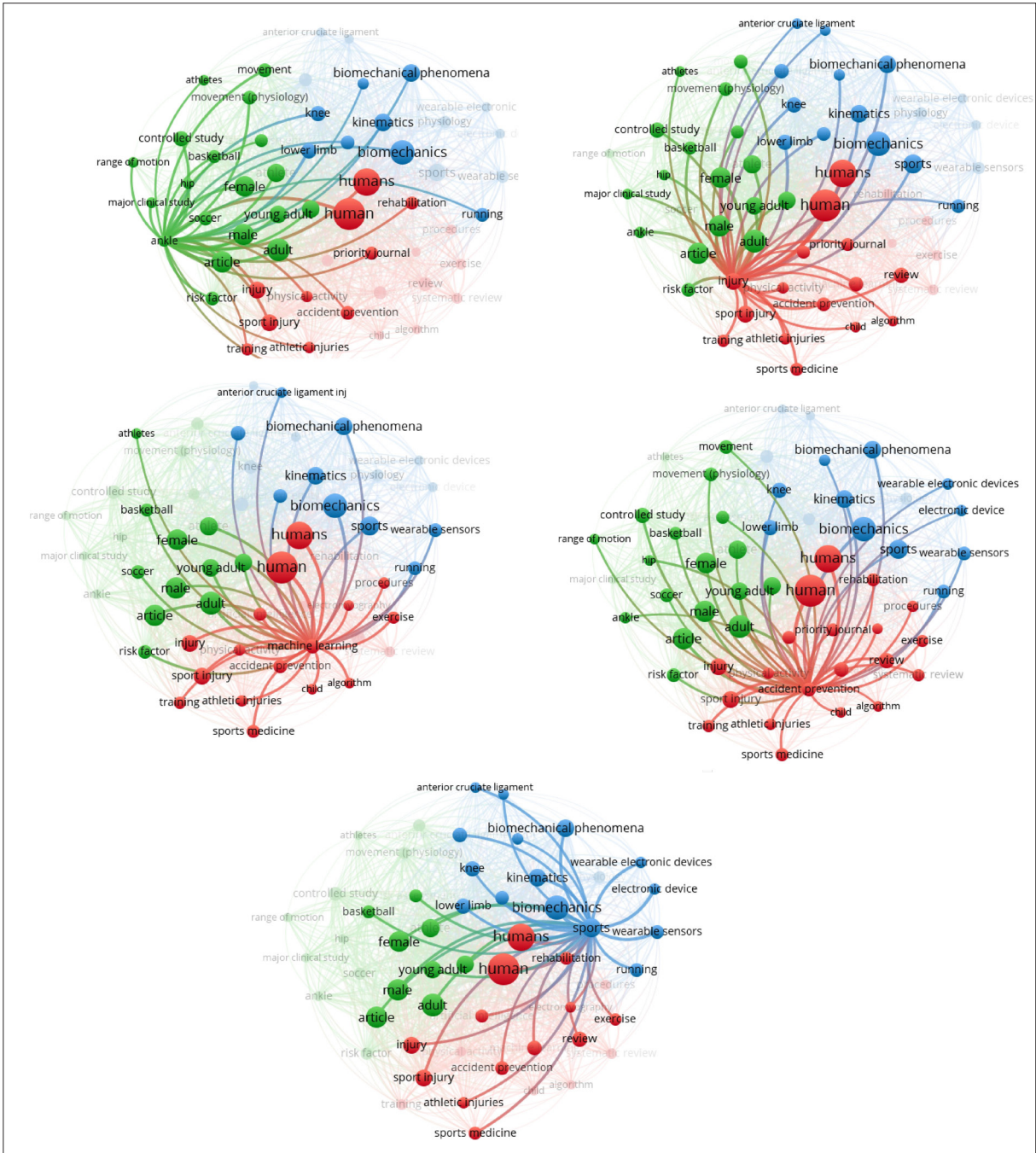
Figure 5 shows the associated words and their different relationships based on the publication date. Additionally, 3-word clusters (green, blue and red) can be observed. The green cluster revolves around gender and age of people, while the blue cluster revolves around biomechanics and the red cluster revolves around the word's human and humans.

