


# Mapping the youth soccer: A bibliometric analysis using R-tool

DIGITAL HEALTH  
Volume 9: 1–16  
© The Author(s) 2023  
Article reuse guidelines:  
[sagepub.com/journals-permissions](https://sagepub.com/journals-permissions)  
DOI: 10.1177/20552076231183550  
[journals.sagepub.com/home/dhj](https://journals.sagepub.com/home/dhj)



Bo Liu , Chang-jing Zhou, Hao-wei Ma and Bo Gong

## Abstract

In recent years, there has been an increase in the scientific production of youth soccer. However, a panoramic map of research on this subject does not exist. The aim of this study was to identify global research trends in youth soccer over time, among the main levels of analysis: sources, authors, documents, and keywords. The bibliometric software Biblioshiny was used to analyze 2606 articles in Web of Science (WoS) published between 2012 and 2021. The main conclusion is that US and UK scholars dominate the research; the topics of research are changing with the real needs, and research on the topic of performance has been of interest to scholars; talent identification and development, performance, injury prevention, and concussion are the studies of interest to scholars in this area. This finding, which offers a global picture of youth soccer research over time, can help future research in this or similar domains.

## Keywords

Youth soccer, bibliometric analysis, bibliometric, WoS

Submission date: 15 December 2022; Acceptance date: 5 June 2023

## Introduction

As the most popular sport in the world, soccer generated enormous social wealth and a sizable human market.<sup>1</sup> There are 270 million (4% of the world's population) active soccer participants worldwide, including referees and officials, according to the FIFA Big Count 2006.<sup>2</sup> Players' payment had increased significantly as a result of the sport's growing professionalization and commercialization,<sup>3</sup> as well as the need for clubs to succeed at a high level of competition.<sup>4</sup> According to the FIFA Global Transfer Report 2021,<sup>5</sup> while there were 185 transactions among FIFA's 211 member associations in 2021, two fewer than in 2020, there were 4544 international moves by clubs, and the total transfer fee was USD 4.86 billion.

The requirement for excellent performance and rising costs have sparked a global arms race to find youth talent.<sup>6</sup> Additionally, knowledge about the performance of youth soccer players can help different professionals and how this information can be used in the future. Considering the definition of youth, according to the World Health Organization and UNICEF age definition criteria for youth, it refers to the group of 15 to 24 years old. However, there are differences in the age range of

adolescents in soccer-related studies, which makes it difficult to form age group criteria for some indicators.<sup>7</sup> In a review from McCalman,<sup>8</sup> the age range of adolescents was 9–23 years old, whereas in another study,<sup>9</sup> the starting age of this age group was 7 years old. Given the difficulty in defining the age range in the current study, our literature was selected by reviewing only articles whose titles contained the word youth. To the best of our knowledge, studies of youth soccer have targeted groups that are more likely to be players with the potential to become professionals rather than recreational and amateur soccer groups in the same age range.

Inspired by deliberate practice theory,<sup>10</sup> researchers have conducted numerous studies related to youth soccer. These studies encompass various aspects such as talent identification and development (TID) in soccer,<sup>11,12</sup> which focuses

School of Athletic Performance, Shanghai University of Sport, Shanghai, China

### Corresponding author:

Bo Gong, School of Athletic Performance, Shanghai University of Sport, Shanghai 200438, China.  
Email: [gbok2001@163.com](mailto:gbok2001@163.com)



on the process of recognizing and nurturing young players' potential. Additionally, the relative age effect<sup>13,14</sup> refers to the phenomenon where players born earlier in the selection year tend to have a competitive advantage due to age-related physical and psychological differences. Furthermore, researchers have also explored incident injuries<sup>15,16</sup> in youth soccer, which refers to injuries that occur during matches or training sessions. Understanding the causes, patterns, and consequences of such injuries is crucial for ensuring the safety and well-being of young soccer players. In addition to these areas of study, researchers have also examined match performance,<sup>17</sup> among other factors, to gain insights into various aspects of youth soccer.

Through literature searches, it is possible to observe the growth in linked-topic research. Since more academic articles are being published every year, it is getting more and more difficult to keep up with everything that is being published.<sup>18</sup> Therefore, it's difficult to observe some information implied behind the research from a holistic perspective, such as research trends, hot spots, collaborator networks, citations, etc. To illustrate the knowledge structures and dynamic evolution of a particular research subject, bibliometric analysis uses statistical methodologies which can give a detailed overview of a study area, topic, or issue as a quantitative method.<sup>19</sup> Enlightened by previous studies, we aimed at mapping youth soccer literature by analyzing the influence of various nations, organizations, writers, journals, and terms using the Biblioshiny in the present study. The study's findings may offer broad perspectives and recommendations about the current state of youth soccer research worldwide.

## Methodology

The bibliometric method is widely used in multiple fields, such as climate change,<sup>20</sup> patent mining,<sup>21</sup> sports,<sup>22</sup> with commonly used software, for instance, CiteSpace,<sup>23</sup> VOSviewer,<sup>24</sup> SciMAT,<sup>25</sup> and R-tool.<sup>26</sup>

The purpose of this article is to evaluate publications, citations, and information sources addressing the developments in the field of youth soccer by following the steps of a bibliometric study. To characterize the bibliographic data, number of publications, citations, authors, and citations per nation, this paper employs bibliometric indicators. The study also examines the intellectual frameworks that show how articles are received by the scientific community as well as the social networks of authors, universities, and collaboration between nations.

The Biblioshiny, an R package, was used for statistical processing. The process was divided into two parts:

1. Data collection; and
2. Bibliometric analysis.

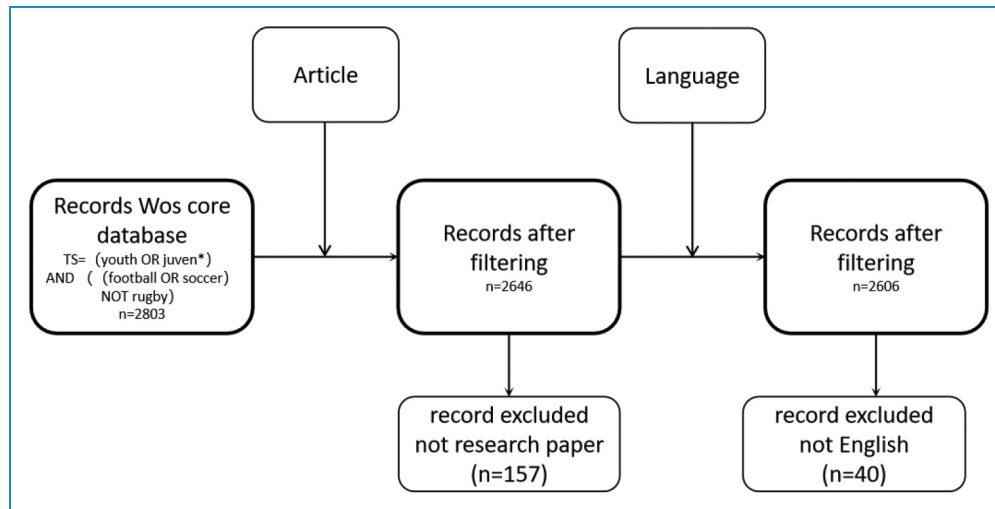
## Data collection

In this process, published research was retrieved in the Web of Science (WoS) core collection by Clarivate, the most widely recommended<sup>27</sup> and reliable<sup>28</sup> international bibliometric database. The search strategy was used: TS(topic search) = (youth OR juven\*) AND ((football OR soccer) NOT rugby) and time span = Jan.1 2012–Dec.31 2021. This search took place on Nov.15 2022.

In order to enhance the precision of this investigation, the screening process of the retrieved papers was conducted by BL and CJZ independently based on title and abstract, with oversight from BG, to ascertain adherence to the predetermined inclusion and exclusion criteria. When there is a discrepancy between the two reviewers, an internal discussion is first conducted to reach a consensus. If a disagreement persists after the discussion, a third reviewer, BG, will make the final decision. The inclusion criteria for this study were as follows: (1) the articles were required to have a specific focus on the domain of youth soccer; (2) only articles written in English were considered eligible. On the other hand, the exclusion criteria were as follows: (1) records pertaining to announcements and book reviews were excluded, with a focus solely on regular research papers; conference papers were also excluded to maintain the focus on journal distribution analysis; (2) studies that primarily concentrated on competitive youth soccer populations were included, while those focusing on recreational soccer populations were excluded. Total records retrieved were 2803, reviews, books and other languages were eliminated as we want to map the articles written in English, then the final number of articles recorded between 2012 and 2021 in this work is 2606 (Figure 1).

## Analysis

In the analysis process, bibliometric techniques were used to extract networks of different levels (Table 1). Co-citation analysis was first put up by American intelligence scientist Small as a technique to gauge the degree of link between documents. Two (or more) papers that are simultaneously cited by one or more subsequent papers are said to constitute a co-citation relationship.<sup>29</sup> For example, literature A cites both literature C and D. At this point, literature C and D are cited together. The number of documents that cite both is called co-citation intensity, and in this case the co-citation intensity is 1, because only document A cites both C and D. Co-word analysis is a technique for examining the actual content of a publication itself. Words in co-word analysis are usually derived from "author keywords," but in its absence, notable words can also be extracted from "article title," "abstract," and the words of interest can also be extracted from "article title," "abstract," and "full text" for analysis. Similar to co-citation analysis, co-word analysis assumes that words that



**Figure 1.** Data collection flow diagram.

frequently occur together are thematically related to each other.<sup>30</sup> A co-author network is a common form of scientific collaboration network, which is a network established through the cooperative relationship of the authors of a paper. If two authors co-authored one or more papers, a contiguous relationship is created between the two authors. Co-authorship networks can provide a powerful tool for the study of the structure of academic relationship networks.<sup>31</sup>

## Result and discussion

### Research trends

According to the scientific production associated with youth soccer from 2012 to 2021 (Figure 2) associated with bibliometric analysis (Table 2), studies in the last decade have grown rapidly at an annual rate of 14.86%. As is shown in Table 2, during the previous decades, 2455 papers were written by 7077 authors worldwide. Interestingly, the analysis revealed large figures for co-author per document and international co-authorships. This indicates that the authors in this field of study prefer to collaborate and that there is a high percentage of international collaboration, which is in line with the trend of globalization of soccer.

