

MAPPING FOOTBALL RESEARCH FOR SPORT EDUCATION: A BIBLIOMETRIC ANALYSIS OF THE DERGIPARK DATABASE

**MAPEAMENTO DE PESQUISAS SOBRE FUTEBOL PARA A EDUCAÇÃO
ESPORTIVA: UMA ANÁLISE BIBLIOMÉTRICA DA BASE DE DADOS DERGIPARK**

**MAPEO DE INVESTIGACIONES SOBRE FÚTBOL PARA LA EDUCACIÓN
DEPORTIVA: UN ANÁLISIS BIBLIOMÉTRICO DE LA BASE DE DATOS
DERGIPARK**

iD

Turhan TOROS¹

e-mail: turhantoros@yahoo.com

iD

Emre SERIN²

e-mail: emreserin@mersin.edu.tr

iD

Arif Mert OZKAN³

e-mail: ozkanarifmert@gmail.com

iD

Mehmet BAYANSALDUZ⁴

e-mail: mehmetbayansalduz@topkapi.edu.tr

iD

Ali Emre EREN⁵

e-mail: emre_01@gmail.com

iD

Mihriay MUSA⁶

e-mail: mihriaymusa@topkapi.edu.tr

iD

Erkan GÜLGÖSTEREN⁷

e-mail: egulgosteren@mersin.edu.tr

How to reference this paper:

Toros, T., Serin, E., Ozkan, A. M., Bayansalduz, M., Eren, A. E., Musa, M., & Gülgösteren, E. (2026). Mapping football research for sport education: a bibliometric analysis of the Dergipark database. *Revista on line de Política e Gestão Educacional*, 30(esp2), e026086. <https://doi.org/10.22633/rpge.v30iesp2.21394>



| Submitted: 24/04/2026

| Revisions required: 28/06/2026

| Approved: 17/06/2026

| Published: 01/08/2026

¹ Mersin University, Mersin – Turkey. Faculty of Sport Sciences.

² Mersin University, Mersin – Turkey. Faculty of Sport Sciences.

³ Mersin University, Mersin – Turkey. Faculty of Sport Sciences.

⁴ Istanbul Topkapi University, Istanbul – Turkey. Faculty of Sport Sciences.

⁵ Mersin University, Mersin – Turkey. Faculty of Sport Sciences.

⁶ Istanbul Topkapi University, Istanbul – Turkey. Faculty of Sport Sciences.

⁷ Mersin University, Mersin – Turkey. Faculty of Sport Sciences.

Editor: Prof. Dr. Sebastião de Souza Lemes
Deputy Executive Editor: Prof. Dr. José Anderson Santos Cruz

ABSTRACT: A corpus of 1,411 articles indexed in DergiPark and published between 1967 and 2025 was compiled through searches using the terms “futbol,” “football,” and “soccer.” The bibliometric analysis combined performance indicators and science mapping. An exponential growth in research output was observed, with 46.92% of the articles published between 2021 and 2025. The studies were distributed across 343 journals and 110 disciplinary categories, highlighting the multidisciplinary nature of football as a research topic. The proportion of articles published in English increased from less than 5% before 2010 to 35.8% between 2021 and 2025, indicating growing internationalization. Sports Medicine predominated (38.31%), although its relative share declined as Sports Training, Sports Activity Management, Exercise Sciences, and Educational Studies expanded. The findings point to the need to broaden methodological approaches, strengthen interdisciplinary research teams, and adopt bilingual publication strategies to enhance the educational and international impact of Turkish research on football.

KEYWORDS: Football. Bibliometric analysis. DergiPark. Sport education.