**Figure 5**  
Words associated with date of publication, and Cluster of words generated



From the Boolean chains, the words “Ankle”, “Injury”, “Machine Learning”, “Prevention” and “Sport” are selected. The mentioned words are focused to visualize the relationship with other keywords (Figure 6).

**Figure 6**  
 Focus on the words “Ankle”, “Injury”, “Machine Learning”, “Prevention” and “Sport”

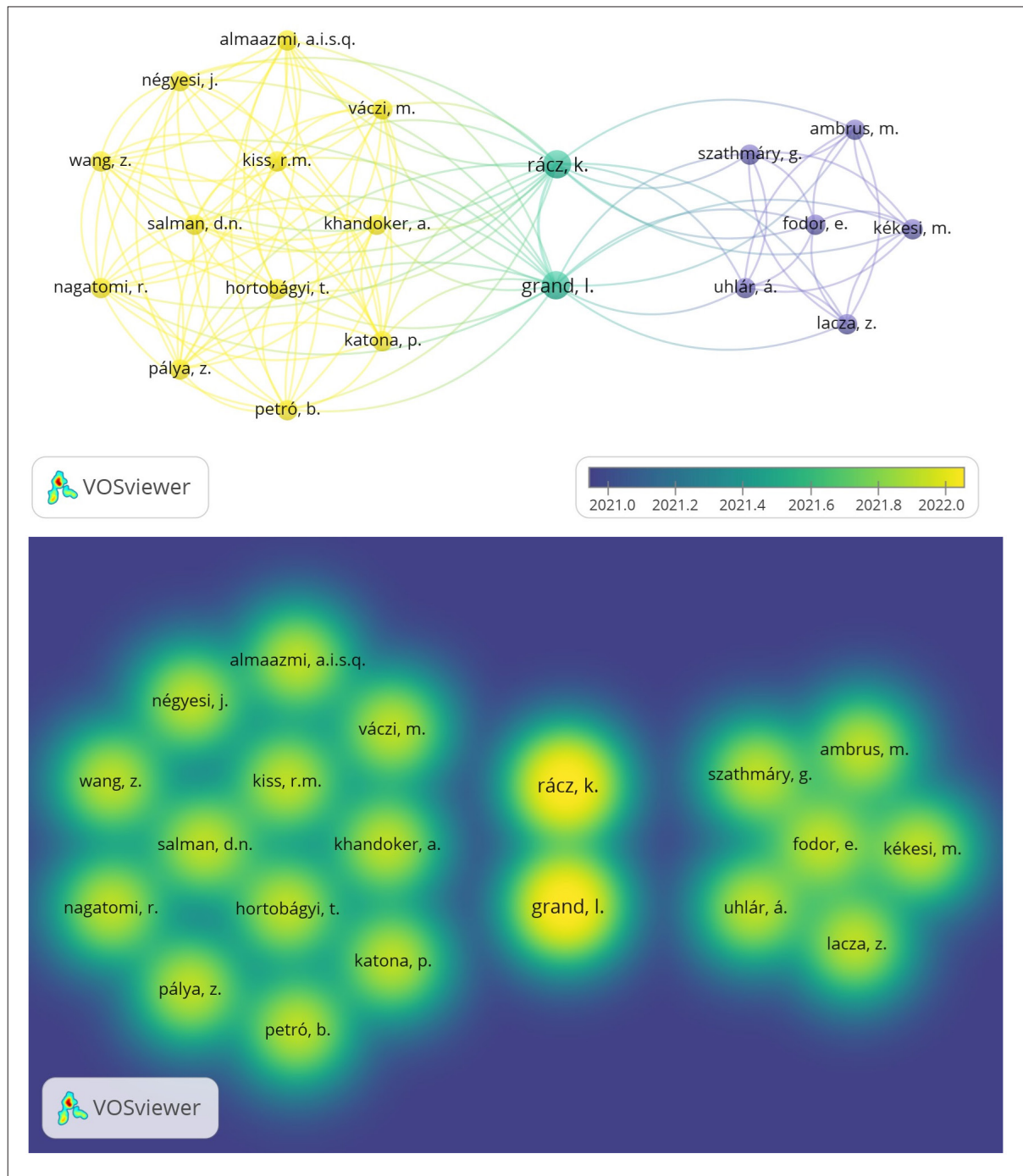


**Results with VosViewer based on the authors**

Figure 7 shows the results obtained based on the authors with the highest number of publications for the analyzed articles. In the upper part, the relationship of the authors is shown based on the publication dates, and the heat diagram based on the frequency of citations.