### Sources

With regard to sources, the most productive journals are shown in Table 3. In the past ten years, the number of articles published in the Journal of Sports Science and Journal of Strength and the Conditioning Research were exceptionally high, 168 and 138, respectively, accounting for 6.8% and 5.6% of the total number of articles published.

With regard to the most local citation sources, 15 journals were cited more than 1000 times. Journal of Sports Science, Journal of Strength and Conditioning Research, and British Journal of Sports Medicine lead the top tree, with the number of citations being 6870, 5426, and 4470, respectively, far exceeding that of peer journals. The H-index ranking of these three journals has not changed since the perspective is turned to Source Impact. An interesting discovery is that although Sports Medicine ranks 18th in the number of publications (with a number of 36), its H-index is the same as that of British Journal of Sports Medicine with a value of 18. As such we can infer that although Sports Medicine does not publish a high volume of articles in youth soccer, they are all of the high quality.

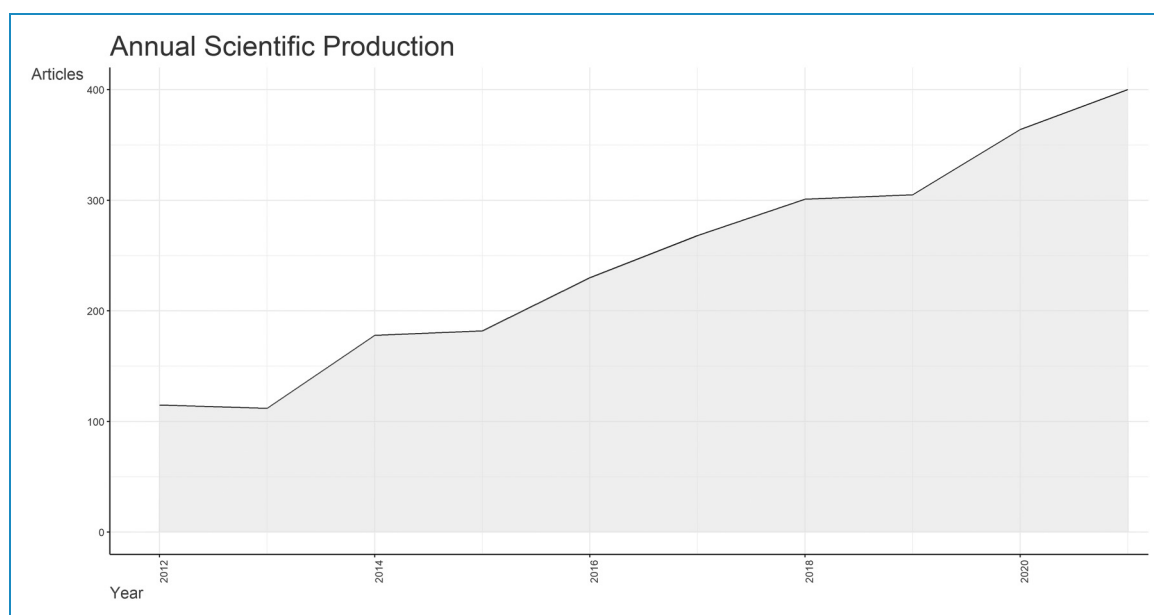
When concerning the Source Dynamics of the top 10 journals (Figure 3), the cumulative publication volume of most journals showed a steady growth trend except two. The British Journal of Sports Medicine was firmly in third place in terms of cumulative number of publications until 2016, however, it was overtaken by seven journals including the International Journal of Environment Research and Public Health, International Journal of Sports Science & Coaching and Journal of Human Kinetics. The journal that has seen the most increase in the number of papers published in the area of youth soccer over the past few years is the International Journal of Environment Research and Public Health, ranking from tenth in 2019 to the third in 2021, closely following the Journal of Sports Science and Journal of Strength and Conditioning Research.

Three clusters were discovered in the source co-citation network (Figure 4). In general, the journals at each cluster's furthest edges show a weaker relationship to the other cluster, while the size of the nodes shows the intensity of connection.<sup>32</sup> It is worth noting that cluster 3 is located in the middle of the other two clusters, connecting the other two clusters together, in

**Table 1.** Specifications of the analysis.

| Analysis level | Metrics   | Analysis unit           | Bibliometric technique       | Statistical technique | Structure                 |
|----------------|---|-------------------------|------------------------------|-----------------------|---------------------------|
| Source         | Source dynamics and most productive sources   | Journal                 | Co-citation                  | Network               | Conceptual                |
| Authors        | Most productive authors and annual production per author<br>Most productive countries | Authors<br>Countries    | Co-citation & collaboration  | Network               | Intellectual and social   |
| Documents      | Most productive countries<br>Most cited documents                                     | Countries<br>References | Collaboration<br>Co-citation | Network               | Social<br>Intellectual    |
| Keyword        | Most frequent words   | References & keywords   | Co-citation & co-word        | Network               | Intellectual & conceptual |

Source: Adapted from “Bibliometric: An R-tool for comprehensive science mapping analysis.”<sup>18</sup>

**Figure 2.** Growth in sources.

addition to the fact that cluster 3 shows an advantage, both in terms of betweenness and number of publications. Analysis shows that more than half of the top ten productive journals are located in cluster 3, led by the Journal of Sport Science and Sports Medicine. Although the number of publications in cluster 1 is not greater compared to cluster 3, the betweenness of cluster 1 yields a larger number than other clusters, which is created by the British Journal of Sports Medicine and *Medicine & Science in Sports & Exercises*. On the contrary, the journals in cluster 2 are inferior to the other two clusters both in terms of number of publications and centrality.

According to the titles of the journals, the red cluster is dominated by sports medicine, such as the *British Journal of Sports Medicine*, *American Journal of Sports Medicine*, *Clinical Journal of Sports Medicine*, etc. These journals are more inclined to publish research on medical topics, such as the ECG of high-level junior soccer players,<sup>33</sup> the left ventricular function of pre-adolescent soccer players,<sup>34</sup> head impacts.<sup>35</sup> *Journal of Sports Sciences*, the *Journal of Strength and Conditioning Research* and *Sports Medicine* are in the green cluster, are more focused on the topics of physical fitness,<sup>36</sup> training,<sup>37</sup> and sports medicine.<sup>38</sup> The blue clusters are dominated by

**Table 2.** Main information of the papers.

| Description                    | No.       |
|--------------------------------|-----------|
| Time span                      | 2012–2021 |
| Documents                      | 2455      |
| Sources                        | 391       |
| Annual growth rate %           | 14.86     |
| Keywords plus (ID)             | 3426      |
| Author's keywords (DE)         | 3922      |
| Average citation per document  | 19.83     |
| Authors                        | 7077      |
| Authors per document           | 2.86      |
| Co-authors per document        | 4.97      |
| International co-authorships % | 38.13     |
| Publications per year          | 24.55     |

**Table 3.** Top 10 productive journals in youth soccer.

| Sources   | Articles |
|---|----------|
| JOURNAL OF SPORTS SCIENCES  | 168      |
| JOURNAL OF STRENGTH AND CONDITIONING RESEARCH                     | 138      |
| INTERNATIONAL JOURNAL OF ENVIRONMENTAL RESEARCH AND PUBLIC HEALTH | 94       |
| INTERNATIONAL JOURNAL OF SPORTS SCIENCE & COACHING                | 82       |
| JOURNAL OF HUMAN KINETICS   | 75       |
| PLOS ONE  | 66       |
| SCIENCE AND MEDICINE IN FOOTBALL                                  | 62       |
| FRONTIERS IN PSYCHOLOGY   | 48       |
| SCANDINAVIAN JOURNAL OF MEDICINE & SCIENCE IN SPORTS              | 48       |
| BRITISH JOURNAL OF SPORTS MEDICINE                                | 47       |

sports psychology topics such as motivation,<sup>39</sup> attitudes toward moral decision,<sup>40</sup> represented by journals such as

Psychology of Sport and Exercise, Journal of Applied Sport Psychology, Sport Psychologist.

### Authors

Regarding the analysis at the author level, the results display first the most relevant authors and some bibliometric indicators (Table 4), secondly the author collaboration network, finally the authors' production over time (Figure 5) and co-citation network (Figure 6).