RESUMO: Foi reunido um corpus de 1.411 artigos indexados na DergiPark, publicados entre 1967 e 2025, mediante buscas pelos termos “futbol”, “football” e “soccer”. A análise bibliométrica combinou indicadores de desempenho e mapeamento científico. Observou-se crescimento exponencial da produção, com 46,92% dos artigos publicados entre 2021 e 2025. Os estudos foram distribuídos em 343 periódicos e 110 categorias disciplinares, evidenciando o caráter multidisciplinar do futebol como objeto de pesquisa. A participação de artigos em inglês aumentou de menos de 5% antes de 2010 para 35,8% entre 2021 e 2025, indicando crescente internacionalização. A Medicina Esportiva predominou (38,31%), embora tenha reduzido sua participação relativa diante da expansão de Treinamento Esportivo, Gestão de Atividades Esportivas, Ciências do Exercício e Estudos em Educação. Os resultados apontam a necessidade de ampliar abordagens metodológicas, fortalecer equipes interdisciplinares e adotar estratégias bilingües para ampliar o impacto educacional e internacional da produção turca sobre futebol.

PALAVRAS-CHAVE: Futebol. Análise bibliométrica. DergiPark. Educação esportiva.

RESUMEN: Se reunió un corpus de 1.411 artículos indexados en DergiPark y publicados entre 1967 y 2025, mediante búsquedas de los términos “futbol”, “football” y “soccer”. El análisis bibliométrico combinó indicadores de desempeño y mapeo científico. Se observó un crecimiento exponencial de la producción científica, con un 46,92 % de los artículos publicados entre 2021 y 2025. Los estudios se distribuyeron en 343 revistas y 110 categorías disciplinares, lo que evidencia el carácter multidisciplinario del fútbol como objeto de investigación. La proporción de artículos publicados en inglés aumentó de menos del 5 % antes de 2010 al 35,8 % entre 2021 y 2025, lo que indica una creciente internacionalización. La Medicina Deportiva predominó (38,31 %), aunque redujo su participación relativa ante la expansión del Entrenamiento Deportivo, la Gestión de Actividades Deportivas, las Ciencias del Ejercicio y los Estudios Educativos. Los resultados señalan la necesidad de ampliar los enfoques metodológicos, fortalecer los equipos de investigación interdisciplinarios y adoptar estrategias de publicación bilingüe para aumentar el impacto educativo e internacional de la producción turca sobre el fútbol.

PALABRAS CLAVE: Fútbol. Análisis bibliométrico. DergiPark. Educación deportiva.

INTRODUCTION

Bibliometric studies are statistical methods that quantitatively assess scientific production on a specific topic and analyze general trends across publications, researchers, journals, institutions, countries or regions, and keywords (Alkan et al., 2025; Kaya & Algin, 2022). Such studies serve a dual function: they help identify knowledge gaps and stimulate new research directions, while also guiding researchers in evaluating the scope and orientation of contributions within a given field. For scholars preparing new work, bibliometric mapping reveals which types of studies have been pursued and where unexamined questions remain, enabling more informed decisions about research design and topic selection (Jacobs, 2001). In interdisciplinary fields such as sport sciences, where methodological traditions and thematic priorities vary widely across sub-disciplines, bibliometric analysis offers a particularly valuable lens for examining both methodological diversity and thematic concentration. Passas (2024) identifies seven main steps in bibliometric research: defining research objectives, literature search and data collection, data cleaning, selection of techniques, data analysis, visualization, and interpretation. Bibliometric analysis is especially suited to broad reviews of large datasets, where it outperforms systematic literature reviews in scope and objectivity (Algin, 2024; Öztürk et al., 2024; Passas, 2024), while systematic reviews remain more appropriate for narrow, in-depth examination of specific research questions.

Sport, derived from the Latin *deportare* (play or amusement) through French *desport*, encompasses all physical activities undertaken for purposes of play and competition (Polat & Bingöl, 2022; Toros, 2025). Beyond its physical benefits, regular sport participation has been linked to higher extraversion, conscientiousness, and agreeableness (Tazegül, 2014; Toros & Şekeroğlu, 2026). Among sport disciplines, football holds particular cultural prominence worldwide and occupies a central position in Turkish society. The English term *association football* distinguishes the sport from other football codes, while the informal abbreviation *soccer* emerged in the 1880s and remains current in several English-speaking countries. In Turkish, the standard term is *futbol*, derived from the English compound of *foot* and *ball*, though the calque *ayak topu* also exists. Such terminological variation informed the present study's search strategy, which employed all three terms ("futbol," "football," and "soccer") to ensure broad coverage of the relevant literature.