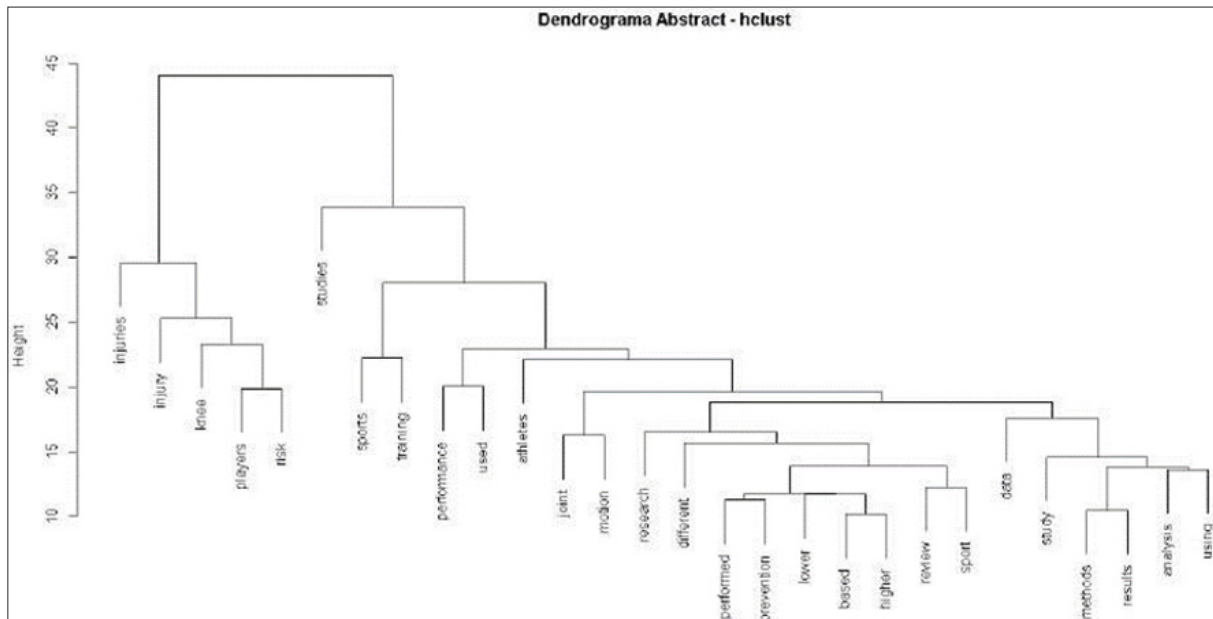
**Figure 7**  
 Clusters generated by Authors (top), authors in relation to publication dates (middle),  
 authors' citation frequency (bottom)



**Results with text analysis using R language**

Using the R programming language, text mining is performed on the abstracts of the articles. A word cloud diagram is generated from the frequency and centrality in the use of the words, where its size indicates the frequency of repetition of the word, and the centrality indicates the closest relationship with the surrounding words. This is shown in Figure 8.