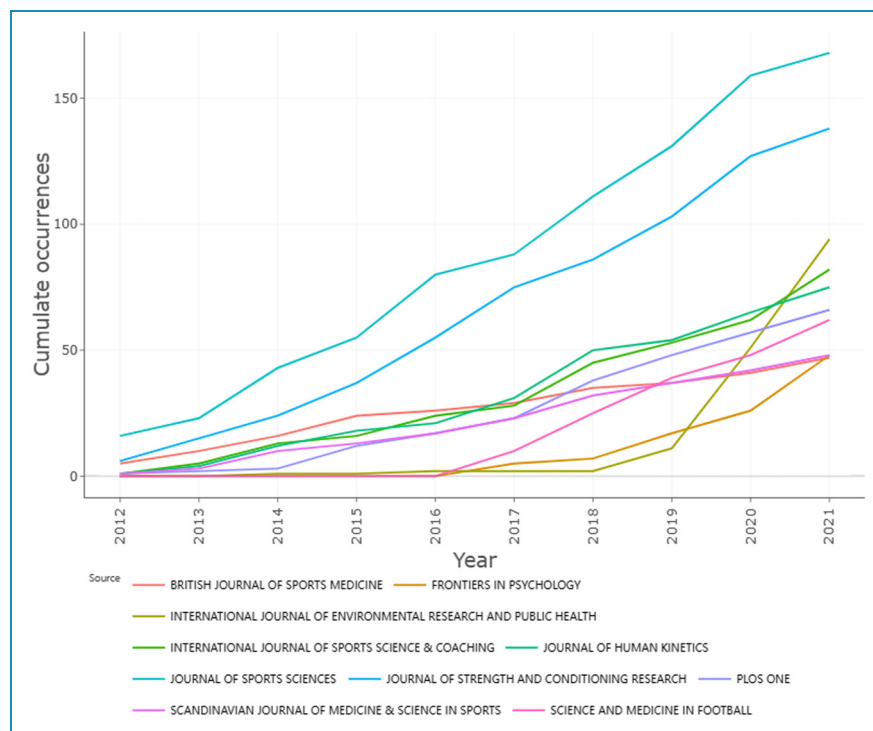
As can be observed in Table 4, there is consistency in the tendency of the number of authors' publications and the fractionalized frequency. When comparing the H-index and the total citations, the figures show some fluctuations. Clemente FM ranks fifth in the top 10 over the last decade, but it has a low citation number of merely 427, in contrast, Vaeyens R has a publication of 30, but it has a high H-index of 21, only 2 points lower than the top two authors. When analyzing the total citations of him, the same high number is observed, which indicates that this author has not only a quantitative advantage but also a higher quality of articles.

Top-authors' productive over time can effectively reflect the time distribution and quantity characteristics of the authors' articles over a certain period of time. As shown in Figure 5, five authors such as Lloyd RS, Oliver JL, and Malina RM are both productive and continuous, they have had an annual output of papers for the past decade and have sustained a high total citation. Clemente FM has only been involved in the field of youth soccer since 2016, however, he has the best performance in terms of output and total citations in the last two years. He has 28 research outputs in 2020 and 2021, which explains why he has higher outputs, but does not rank high in total citations in most relevant authors.

By analyzing the collaboration network, a total of 10 clusters were generated and most of them were interconnected with each other, except for cluster 4 and cluster 10. Cluster 1 had the largest number of highly productive authors, accounting for half of the top ten, and the author studies in this cluster focused on maturation,<sup>41</sup> relative age effects,<sup>42</sup> mental fatigue,<sup>43</sup> and injuries<sup>44</sup> in youth soccer players. Although the number of authors in cluster 2 was not as high as in cluster 1, the authors with the highest number of publications Lloyd RS, Oliver JL, and Myer GD were in cluster 2, and their research topics focused on youth physical development model,<sup>45</sup> injury risk screening,<sup>46</sup> bio-banding,<sup>47</sup> and neuromuscular control.<sup>48</sup> Authors in cluster 3 are keen on themes related to small-side games,<sup>49</sup> plyometric training,<sup>50</sup> and physical fitness training.<sup>51</sup>

### Country collaboration, production, and citation

The country cooperation social networks, national and international collaboration, as well as productivity and



**Figure 3.** Source dynamics of the top 10 productive journals.

citation networks, were investigated to analyze the nations with the highest levels of production based on the country of authors.

By analyzing the network of collaboration between countries, Figure 7 shows the collaboration network among the top 30 countries, the size of the nodes reflects the amount of collaboration, the larger the nodes the more collaboration, while the thickness of the linkage reflects the number of collaboration, the thicker the linkage, the greater the number. Countries with the most collaboration with others are the United States (US) and the United Kingdom (UK). The UK has collaboration between 47 countries, while the figure for the US is 44. The most frequent collaborations between countries were between the US and the UK, as well as the UK and Australia, with 86 and 60 papers collaborating, respectively.

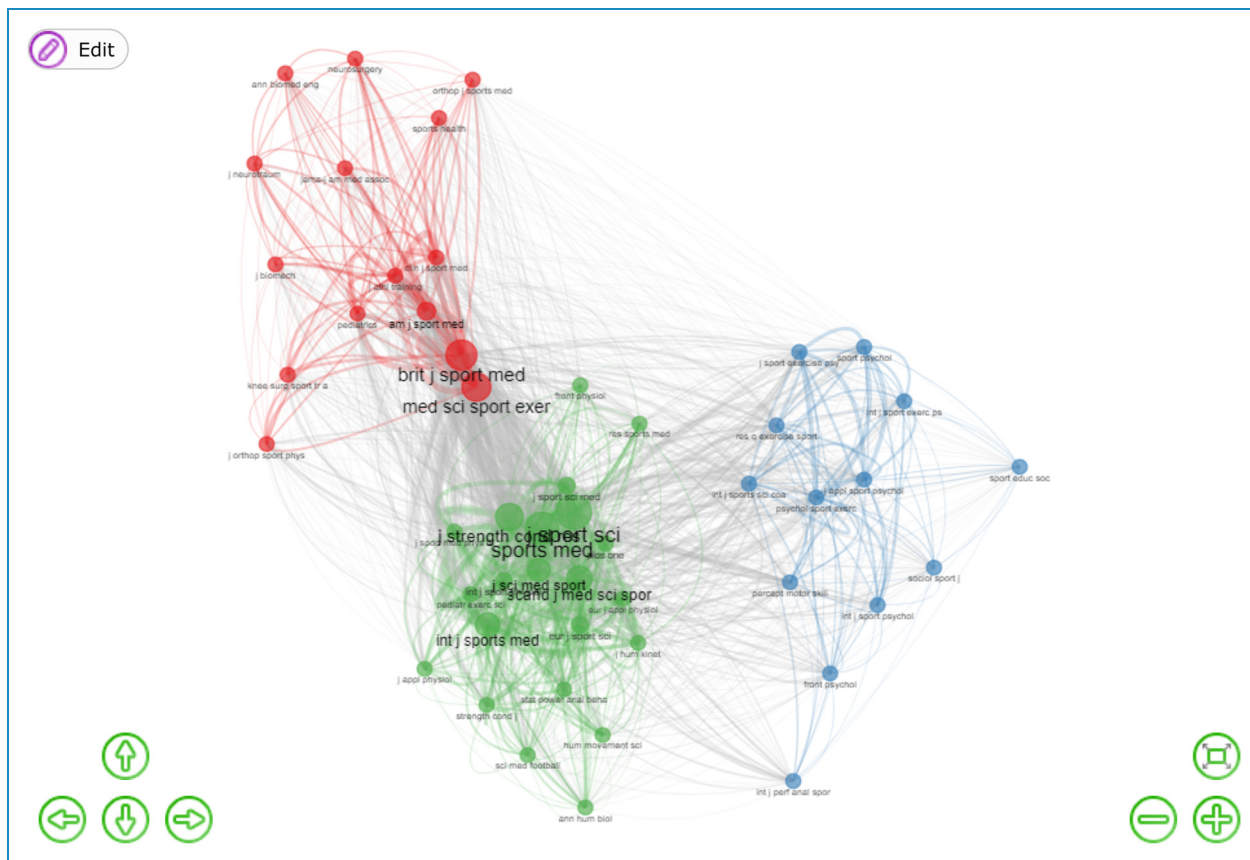
With regard to the most cited countries (Table 5), the USA and the UK show dominance again. The USA had the highest total number of citations at 11,289, followed by the UK at 7929. Canada, Australia, and Spain are in third to fifth place, but their total citations are further away from the top two. An interesting phenomenon was discovered when inspecting the average citations per year (ACP), only Switzerland, Netherlands, and Canada of the top 10 countries in terms of total citations appear in the top 10 ranking of the ACP. Colombia, Kuwait, and Nigeria ranked in the top three in the ACP, notwithstanding the fact that their total citations were not as high.

### Cited local documents

The most 10 local cited documents (Table 6) are mainly from Journal of Sports Sciences, British Journal of Sports Medicine, Annals of Biomedical Engineering, Strength and Conditioning Journal, and Journal of Strength and Conditioning Research. Taking into account of the document clusters coupled by abstracts, two clusters were identified (Figure 8). The literature in cluster 1 (blue cluster) is centered on Sarmiento H<sup>52</sup> and focuses on research related to youth talent identification. Compared to cluster 1, scholars are more focused on injury prevention in football, with Steffen as the main representative.<sup>53</sup> Cluster 2 has a high impact factor and centrality observed from the axes, indicating that this is a well-defined and long-running set of research topics. Despite the low impact factor and centrality of cluster 1, Cobo M J<sup>54</sup> advocates that a dynamic analysis of the literature in this zone is needed to understand their contribution. Inspired by this claim, we analyzed the normalized local citation score of Sarmiento H, which was as high as 10.24, a figure that represents a much higher local citation than his peers.

### Keyword

The most frequent words used by the authors are soccer (football), youth, concussion, youth (sport), concussion, adolescents, talent identification, performance, injury prevention, and so on (Figure 9). An analysis of research



**Figure 4.** Co-citation network of sources. Red cluster = cluster 1; blue cluster = cluster 2; green cluster = cluster 3.