In educational terms, football should not be understood only as a competitive sport or a professional entertainment industry. It also functions as curricular content in physical education, as a pedagogical resource for the development of motor, social, cognitive, and affective

competencies, and as a field of professional training for teachers, coaches, and sport managers. Research on football may therefore inform decisions about curriculum design, pedagogical sequencing, inclusion, school sport, youth development, and the continuing education of professionals who work in educational and sport institutions. This educational dimension is consistent with the broader literature on physical education and sport pedagogy, which emphasizes that school sport and physical education can contribute to learning, participation, cooperation, health literacy, and social development when they are intentionally planned and critically mediated (Armour, 2011; Bailey et al., 2009; Kirk, 2010; UNESCO, 2015).

The concept of bibliometrics itself derives from the Greek and Latin roots *biblion* (book) and *metrikos* (measurement), and the term was first introduced by Pritchard (1969), who defined it as “the application of mathematical and statistical methods to books and other media of communication.” Pritchard (1969) argued that bibliometrics functions as a guide within the written communication process and enables the study of a scientific discipline’s nature and development. Sengupta (1992) elaborated on this definition by locating bibliometrics within a family of related measurement disciplines, including informetrics, scientometrics, and librametrics. Diodato (1994, as cited in Köseoğlu et al., 2016) extended the conceptual framework further, describing bibliometrics as a method that permits the quantitative examination of scientific productivity, in which parameters such as citations, authors, keywords, topics, and methods are statistically categorized, described, and evaluated. It also enables the collection and interpretation of historical usage statistics for research published in books and journals (Raisig, 1962), and Roy and Basak (2013) characterize it as a discipline that employs quantitative methods to assess the scientific communication process through written documents. Through bibliometric analysis, the development of a body of literature can be observed and quantitative examinations can be conducted to contribute to the relevant field (Okubo, 1997; Serin et al., 2025). Put more broadly, bibliometrics draws on mathematical and statistical methods to analyze a field’s development process and to investigate who researches what, where, and how within a scientific framework (Gulgösteren et al., 2025; Sanchez et al., 2017; Toros, 2025).

Recent methodological advances have formalized bibliometric research into two main analytical categories: performance analysis, which evaluates the productivity and impact of scientific constituents, and science mapping, which reveals relationships between research constituents through network analysis (Donthu et al., 2021; Öztürk et al., 2024). The parameters typically examined in bibliometric studies include article topic, year of publication, research

approach, author titles and count, institutional affiliations, citation counts, reference counts, publication language, journal distribution, the conceptual or empirical nature of articles, sub-fields, research areas, publication volume, authorship patterns, keywords, and the distribution of journals and authors by disciplinary affiliation (Cevizkaya et al., 2014; Çiçek & Kozak, 2012; Özel & Kozak, 2012; Şakar & Cerit, 2013; Sökmen & Özkanlı, 2018; Yetiş & Çokal, 2018). Bibliometric studies targeting graduate theses have been conducted in fields such as gastronomy, destination management, recreation, tourist guidance, sustainable tourism, general tourism, and tourism geography (Ayar & Koç, 2018; Aydın & Aksöz, 2019; Bao, 2002; Meyer-Arendt, 2000; Şahin & Acun, 2015; Sünnetçioğlu et al., 2017; Tayfun et al., 2016). Within sport sciences, however, football-specific bibliometric mapping in the Turkish context remains scarce. International bibliometric work on football research exists, but such studies draw primarily from Scopus and Web of Science and therefore underrepresent Turkish-language scholarship published through national platforms.