**Figure 8**  
Word dendrogram generated from the abstracts



As can be seen in Figure 8, there is a relationship between the words “studies” and “injuries”, as well as a close interaction with the word “injuries” and “sports”.

## Discussion

Regarding the number of articles published on this topic, a growth in the number of publications from 2016 can be evidenced. In this regard, Llamas (2021) states that the use of technological tools in sport is increasing for the training and control of athletes, in terms of injury prediction, the use of mathematical models can provide relevant information on prevention strategies, however, it is necessary to continue creating and implementing specific computer applications for each sport. This may explain the increase in the research carried out, given that a greater number of studies are required to consolidate these tools as injury prediction instruments.

In reference to the most cited articles, the study “Foot Plantar Pressure Measurement System: A Review” by the authors Abdul Razak et al. (2012) is the article with the highest number of citations, which reviews the features of the sensors and systems for measuring the plantar pressure of the foot and the proposal of a wireless system to measure high pressure distributions under the foot. Wilzman et al. (2022) mention that the measurement of plantar pressure in runners can help predict overuse injuries with the assistance of machine learning, although further research should be done on the characteristics and relationships between healthy and injured subjects to make the model more interpretable.

Additionally, the study by Razak et al. (2021) presented as keywords: “plantar foot pressure”, “pressure sensor”, “wireless system”, such terms have been used in studies looking for predictive results using plantar pressure distributions for risk estimation and prevention strategies in overuse bone injuries (Wilzman et al., 2022), on the other hand, pressure sensors play an important role in the modeling of sports shoes to decrease the complex biomechanical and footwear interactions (Song et al., 2022). Plantar pressure systems have been used in different studies related to diagnosis, injury prevention and rehabilitation of the foot (Zhao et al., 2020), which may explain the references of this research.

The journal “Sensors” presents the highest number of published articles, followed by the journal “Sports Medicine”; the former specializes in presenting contributions of authors related experimental and theoretical results on sensor technology and its applications, while “Sports Medicine” publishes on topics related to injury prevention, exercise for health, training, nutrition, and drugs in sport. Advances in sensor technology and greater ease of availability and acquisition have led to an increase in research related to the evaluation and monitoring of athletes’ performance to help control the risk of injury (De Fazio et al., 2023), a situation that may be related to the increase in publications on sensors and sport.

Regarding the evolution of words and networks generated, this study shows an increase in studies that relate words such as: sport, ankle joint, risk factors and technology (mobile applications, computational methods, wireless communication). This trend may be due, as mentioned by Li and Kang (2022) that foot joints generate 10- 30% of all sports injuries,



furthermore, Kobayashi et al. (2016) stated that, although, there are different intrinsic risk factors related to ankle ligament injuries, no consensus has been reached so far on the predictive value that each of these has on the incidence of injury. Despite recognizing the importance of incorporating technology into ankle rehabilitation processes, and although advances in the use of AI in ankle and foot surgery are promising, new predictive models that present greater external validation and clinical benefit should continue to be explored through research (Gupta et al., 2023).

Regarding the evolution of words, the relationship between risk factors and prospective studies should be highlighted, as well as the relationship between ankle injury, knee injury and hip injury. Regarding the former, prospective studies have the advantage that the registration of data is more efficient, since at the beginning of the study the exposure factor may or may not have been present (Moreno & Gutiérrez, 2020), but the research effect has not yet been presented, requiring a follow-up period forward (Salazar et al., 2019), which may be easier for the control of measurements and the analysis of the different relationships between variables understanding their complexity (Pizarro & Masson, 2020). The above, may represent an advantage in the use of this type of methodologies for the researcher who seeks to predict ankle injuries from the analysis of the interaction between the different risk factors.

Regarding the relationships between ankle injury and the other joints of the lower limbs, research has focused its efforts on understanding the involvement and relationship that each of the joints has in different biomechanical actions that allow better decisions to be made in the training and rehabilitation of athletes through a more efficient and integrative clinical reasoning (Kotsifaki et al., 2021).