**Table 4.** Most relevant authors.

| Ranking | Author             | No. | H-index | Fractionalized | Total citation |
|---------|--------------------|-----|---------|----------------|----------------|
| 1       | LLOYD RS           | 45  | 23      | 9.16           | 1866           |
| 2       | OLIVER JL          | 44  | 23      | 9.15           | 1904           |
| 3       | MALINA RM          | 41  | 20      | 6.84           | 1419           |
| 4       | MYER GD            | 38  | 22      | 6.65           | 1635           |
| 5       | CLEMENTE FM        | 36  | 12      | 7.20           | 427            |
| 6       | RAMIREZ-CAMPILLO R | 36  | 18      | 5.37           | 902            |
| 7       | COELHO-E-SILVA MJ  | 33  | 14      | 4.63           | 557            |
| 8       | VAEYENS R          | 30  | 21      | 4.91           | 1131           |
| 9       | FRANSEN J          | 28  | 15      | 5.10           | 728            |
| 10      | FIGUEIREDO AJ      | 27  | 13      | 4.62           | 801            |

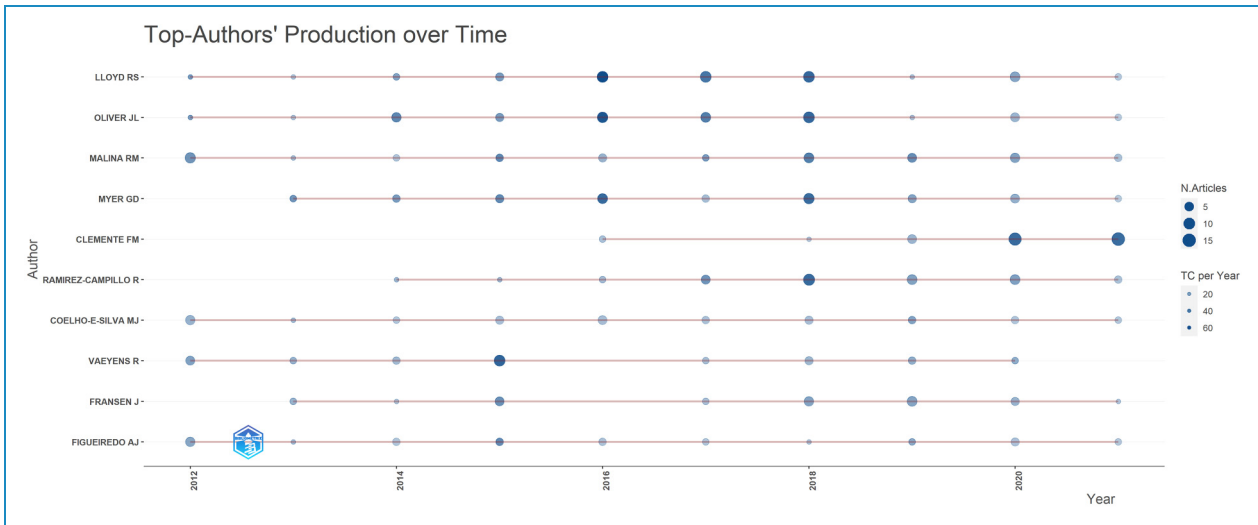


Figure 5. Top authors' production over time.

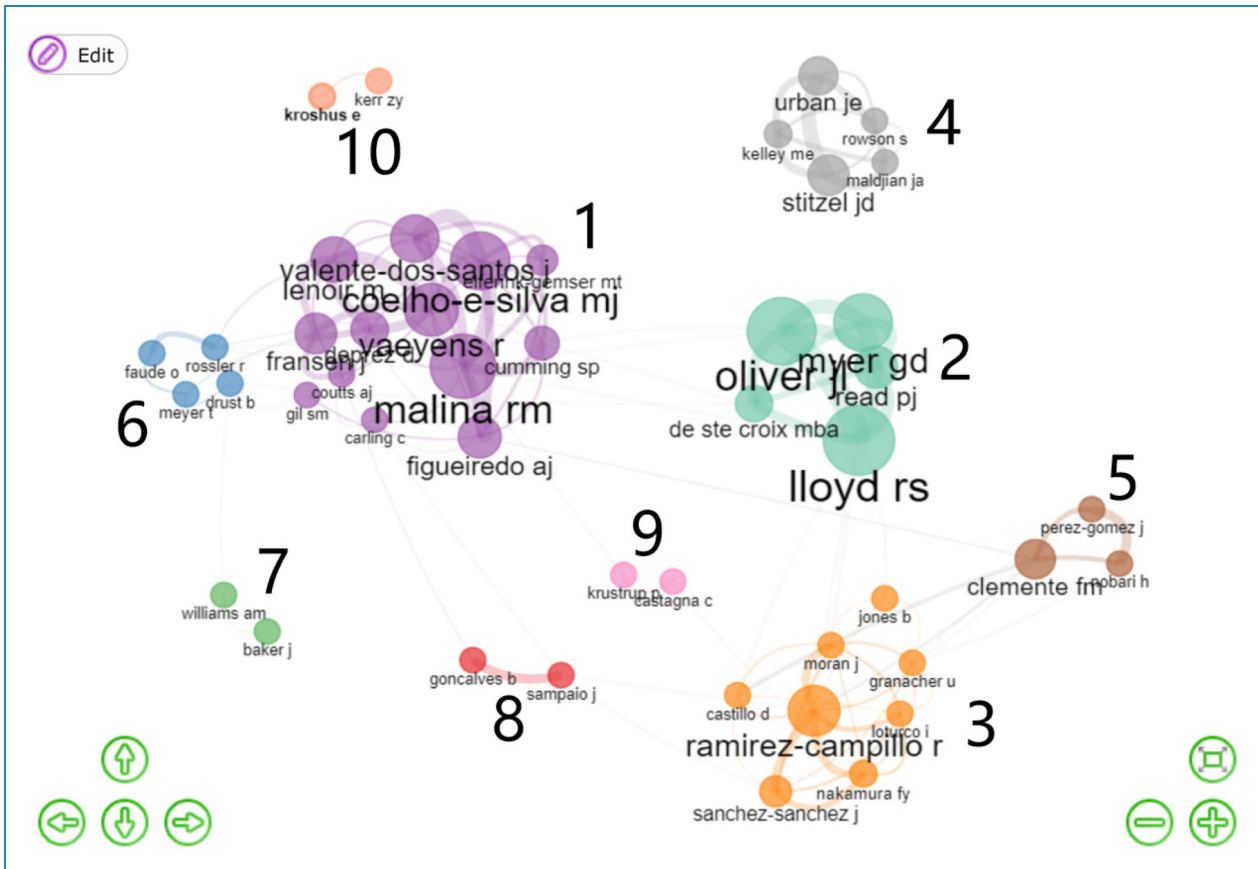


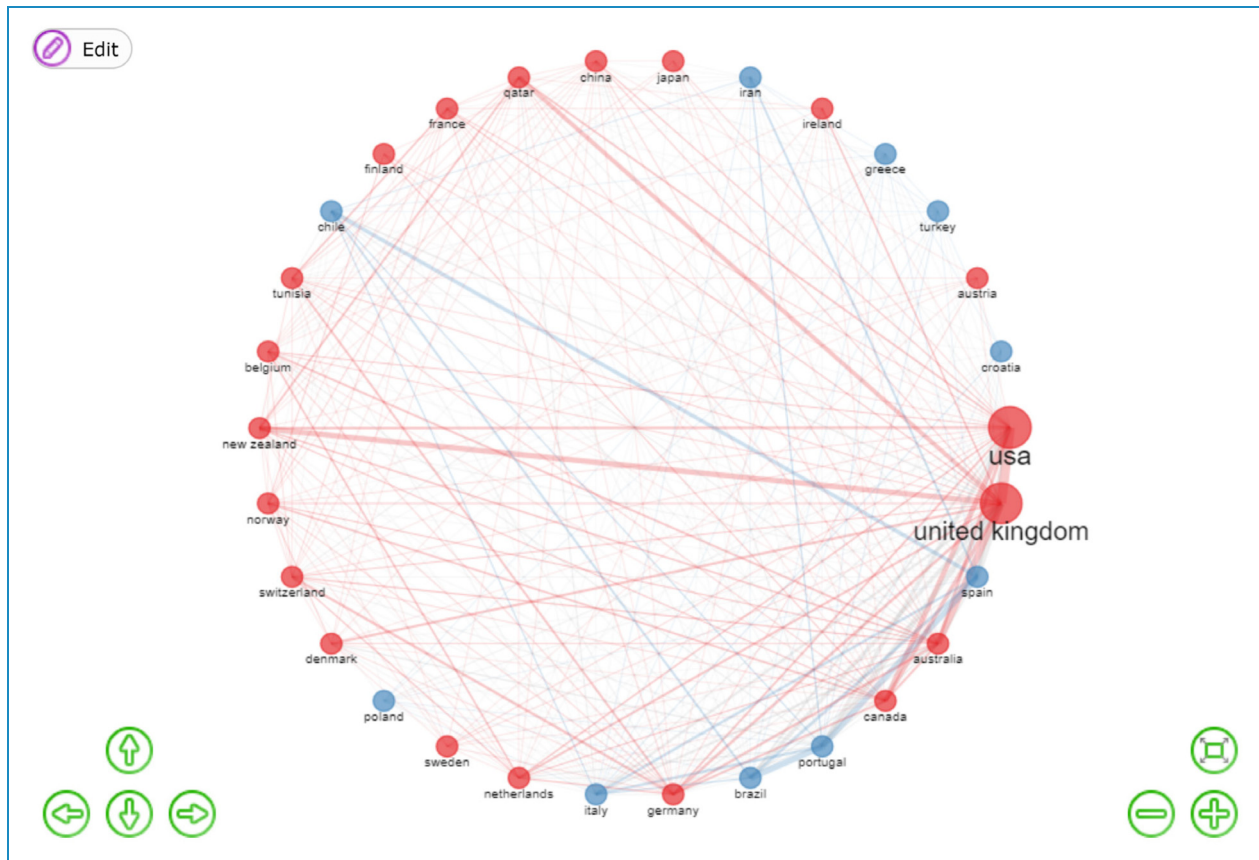
Figure 6. Author collaboration network.

trends (Figure 10) shows that the number of studies on themes velocity, blood lactate, high-intensity running, and aerobic endurance have decreased over the years since 2013, and for comparison, the hot spots themes in recent

years are head impact, coach-athlete relationship, and performance.