Such a gap creates an opportunity both to increase the visibility of domestic academic production and to identify research vacancies that future studies might fill, especially in relation to physical education, sport pedagogy, teacher education, and educational management. DergiPark, with its open-access policy and its role as the primary hosting infrastructure for Turkish academic journals under Tübitak-Ulakbim, constitutes an especially suitable source for bibliometric examination of football-themed publications in Turkey. The present study analyzes all research articles containing the words “futbol,” “football,” or “soccer” in their titles or keywords published in the DergiPark database, examining publication trends, journal distribution, language patterns, disciplinary structure, thematic content, authorship profiles, and citation impact through bibliometric criteria. Following the analytical framework proposed by Donthu et al. (2021) and Öztürk et al. (2024), the study combines performance analysis with science mapping to capture both the productivity profile and the thematic structure of the field. The interpretation also examines how football research can support knowledge production in sport education and how the observed disciplinary distribution reveals opportunities for stronger articulation between sport sciences and education. The analysis pursues eight research questions:

- *RQ1: How has the volume of football-related publications in DergiPark changed over time?*
- *RQ2: Which journals concentrate the highest share of football-related output?*

- *RQ3: What is the distribution of publication language, and how has it changed over time?*
- *RQ4: Which disciplinary subject categories characterize football research in DergiPark?*
- *RQ5: What thematic patterns emerge from natural language processing of article titles?*
- *RQ6: What authorship patterns characterize the corpus?*
- *RQ7: Which articles have generated the highest citation impact?*
- *RQ8: What educational implications can be derived from the disciplinary and thematic mapping of football research in DergiPark?*

Findings from this analysis are expected to contribute to the development of academic production in the field of football and to encourage interdisciplinary collaboration by providing a structured baseline against which future bibliometric work can measure change. More specifically, the study seeks to support the educational use of bibliometric evidence by identifying which parts of the football literature can inform physical education, sport pedagogy, teacher education, and the institutional management of sport-related educational programs.

METHODOLOGY

Research design

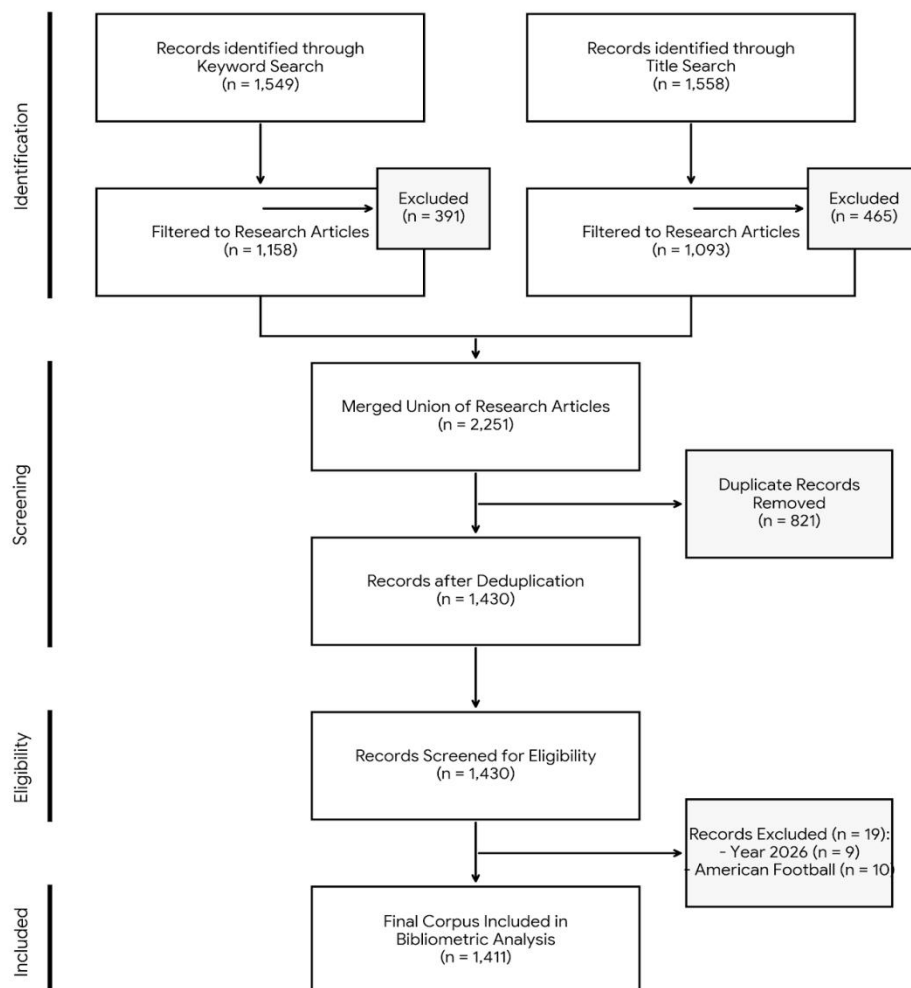
A bibliometric analysis design was adopted for this study, drawing on the frameworks proposed by Donthu et al. (2021) and Öztürk et al. (2024). Bibliometric analysis enables the quantitative characterization of recorded scholarly discourse (Schrader, 1981) and has been applied widely to map disciplinary development across the social sciences (Sanchez et al., 2017). Following Donthu et al.'s (2021) two-category typology, the analysis combined performance analysis (frequency distributions of publications, journals, languages, subject categories, authors, and citation counts) with science mapping (title-based bigram extraction and word co-occurrence network analysis) to capture both the productivity profile and the thematic structure of football research in DergiPark. Passas (2024) identifies seven procedural steps for bibliometric research: defining objectives, literature search, data cleaning, technique selection, analysis, visualization, and interpretation. Each of these steps is documented in the subsections that follow.