For their part, the word foci allow us to establish a trend in recent years of networking and interaction between "Ankle", "Injury", "Machine Learning", "Prevention" and "Sport". This may be mainly due to individual differences, multiple variables and factors and non-linear relationships between them (de Leeuw et al., 2022), have allowed research progress to focus its efforts on machine learning as a tool to identify patterns and create predictive models from the analysis of large data sets, however, research is currently required to incorporate this technology in the prediction of sports injuries (Luu et al., 2020), which makes it an attractive topic for researchers and professionals in sports science.

On the other hand, the scientific collaboration networks show the relationship between authors, publication dates and citation frequency, being the centrality of the network marked by two authors: "Rácz" and "Grand", who present the highest number of links, identifying leadership and being "star" nodes within the established links. It is important to mention that the greater the number of nodes and relationships, the greater the density of scientific collaboration, generating greater access to information, favoring the availability of resources and scientific productivity (Achury-Saldaña et al., 2022).

Additionally, this analysis shows the dendrogram from the abstracts, it is important to remember that this type of diagram creates clusters in the form of a tree that allows organizing the data, relating the sets, and hierarchizing them (Backhaus et al., 2023). From the graph it can be interpreted that injuries and studies have a more relevant relationship, just as studies on this subject are developed in sport, where there is a relationship with performance and training. This may be due to the increase in studies that seek to predict not only injuries but also the level of performance presented by the athlete, an example of this is the study by Hoog et al. (2021) where they used machine learning to determine the age-related decline in performance.

## Conclusions

The present bibliometric analysis allows establishing that there is an increase in research trends towards the prevention of ankle injuries, using neural networks, algorithms and computational methods that facilitate the prediction of sports injuries.

The increase of studies in this area relates words such as: sport, ankle joint, risk factors and technology (mobile applications, computational methods, wireless communication), showing the importance and progress in the use of technological resources, such as machine learning. However, there is still much to be developed and researched in this field, being an opportunity for sports science professionals to explore, innovate and research on the management, control, evaluation, and training of athletes through AI.

The analysis of scientific evidence establishes the importance of prospective studies, which may occur due to the interaction between the words "Ankle", "Injury", "Machine Learning", "Prevention" and "Sport", non-linear and differential variables within a context, allowing a more efficient data recording, a forward tracking that allows a better understanding of the complex interaction between different risk factors.

## Ethics Committee Statement

Not applicable because this is a bibliometric analysis study.

## Conflict of Interest Statement

The authors declare that they have no conflicts of interest.

## Funding

The study did not receive any external funding.

## Authors' Contribution

Conceptualization: Marlon Felipe Burbano, Andres Felipe Villaquiran Hurtado, Nancy Janneth Molano; Data Curation: Marlon Felipe Burbano, Andres Felipe Villaquiran Hurtado; Formal analysis: Marlon Felipe Burbano, Andres Felipe Villaquiran Hurtado; Research: Viviana Marcela Celis and Jeffrey Alexander Hoyos, Nancy Janneth Molano; Methodology: Marlon Felipe Burbano, Andrés Felipe Villaquiran Hurtado, Viviana Marcela Celis and Jeffrey Alexander Hoyos, Nancy Janneth Molano; Software: Marlon Felipe Burbano; Supervision: Andres Felipe Villaquiran Hurtado; Validation: Marlon Felipe Burbano, Andrés Felipe Villaquiran Hurtado, Viviana Marcela Celis and Jeffrey Alexander Hoyos; Visualization: Marlon Felipe Burbano, Andrés Felipe Villaquiran Hurtado, Viviana Marcela Celis and Jeffrey Alexander Hoyos; Writing - Marlon Felipe Burbano, Andrés Felipe Villaquiran Hurtado, Viviana Marcela Celis and Jeffrey Alexander Hoyos, Nancy Janneth Molano; Writing - proofreading and editing: Marlon Felipe Burbano, Andrés Felipe Villaquiran Hurtado, Viviana Marcela Celis and Jeffrey Alexander Hoyos, Nancy Janneth Molano.