According to the co-word analysis (Figure 11), four clusters were found, reflecting four main areas of research. As





**Figure 7.** Country collaboration network.

**Table 5.** Most cited countries.

| Ranking | Country        | TC     | Country      | ACP   |
|---------|----------------|--------|--------------|-------|
| 1       | USA            | 11,289 | COLOMBIA     | 62.33 |
| 2       | UNITED KINGDOM | 7929   | KUWAIT       | 55    |
| 3       | CANADA         | 3612   | NIGERIA      | 46.5  |
| 4       | AUSTRALIA      | 3587   | BELGIUM      | 34.36 |
| 5       | SPAIN          | 2689   | SWITZERLAND  | 30.95 |
| 6       | PORTUGAL       | 2446   | MALAYSIA     | 30    |
| 7       | GERMANY        | 2105   | SAUDI ARABIA | 28    |
| 8       | NETHERLANDS    | 1472   | QATAR        | 27.46 |
| 9       | SWITZERLAND    | 1362   | NETHERLANDS  | 25.38 |
| 10      | BRAZIL         | 1301   | CANADA       | 25.08 |

TC: total citations; ACP: average citations per year.

the co-occurrences network was extracted from the keywords of the authors, therefore, the analysis provides an understanding of the research areas that academics are focusing on.

In cluster 1, scholars have conducted research mainly on the TID of youth soccer. It is revealed that the broad topic of TID is a vibrant area of research, ranging from a large number of books, systematic reviews, 2018, narrative reviews, and empirical papers published in the last 20 years.<sup>55</sup> The sub-themes in this area are predictors for selection of the youth players,<sup>56</sup> relative age effect,<sup>57</sup> biological maturity,<sup>58</sup> and methodological issues.<sup>7</sup>

The studies in cluster 2 are relatively comprehensive, as the most centralized betweenness words are in this cluster, such as soccer, football, and youth. This cluster contains both sports performance topics in cluster 3 and covers injury prevention studies in cluster 4. Additionally, this cluster contains studies related to training<sup>36</sup> and fatigue.<sup>59</sup>

The theme of cluster 3 is sports performance. In terms of physical performance, indicators such as speed,<sup>60</sup> strength,<sup>61</sup> and agility<sup>62</sup> are included. A common testing protocol is to use Global Positioning System<sup>63</sup>

**Table 6.** The most local cited documents.

| Ranking | Document   | Year | LC | GC  | LC/GC | Source  |
|---------|--|------|----|-----|-------|---|
| 1       | High adherence to a neuromuscular injury prevention program (FIFA 11+) improves functional balance and reduces injury risk in Canadian youth female football players: a cluster randomized trial                                   | 2013 | 66 | 212 | 31.13 | British Journal of Sports Medicine            |
| 2       | Talent identification in youth soccer  | 2012 | 65 | 162 | 40.12 | Journal of Sports Sciences                    |
| 3       | Biological maturation of youth athletes: assessment and implications   | 2015 | 62 | 264 | 23.48 | British Journal of Sports Medicine            |
| 4       | Head impact exposure in youth football: Elementary school ages 9–12? Years and the effect of practice structure  | 2013 | 60 | 133 | 45.11 | Annals of Biomedical Engineering              |
| 5       | Head impact exposure in youth football   | 2012 | 57 | 147 | 38.78 | Annals of Biomedical Engineering              |
| 6       | The youth physical development model<br>A new approach to long-term athletic development   | 2012 | 55 | 268 | 20.52 | Strength and Conditioning Journal             |
| 7       | The relative age effect in European professional soccer: Did ten years of research make any difference?  | 2012 | 51 | 101 | 50.5  | Journal of Sports Sciences                    |
| 8       | Interrelationships among invasive and non-invasive indicators of biological maturation in adolescent male soccer players   | 2012 | 48 | 107 | 44.86 | Journal of Sports Sciences                    |
| 9       | Biological maturation, morphology, fitness, and motor coordination as part of a selection strategy in the search for international youth soccer players (age 15–16 years)  | 2012 | 46 | 115 | 40    | Journal of Sports Sciences                    |
| 10      | A retrospective study on anthropometrical, physical fitness, and motor coordination characteristics that influence dropout, contract status, and first-team playing time in high-level soccer players aged eight to eighteen years | 2015 | 45 | 79  | 56.96 | Journal of Strength and Conditioning Research |

LC: local citations; GC: global citations.

and other devices to monitor heart rate<sup>64</sup> and other relevant athletic indicators<sup>65</sup> during small-side games,<sup>66</sup> which refer to soccer matches played with fewer players on a smaller field, promoting enhanced player involvement, skill development, and tactical understanding.

Cluster 4 has fewer nodes, compared to other clusters. However, in this cluster, the research is more focused on the theme of concussion in youth soccer.<sup>67</sup> Compared to other sports injuries, concussions are not as easily noticed and have a significantly negative impact on players' health. Concussion risk has been a hot topic in soccer, with several retired players falling ill from the effects of head impacts in recent years,<sup>68</sup> which has sparked public concern. Research hotspots in the field of youth soccer

regarding concussion include prevention and identification strategies,<sup>69</sup> clinical management and rehabilitation,<sup>70</sup> long-term consequences and monitoring,<sup>71</sup> education and awareness,<sup>72</sup> and policy development.<sup>73</sup>

### Theme map and thematic evolution

Thematic map was analyzed to understand the research themes (Figure 12). The results of the analysis show that studies related to youth soccer over the decade can be divided into five thematic clusters, and their labels for representative are children, performance, reliability, perception, and epidemiology. To better understand the trend of theme change, we analyzed the thematic evolution of the publications (Figure 13). There is an

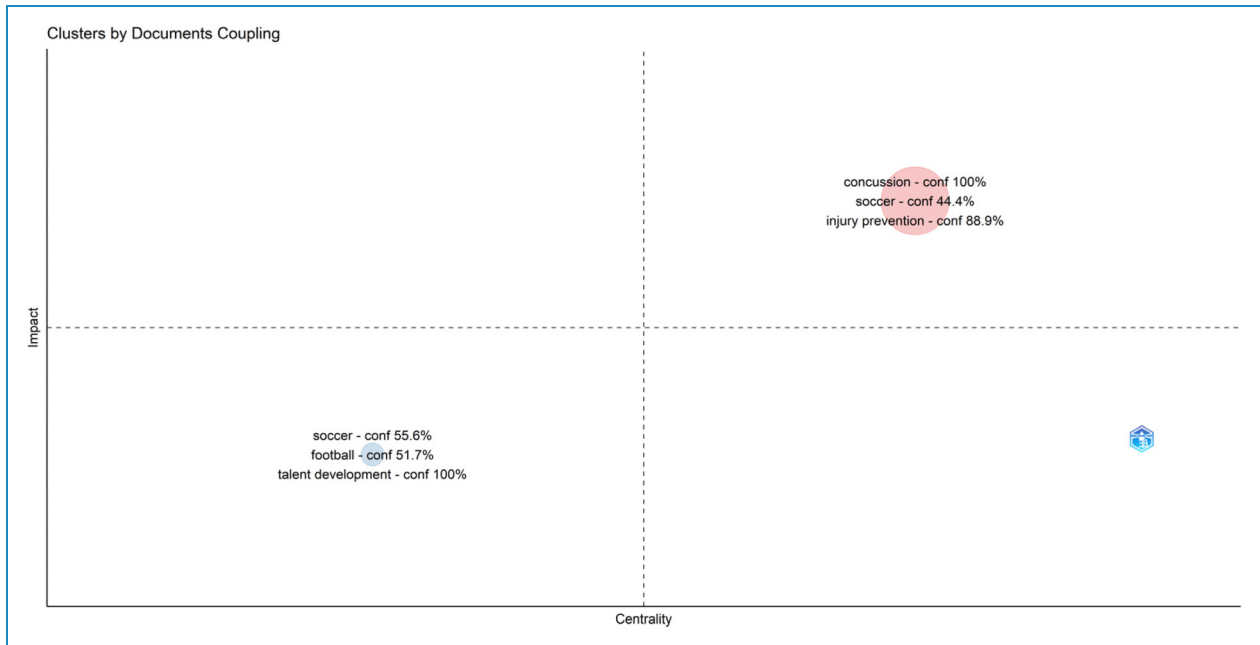


Figure 8. Documents coupling.  
Blue = cluster 1, red = cluster 2.



Figure 9. Word cloud of youth soccer.

increasing trend in the number of subjects, with themes developing from four in the first phase to six in 2021. It is noteworthy that themes of performance exist at all phases, and the proportion of studies shows an upward

trend. Research topics in this field have evolved from epidemiology, football, sport, and performance in 2012 to performance, reliability, concussion, risk-factors in 2021.