Data source and population

The study population comprised all research articles indexed in the DergiPark electronic journal system, which operates under the TÜBİTAK National Academic Network and Information Centre (ULAKBİM). DergiPark was selected for two reasons: it is the most extensive open-access repository for Turkish-language academic journals, and its structured metadata fields permit systematic extraction of publication records.

Search strategy and sample

Figure 1.
PRISMA Flowchart of the Data Collection Process



Note. The Authors.

Data extraction and classification

Metadata for each article (title, DOI, author names, journal name, publication year, page range, subject category, primary language, and publication type) were exported through DergiPark's CSV download function, cleaned, and organized via Python. DergiPark's primary language metadata field was available for 1,182 of 1,411 articles. For the remaining 229 articles, language was verified manually through full-text inspection. All 229 were confirmed as Turkish. DergiPark assigns subject categories to articles based on journal-level classification. Subject data was available for 890 of 1,411 articles (63.1%). Articles without subject classification were concentrated in older publication years, reflecting incomplete metadata in retroactively indexed journals.

Thematic Analysis

Thematic patterns were examined through computational text analysis of article titles (Moral-Muñoz et al., 2020). English-language title text was extracted for 689 articles. Two techniques were applied.

Bigram frequency analysis: Titles were tokenized after removing stopwords and common methodological terms. Two-word phrases were extracted and ranked by frequency. Syntactic fragments arising from title grammar rather than substantive content were filtered manually.

Word co-occurrence network analysis: Title tokens were lemmatized to merge plural and singular forms. Co-occurrence was calculated as the number of titles in which two words appeared together. A network graph was constructed with nodes representing words appearing in 18 or more titles and edges representing co-occurrence frequencies of 5 or higher. Node size was scaled to word frequency, edge width to co-occurrence strength, and node color to weighted degree centrality.

Citation and Data Analysis

Citation impact was assessed through a Google Scholar search using the query site:dergipark.org.tr “futbol” OR “football” OR “soccer”, conducted in March 2026. The ten articles with the highest citation counts were identified and recorded.

Descriptive statistics (frequencies and percentages) were computed for each parameter. Five-year period aggregations were calculated for publication volume and language trends to

reveal temporal patterns that annual figures obscure. No inferential statistical tests were applied, consistent with the descriptive nature of the bibliometric design.