## Data Availability Statement

The data that support the findings of this study are available in the following repository: <https://github.com/mfurbano/BibliometricAnkleInjury2024>

## Acknowledgements

We thank the Universidad del Cauca for its support in strengthening research.

## References

- Abdul Razak, A. H., Zayegh, A., Begg, R. K., & Wahab, Y. (2012). Foot plantar pressure measurement system: a review. *Sensors*, 12(7), 9884–9912. <https://doi.org/10.3390/s120709884>
- Achury-Saldaña, D. M., Castañeda-Rodríguez, L. A., & Perianes-Rodríguez, A. (2022). Differences in scientific collaboration and their effects on research influence: a quantitative analysis of nursing publications in Latin America (Scopus, 2005–2020). *Heliyon*, 8(10), e11047. <https://doi.org/10.1016/j.heliyon.2022.e11047>
- Adán, L., García-Angulo, A., Gómez-Ruano, M. A., Sainz de Baranda, P., & Ortega-Toro, E. (2020). Análisis bibliométrico de la producción científica en fútbol femenino. *Journal of Sport & Health Research*, 12(3), 302–317. <https://recyt.fecyt.es/index.php/JSHR/article/view/83561>
- Arruda, H., Silva, E. R., Lessa, M., Proença, D., Jr, & Bartholo, R. (2022). VOSviewer and bibliometrix. *Journal of the Medical Library Association: JMLA*, 110(3), 392–395. <https://doi.org/10.5195/jmla.2022.1434>
- Backhaus, K., Erichson, B., Gensler, S., Weiber, R., & Weiber, T. (2023). Cluster Analysis. In K. Backhaus, *Multivariate Analysis* (pp. 453–532). SpringerGabler. [https://doi.org/10.1007/978-3-658-40411-6\\_8](https://doi.org/10.1007/978-3-658-40411-6_8)
- De Fazio, R., Mastronardi, V. M., De Vittorio, M., & Visconti, P. (2023). Wearable sensors and smart devices to monitor rehabilitation parameters and sports performance: an overview. *Sensors (Basel, Switzerland)*, 23(4), 1856. <https://doi.org/10.3390/s23041856>
- de Leeuw, A., van der Zwaard, S., van Baar, R., & Knobbe, A. (2022). Personalized machine learning approach to injury monitoring in elite volleyball players. *European Journal of Sport Science*, 22(4), 511–520. <https://doi.org/10.1080/17461391.2021.1887369>
- Delahunt, E., & Remus, A. (2019). Risk factors for lateral ankle sprains and chronic ankle instability. *Journal of Athletic Training*, 54(6), 611–616. <https://doi.org/10.4085/1062-6050-44-18>
- Feria-Arias, E., Boukhemis, K., Kreulen, C., & Giza, E. (2018). Foot and ankle injuries in soccer. *American Journal of Orthopedics*, 47(10). <https://doi.org/10.12788/ajo.2018.0096>
- Gulbrandsen, M., Hartigan, D. E., Patel, K. A., Makovicka, J. L., Tummala, S. V., & Chhabra, A. (2019). Ten-year epidemiology of ankle injuries in men's and women's collegiate soccer players. *Journal of Athletic Training*, 54(8), 881–888. <https://doi.org/10.4085/1062-6050-144-18>