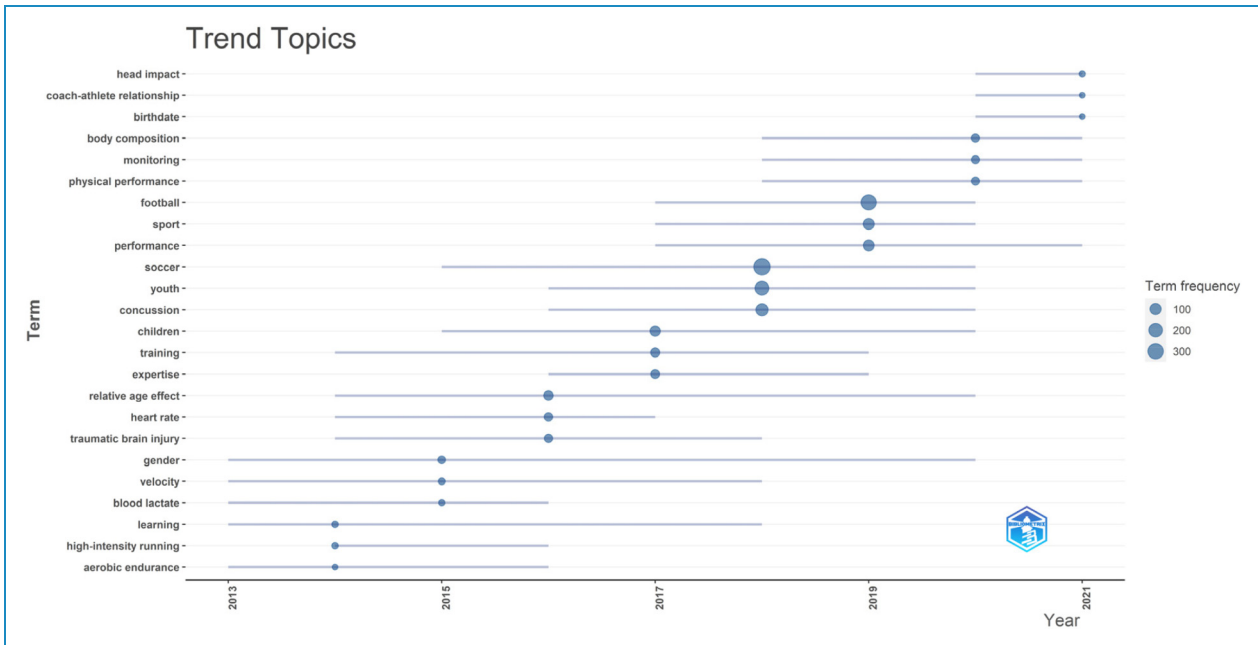


Figure 10. Trend topic of youth soccer.

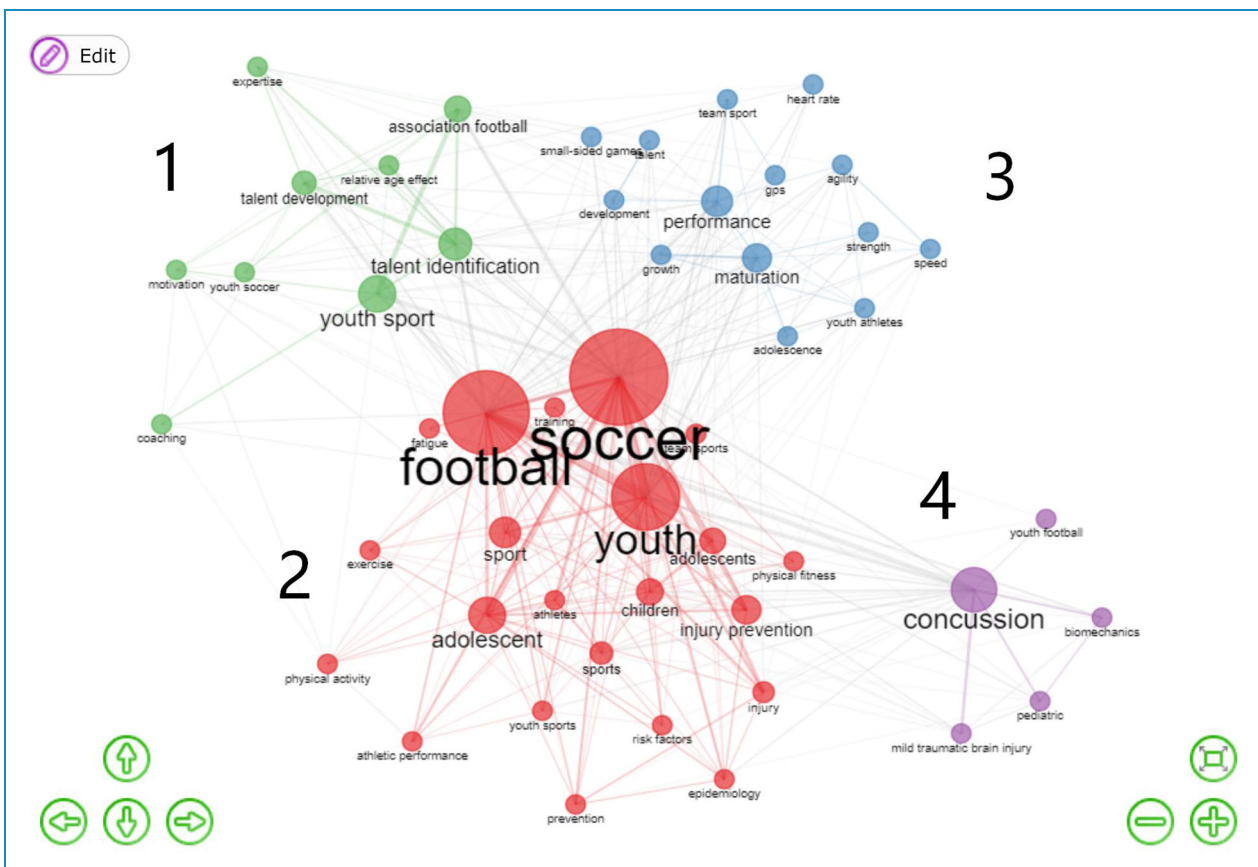


Figure 11. Keywords co-occurrences network.

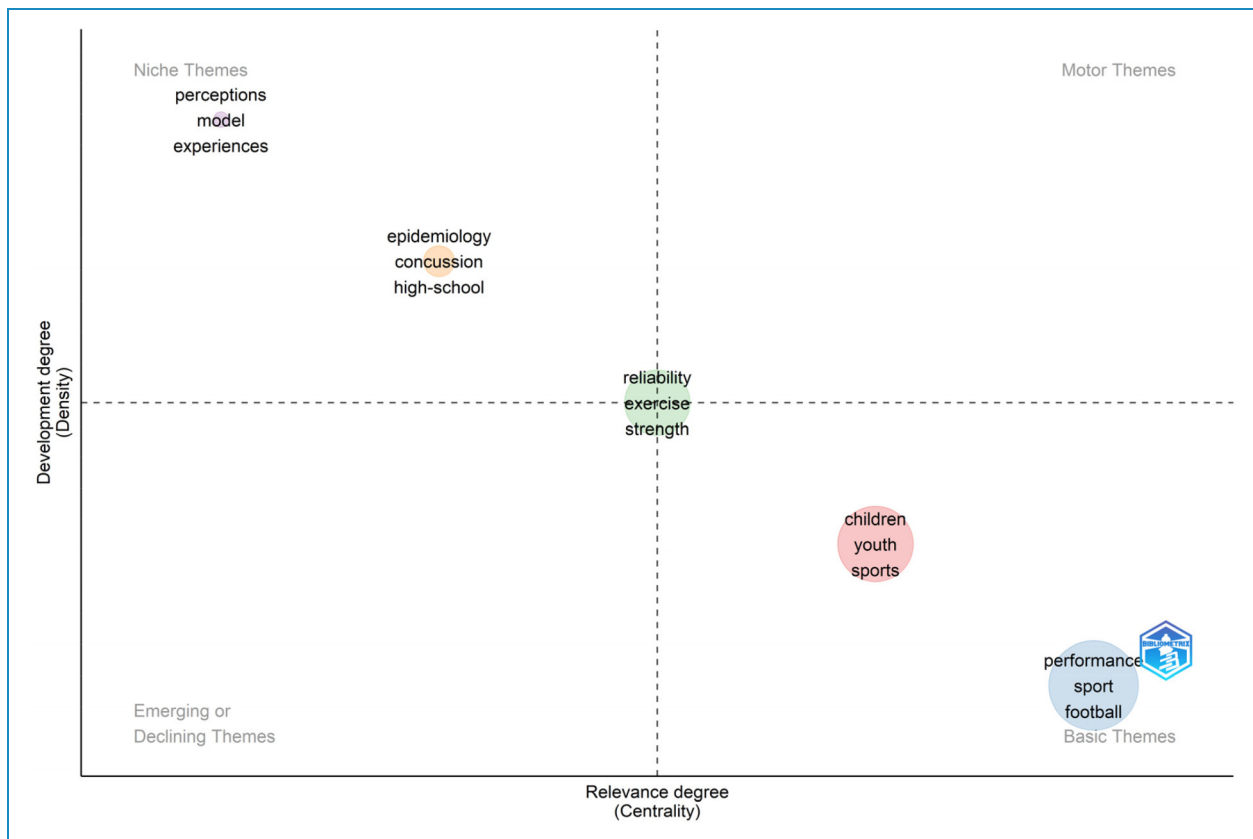


Figure 12. Thematic map.

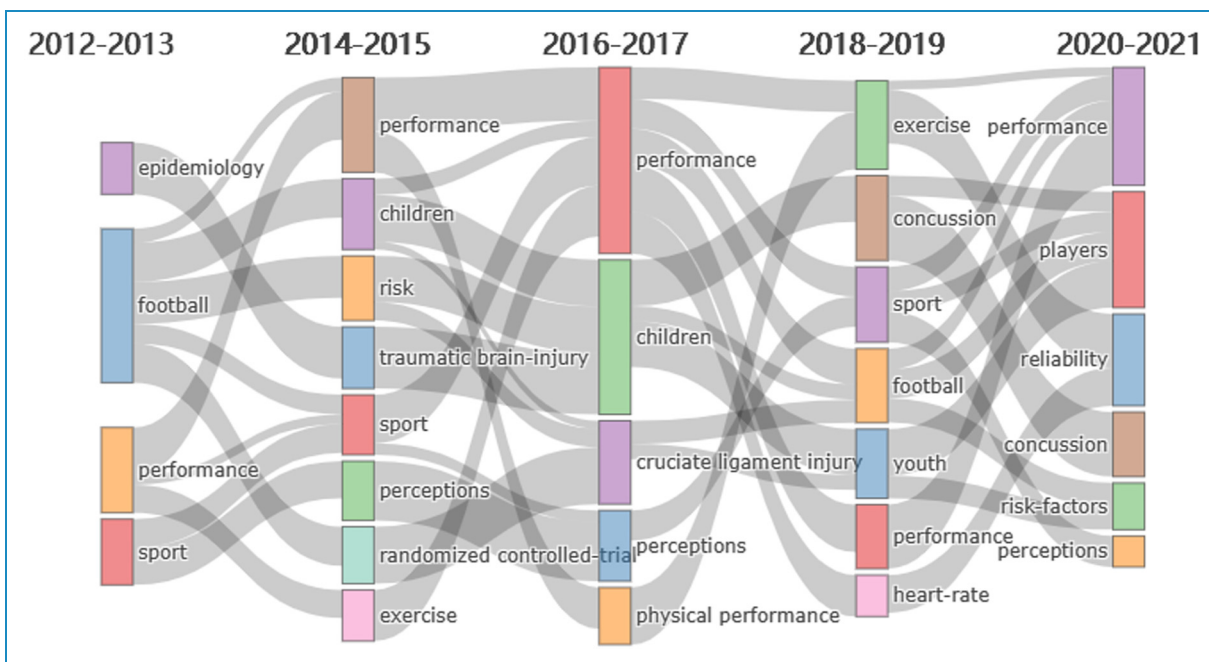


Figure 13. Thematic evolution.

## Conclusions

This bibliometric study examined global trends in scientific production on youth soccer from 2012 to 2021. The analysis revealed an increase in the number of sources, authors, and documents, indicating growing interest in youth soccer research. Collaboration among authors increased, with a significant international collaboration rate of 38.13%. Journals such as *Journal of Sports Sciences* and *Journal of Strength and Conditioning Research* played prominent roles in publishing youth soccer articles. Key authors like Lloyd RS, Malina RM, and Clemente FM contributed significantly to the field. Important research themes focused on talent identification, physical fitness, and sports injury prevention in youth soccer players. The findings highlight the international scope and evolving research topics within the field of youth soccer. Future research in this area could explore emerging areas such as head impact, coach-athlete relationships, relative age effects, body composition, monitoring, and performance, further advancing our understanding of youth soccer and its implications for player development and injury prevention.

## Limitations

This study has several limitations that should be acknowledged. Firstly, the literature search was conducted using the WoS database, which may result in the omission of relevant publications not indexed in WoS. Although efforts were made to minimize this limitation by using one of the most influential databases, it is possible that some relevant studies were not included in the analysis. Secondly, it is important to note that the term “youth” is commonly used in international journal publications to refer to the age group of interest in youth soccer research. However, the age range associated with the term “youth” varies across different countries and contexts. This poses challenges in accurately defining the age range in this study. It is recommended that future research on youth soccer clearly report the specific age range under investigation to ensure consistency and comparability across studies. Despite these limitations, the findings of this study provide valuable insights into the trends and research themes in the field of youth soccer. The limitations do not undermine the overall assessment of the trends or the implications for future research in this area.

**Acknowledgements:** The authors thank the National Social Science Foundation of China for funding this study.

**Contributorship:** Bo Gong proposed the research design, Bo Liu conducted the article, and Changjing Zhou and Haowei Ma proposed revisions to the paper. All authors participated in the

revision and finalization of the manuscript, and read and agreed to the published version of the manuscript.

**Consent statement:** All contributing authors of this manuscript have given the submission consent. In addition, patient consent is not necessary for this manuscript, as it is a bibliometric study.

**Declaration of conflicting interests:** The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

**Ethical approval:** Not applicable, because this article does not contain any studies with human or animal subjects.

**Funding:** The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the National Social Science Foundation of China (grant number 20BTY089).

**Guarantor:** BL.

**ORCID ID:** Bo Liu  <https://orcid.org/0000-0002-3869-2845>

## References

1. Terekli S and Cobanoglu HO. Developing economic values in football: example of Turkish Football Federation. *Open Access Library J* 2018; 05: 1–14.
2. Fifa C. FIFA Big Count 2006: 270 million people active in football. *FIFA Commun Div Inf Serv* 2007; 31: 1.
3. Allison R and Pope S. Becoming fans: socialization and motivations of fans of the England and US women’s national football teams. *Sociol Sport J* 2021; 39: 287–297.
4. Lago U, Simmons R and Szymanski S. *The financial crisis in European football: an introduction*. Thousand Oaks, CA: Sage Publications Sage CA, 2006, pp. 3–12.
5. FIFA. FIFA global transfer report 2021. 2022. <https://digitalhub.fifa.com/m/2b542d3b011270ff/original/FIFA-Global-Transfer-Report-2021-2022-indd.pdf>
6. De Bosscher V. *The global sporting arms race: An international comparative study on sports policy factors leading to international sporting success*. Oxford: Meyer & Meyer Verlag, 2008.
7. Bergkamp TL, Niessen ASM, den Hartigh RJ, et al. Methodological issues in soccer talent identification research. *Sports Med* 2019; 49: 1317–1335.
8. McCalman W, Crowley-McHattan ZJ, Fransen J, et al. Skill assessments in youth soccer: a scoping review. *J Sports Sci* 2022; 40: 667–695.
9. Reeves MJ and Roberts SJ. A bioecological perspective on talent identification in junior-elite soccer: a Pan-European perspective. *J Sports Sci* 2020; 38: 1259–1268.
10. Helsen WF, Starkes JL and Hodges NJ. Team sports and the theory of deliberate practice. *J Sport Exerc Psychol* 1998; 20: 12–34.
11. Kelly AL and Williams CA. Physical characteristics and the talent identification and development processes in male