- Gupta, P., Kingston, K. A., O'Malley, M., Williams, R. J., & Ramkumar, P. N. (2023). Advancements in artificial intelligence for foot and ankle surgery: a systematic review. *Foot & Ankle Orthopaedics*, 8(1), 247301142211510. <https://doi.org/10.1177/24730114221151079>
- Halabchi, F., & Hassabi, M. (2020). Acute ankle sprain in athletes: clinical aspects and algorithmic approach. *World Journal of Orthopaedics*, 11(12), 534–558. <https://doi.org/10.5312/wjo.v11.i12.534>
- Hoog Antink, C., Braczynski, A. K., & Ganse, B. (2021). Learning from machine learning: prediction of age-related athletic performance decline trajectories. *GeroScience*, 43(5), 2547–2559. <https://doi.org/10.1007/s11357-021-00411-4>
- Jauhainen, S., Kauppi, J.-P., Leppänen, M., Pasanen, K., Parkkari, J., Vasankari, T., Kannus, P., & Äyrämö, S. (2021). New machine learning approach for detection of injury risk factors in young team sport athletes. *International Journal of Sports Medicine*, 42(02), 175–182. <https://doi.org/10.1055/a-1231-5304>
- Kaminski, T. W., Needle, A. R., & Delahunt, E. (2019). Prevention of lateral ankle sprains. *Journal of Athletic Training*, 54(6), 650–661. <https://doi.org/10.4085/1062-6050-487-17>
- Kobayashi, T., Tanaka, M., & Shida, M. (2016). Intrinsic risk factors of lateral ankle sprain: a systematic review and meta-analysis. *Sports Health: A Multidisciplinary Approach*, 8(2), 190–193. <https://doi.org/10.1177/1941738115623775>
- Kolokotsios, S., Drousia, G., Koukoulithras, I., & Plexousakis, M. (2021). Ankle injuries in soccer players: a narrative review. *Cureus*, 13(8), e17228. <https://doi.org/10.7759/cureus.17228>
- Kotsifaki, A., Korakakis, V., Graham-Smith, P., Sideris, V., & Whiteley, R. (2021). Vertical and horizontal hop performance: contributions of the hip, knee, and ankle. *Sports Health: A Multidisciplinary Approach*, 13(2), 128–135. <https://doi.org/10.1177/1941738120976363>
- Li, X., & Kang, M. (2022). Athletes' Ankle injury features and rehabilitation methods based on internet big data. *Journal of Nanomaterials*, 2022, 1–9. <https://doi.org/10.1155/2022/7904739>
- Llamas, M. D. C. (2021). Modelización matemática para la predicción y prevención de lesiones deportivas. *Retos: Nuevas Tendencias en Educación Física, Deportes y Recreación*, 39, 681–685. <https://doi.org/10.47197/retos.v0i39.81315>
- López-Robles, J. R., Cobo, M. J., Gutiérrez-Salcedo, M., Martínez-Sánchez, M. A., Gamboa-Rosales, N. K., & Herrera-Viedma, E. (2021). 30th Anniversary of applied intelligence: a combination of bibliometrics and thematic analysis using SciMAT. *Applied Intelligence*, 51(9), 6547–6568. <https://doi.org/10.1007/s10489-021-02584-z>
- Luu, B. C., Wright, A. L., Haeberle, H. S., Karnuta, J. M., Schickendantz, M. S., Makhni, E. C., Nwachukwu, B. U., Williams, R. J., & Ramkumar, P. N. (2020). Machine learning outperforms logistic regression analysis to predict next-season NHL player injury: an analysis of 2322 players from 2007 to 2017. *Orthopaedic Journal of Sports Medicine*, 8(9), 232596712095340. <https://doi.org/10.1177/2325967120953404>
- Moreno, G. A., & Gutiérrez, R. E. (2020). Estudio prospectivo de la tecnología en la educación superior en Colombia al 2050. *Revista Universidad y Empresa*, 22(38), 160–182. <http://dx.doi.org/10.12804/revistas.urosario.edu.co/empresa/a.7583>
- Müller, A. M., Maher, C. A., Vandelanotte, C., Hingle, M., Middelweerd, A., Lopez, M. L., DeSmet, A., Short, C. E., Nathan, N., Hutchesson, M. J., Poppe, L., Woods, C. B., Williams, S. L., & Wark, P. A. (2018). Physical activity, sedentary behavior, and diet-related eHealth and mHealth research: bibliometric analysis. *Journal of Medical Internet Research*, 20(4), e122. <https://doi.org/10.2196/jmir.8954>
- Murtagh, F., & Legendre, P. (2014). Ward's hierarchical agglomerative clustering method: which algorithms implement ward's criterion? *Journal of Classification*, 31(3), 274–295. <https://doi.org/10.1007/s00357-014-9161-z>
- Nuyts, L., De Brabandere, A., Van Rossom, S., Davis, J., & Vanwanseele, B. (2022). Machine-learned-based prediction of lower extremity overuse injuries using pressure plates. *Frontiers in Bioengineering and Biotechnology*, 10. <https://doi.org/10.3389/fbioe.2022.987118>
- Olmedilla, A., Andreu, M. D., Abenza, L., Ortín, F. J., & Blas, A. (2006). Lesiones y factores deportivos en futbolistas jóvenes. *Cultura, Ciencia y Deporte*, 2(5), 59–66. <https://doi.org/10.12800/ccd.v2i5.176>
- Olmedilla, A., Andreu, M. D., Ortín, F. J., & Blas, A. (2008). Epidemiología lesional en futbolistas jóvenes. *Cultura, Ciencia y Deporte*, 3(9), 177–183. <https://doi.org/10.12800/ccd.v3i9.165>
- Pizarro, R., & Masson, W. (2020). Estudios prospectivos poblacionales: las fortalezas de las cohortes históricas. *Acta Gastroenterológica Latinoamericana*, 50(4), 382–385.
- R Core Team. (n.d.). R: The R Project for Statistical Computing. <https://www.r-project.org/>