- youth soccer: a narrative review. *Strength Condition J* 2020; 42: 15–34.
12. Williams AM and Reilly T. Talent identification and development in soccer. *J Sports Sci* 2000; 18: 657–667.
  13. García-Rubio J, García-Vallejo A, Arenas-Pareja MdlÁ, et al. From junior to elite in soccer: exploring the relative age effect and talent selection in Spanish youth national teams. *Children* 2022; 9: 1543.
  14. Helsen WF, Van Winckel J and Williams AM. The relative age effect in youth soccer across Europe. *J Sports Sci* 2005; 23: 629–636.
  15. Kucera KL, Marshall SW, Kirkendall DT, et al. Injury history as a risk factor for incident injury in youth soccer. *Br J Sports Med* 2005; 39: 462–462.
  16. Pfitrman D, Herbst M, Ingelfinger P, et al. Analysis of injury incidences in male professional adult and elite youth soccer players: a systematic review. *J Athl Train* 2016; 51: 410–424.
  17. Murr D, Larkin P and Höner O. Decision-making skills of high-performance youth soccer players: validating a video-based diagnostic instrument with a soccer-specific motor response. *German J Exerc Sport Res* 2021; 51: 102–111.
  18. Aria M and Cuccurullo C. Bibliometrix: an R-tool for comprehensive science mapping analysis. *J Informetr* 2017; 11: 959–975.
  19. Daim TU, Rueda G, Martin H, et al. Forecasting emerging technologies: use of bibliometrics and patent analysis. *Technol Forecast Soc Change* 2006; 73: 981–1012.
  20. Sharifi A, Simangan D and Kaneko S. Three decades of research on climate change and peace: a bibliometrics analysis. *Sustainability Sci* 2021; 16: 1079–1095.
  21. Madani F and Weber C. The evolution of patent mining: applying bibliometrics analysis and keyword network analysis. *World Pat Inf* 2016; 46: 32–48.
  22. Jiménez-García M, Ruiz-Chico J, Peña-Sánchez AR, et al. A bibliometric analysis of sports tourism and sustainability (2002–2019). *Sustainability* 2020; 12: 2840.
  23. Chen C. Citespace II: detecting and visualizing emerging trends and transient patterns in scientific literature. *J Am Soc Inf Sci Technol* 2006; 57: 359–377.
  24. Van Eck NJ and Waltman L. Citation-based clustering of publications using CitNetExplorer and VOSviewer. *Scientometrics* 2017; 111: 1053–1070.
  25. Cobo MJ, López-Herrera AG, Herrera-Viedma E, et al. SciMAT: a new science mapping analysis software tool. *J Am Soc Inf Sci Technol* 2012; 63: 1609–1630.
  26. Ahmi A. Bibliometric Analysis using R for Non-Coders: A practical handbook in conducting bibliometric analysis studies using Biblioshiny for Bibliometrix R package. 2022.
  27. Maia SC, de Benedicto GC, do Prado JW, et al. Mapping the literature on credit unions: a bibliometric investigation grounded in Scopus and Web of Science. *Scientometrics* 2019; 120: 929–960.
  28. Saleem F, Khattak A, Ur Rehman S, et al. Bibliometric analysis of green marketing research from 1977 to 2020. *Publications* 2021; 9: 1.
  29. Small H. Co-citation in the scientific literature: a new measure of the relationship between two documents. *J Am Soc Inf Sci* 1973; 24: 265–269.
  30. Callon M, Courtial JP and Laville F. Co-word analysis as a tool for describing the network of interactions between basic and technological research: the case of polymer chemistry. *Scientometrics* 1991; 22: 155–205.
  31. Sun Y, Barber R, Gupta M, et al. Co-author relationship prediction in heterogeneous bibliographic networks. In: 2011 International Conference on Advances in Social Networks Analysis and Mining, 2011, pp. 121–128. IEEE.
  32. Rodríguez-Soler R, Uribe-Toril J and Valenciano JDP. Worldwide trends in the scientific production on rural depopulation, a bibliometric analysis using bibliometrix R-tool. *Land Use Policy* 2020; 97: 104787.
  33. Bessem B, De Bruijn M and Nieuwland W. The ECG of high-level junior soccer players: comparing the ESC vs the Seattle criteria. *Br J Sports Med* 2015; 49: 1000–1006.
  34. Unnithan VB, Rowland TW, George K, et al. Left ventricular function during exercise in trained pre-adolescent soccer players. *Scand J Med Sci Sports* 2018; 28: 2330–2338.
  35. Koerte IK, Nichols E, Tripodis Y, et al. Impaired cognitive performance in youth athletes exposed to repetitive head impacts. *J Neurotrauma* 2017; 34: 2389–2395.
  36. Negra Y, Chaabene H, Fernandez-Fernandez J, et al. Short-term plyometric jump training improves repeated-sprint ability in prepubertal male soccer players. *J Strength Condition Res* 2020; 34: 3241–3249.
  37. Bianchi M, Coratella G, Dello Iacono A, et al. Comparative effects of single vs. double weekly plyometric training sessions on jump, sprint and COD abilities of elite youth football players. *J Sports Med Phys Fitness* 2018; 59: 910–915.
  38. Romann M and Fuchslocher J. Assessment of skeletal age on the basis of DXA-derived hand scans in elite youth soccer. *Res Sports Med* 2016; 24: 185–196.
  39. Sevil-Serrano J, Abós Á, Diloy-Peña S, et al. The influence of the coach's autonomy support and controlling behaviours on motivation and sport commitment of youth soccer players. *Int J Environ Res Public Health* 2021; 18: 8699.
  40. Wagnsson S, Stenling A, Gustafsson H, et al. Swedish Youth football players' attitudes towards moral decision in sport as predicted by the parent-initiated motivational climate. *Psychol Sport Exerc* 2016; 25: 110–114.
  41. Malina RM, Coelho-e-Silva MJ, Martinho DV, et al. Observed and predicted ages at peak height velocity in soccer players. *Plos one* 2021; 16: e0254659.
  42. Figueiredo AJ, Coelho-e-Silva MJ, Cumming SP, et al. Relative age effect: characteristics of youth soccer players by birth quarter and subsequent playing status. *J Sports Sci* 2019; 37: 677–684.
  43. Smith MR, Coutts AJ, Merlini M, et al. Mental fatigue impairs soccer-specific physical and technical performance. *Med Sci Sports Exercise* 2016; 48: 267–276.
  44. Larruskain J, Lekue JA, Diaz N, et al. A comparison of injuries in elite male and female football players: a five-season prospective study. *Scand J Med Sci Sports* 2018; 28: 237–245.
  45. Lloyd RS and Oliver JL. The youth physical development model: a new approach to long-term athletic development. *Strength Condition J* 2012; 34: 61–72.
  46. Read PJ, Oliver JL, Myer GD, et al. Utility of the anterior reach Y-BALANCE test as an injury risk screening tool in elite male youth soccer players. *Phys Ther Sport* 2020; 45: 103–110.
-

47. Cumming SP, Lloyd RS, Oliver JL, et al. Bio-banding in sport: applications to competition, talent identification, and strength and conditioning of youth athletes. *Strength Condition J* 2017; 39: 34–47.
  48. Lloyd RS, Oliver JL, Myer GD, et al. Seasonal variation in neuromuscular control in young male soccer players. *Phys Ther Sport* 2020; 42: 33–39.
  49. Sanchez-Sanchez J, Ramirez-Campillo R, Carretero M, et al. Soccer small-sided games activities vary according to the interval regime and their order of presentation within the session. *J Hum Kinet* 2018; 62: 167–175.
  50. Ramirez-Campillo R, Alvarez C, García-Pinillos F, et al. Optimal reactive strength index: is it an accurate variable to optimize plyometric training effects on measures of physical fitness in young soccer players? *J Strength Condition Res* 2018; 32: 885–893.
  51. Rodríguez-Fernández A, Sanchez-Sanchez J, Ramirez-Campillo R, et al. Relationship between repeated sprint ability, aerobic capacity, intermittent endurance, and heart rate recovery in youth soccer players. *J Strength Condition Res* 2019; 33: 3406–3413.
  52. Sarmento H, Anguera MT, Pereira A, et al. Talent identification and development in male football: a systematic review. *Sports Med* 2018; 48: 907–931.
  53. Steffen K, Emery CA, Romiti M, et al. High adherence to a neuromuscular injury prevention programme (FIFA 11+) improves functional balance and reduces injury risk in Canadian youth female football players: a cluster randomised trial. *Br J Sports Med* 2013; 47: 794–802.
  54. Cobo MJ, López-Herrera AG, Herrera-Viedma E, et al. An approach for detecting, quantifying, and visualizing the evolution of a research field: a practical application to the Fuzzy Sets Theory field. *J Informetr* 2011; 5: 146–166.
  55. Williams AM, Ford PR and Drust B. Talent identification and development in soccer since the millennium. *J Sports Sci* 2020; 38: 1199–1210.
  56. Meylan C, Cronin J, Oliver J, et al. Talent identification in soccer: the role of maturity status on physical, physiological and technical characteristics. *Int J Sports Sci Coach* 2010; 5: 571–592.
  57. Kelly AL, Wilson MR, Gough LA, et al. A longitudinal investigation into the relative age effect in an English professional football club: exploring the ‘underdog hypothesis’. *Sci Med Football* 2020; 4: 111–118.
  58. Towlson C, MacMaster C, Parr J, et al. One of these things is not like the other: time to differentiate between relative age and biological maturity selection biases in soccer? *Sci Med Football* 2022; 6: 273–276.
  59. Sanchez-Sanchez J, Sanchez M, Hernandez D, et al. Fatigue in u12 soccer-7 players during repeated 1-day tournament games—A pilot study. *J Strength Condition Res* 2019; 33: 3092–3097.
  60. Tomáš M, František Z, Lucia M, et al. Profile, correlation and structure of speed in youth elite soccer players. *J Hum Kinet* 2014; 40: 149–159.
  61. Nguyen A-D, Zuk EF, Baellow AL, et al. Longitudinal changes in hip strength and range of motion in female youth soccer players: implications for ACL injury, a pilot study. *J Sport Rehabil* 2017; 26: 358–364.
  62. Negra Y, Chaabene H, Amara S, et al. Evaluation of the Illinois change of direction test in youth elite soccer players of different age. *J Hum Kinet* 2017; 58: 215–224.
  63. Reilly B, Akubat I, Lyons M, et al. Match-play demands of elite youth Gaelic football using global positioning system tracking. *J Strength Condition Res* 2015; 29: 989–996.
  64. Ørntoft C, Larsen MN, Andersen TB, et al. Technical actions, heart rate, and locomotor activity in 7v7 and 8v8 games for female youth soccer players. *J Strength Condition Res* 2016; 30: 3298–3303.
  65. Fernandes-da-Silva J, Castagna C, Teixeira AS, et al. The peak velocity derived from the Carminatti Test is related to physical match performance in young soccer players. *J Sports Sci* 2016; 34: 2238–2245.
  66. Silva P, Esteves P, Correia V, et al. Effects of manipulations of player numbers vs. Field dimensions on inter-individual coordination during small-sided games in youth football. *Int J Perform Anal Sport* 2015; 15: 641–659.
  67. Wahlquist VE and Kaminski TW. Purposeful heading in youth soccer: a review. *Sports Med* 2021; 51: 51–64.
  68. Kerr ZY, Marshall SW, Harding HP Jr, et al. Nine-year risk of depression diagnosis increases with increasing self-reported concussions in retired professional football players. *Am J Sports Med* 2012; 40: 2206–2212.
  69. Indharty RS, Siahaan AM, Susanto M, et al. Prevention of sports-related concussion in soccer: a comprehensive review of the literature. *Ann Med Surg* 2023; 85: 365.
  70. Kirkwood MW, Crossland MM, Howell DR, et al. A longitudinal investigation of symptom recovery following concussion in youth soccer. *J Pediatr* 2020; 220: 207–213. e202.
  71. Vedung F, Hänni S, Tegner Y, et al. Concussion incidence and recovery in Swedish elite soccer—prolonged recovery in female players. *Scand J Med Sci Sports* 2020; 30: 947–957.
  72. Schatz P, Corcoran M, Kontos AP, et al. Youth soccer parents’ perceptions of long-term effects of concussion. *Dev Neuropsychol* 2020; 45: 110–117.
  73. Shenouda C, Hendrickson P, Davenport K, et al. The effects of concussion legislation one year later—what have we learned: a descriptive pilot survey of youth soccer player associates. *PM&R* 2012; 4: 427–435.
-