- Ramkumar, P. N., Luu, B. C., Haeberle, H. S., Karnuta, J. M., Nwachukwu, B. U., & Williams, R. J. (2022). Sports medicine and artificial intelligence: a primer. *The American Journal of Sports Medicine*, 50(4), 1166–1174. <https://doi.org/10.1177/03635465211008648>
- Román-Belmonte, J. M., Corte-Rodríguez, H., & Rodríguez-Merchán, E. C. (2021). Artificial intelligence in musculoskeletal conditions. *Frontiers in bioscience (Landmark edition)*, 26(11), 1340–1348. <https://doi.org/10.52586/5027>
- RPubs. (n.d.). Word Cloud. Retrieved from RPubs - Word Cloud
- Salazar F., P., Manterola, C., Quiroz S., G., García M., N., Otzen H., T., Mora V., M., & Duque P., G. (2019). Estudios de cohortes. 1ª parte. Descripción, metodología y aplicaciones. *Revista de Cirugía*, 71(5). <https://doi.org/10.35687/s2452-45492019005431>
- Shah, S. H. H., Lei, S., Ali, M., Doronin, D., & Hussain, S. T. (2020). Prosumption: bibliometric analysis using HistCite and VOSviewer. *Kybernetes*, 49(3), 1020-1045. <http://dx.doi.org/10.1108/K-12-2018-0696>
- Song, Y., Cen, X., Zhang, Y., Bíró, I., Ji, Y., & Gu, Y. (2022). Development and validation of a subject-specific coupled model for foot and sports shoe complex: a pilot computational study. *Bioengineering*, 9(10), 553. <https://doi.org/10.3390/bioengineering9100553>
- Wilzman, A. R., Tenforde, A. S., Troy, K. L., Hunt, K., Fogel, N., Roche, M. D., Kraus, E., Trikha, R., Delp, S., & Fredericson, M. (2022). Medical and biomechanical risk factors for incident bone stress injury in collegiate runners: can plantar pressure predict injury? *Orthopaedic Journal of Sports Medicine*, 10(6), 232596712211047. <https://doi.org/10.1177/23259671221104793>
- Zhao, S., Liu, R., Fei, C., Zia, A. W., & Jing, L. (2020). Flexible sensor matrix film-based wearable plantar pressure force measurement and analysis system. *PLOS ONE*, 15(8), e0237090. <https://doi.org/10.1371/journal.pone.0237090>