The interpretation of the bibliometric indicators also considered the educational meanings of the corpus. Particular attention was given to subject categories and title terms connected to Sports Training, Studies on Education, youth development, coaching, and professional preparation, since these strands are directly related to physical education curricula, pedagogical practice, and the organization of sport education programs.

RESULTS

A corpus of 1,411 research articles was assembled from two complementary DergiPark searches, merged, deduplicated, and screened for non-association football content. As summarized in Table 1, the dataset spans nearly six decades of publication activity from 1967 to 2025, though output before the mid-1990s was sparse. The volume of literature has accelerated significantly in recent years, culminating in a peak output in 2025 with 157 articles, which alone accounts for 11.13% of the total corpus.

Table 1.
Descriptive Profile of the Corpus

Parameter	Finding
Time span	1967 to 2025
Peak year	2025 ($n = 157$, 11.13%)
Language	Turkish 73.14%, English 26.65%, Other 0.21%
Top journal	SPORMETRE ($n = 102$)
Unique journals	343
Most common author count	Two-author teams (37.63%)
Unique authors	2,386
Subject categories	110 unique categories across 890 classified articles
Leading subject category	Sports Medicine ($n = 341$, 38.31% of classified articles)

Note. The Authors.

Linguistically, the literature is predominantly published in Turkish (73.14%), accompanied by a substantial secondary body of work in English (26.65%). These studies are distributed across a highly dispersed network of 343 unique journals, with *SPORMETRE* emerging as the most productive outlet, publishing 102 of the mapped articles. The dataset features contributions from 2,386 unique authors, indicating a broad base of academic interest.

Collaborative research is the standard, with two-author teams representing the most common authorship model (37.63%). In terms of formatting, the articles are relatively standardized in length, exhibiting a median of 11 and a mean of 12.5 pages. Finally, thematic classification was available for 890 articles across 110 unique subject categories. Among these, Sports Medicine heavily dominates the academic discourse, comprising 38.31% ($n = 341$) of all classified research in this domain. In the following paragraphs, we present detailed analyses of each of these parameters.

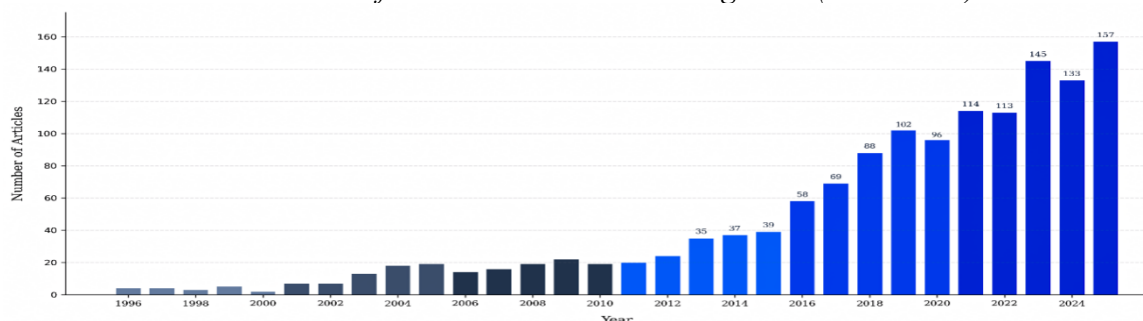
Publication Volume by Year (RQ1): Five-year period distribution of the 1,411 articles is presented in Table 2.

Table 2.
Distribution of Publications by Five-Year Period

Period	<i>n</i>	%
Pre-1996	9	0.64
1996-2000	18	1.28
2001-2005	64	4.54
2006-2010	90	6.38
2011-2015	155	10.98
2016-2020	413	29.27
2021-2025	662	46.92
Total	1,411	100.00

Note. The Authors. Nine articles published before 1996 span 1967, 1978, 1989, 1990, 1992, 1993, and 1994.

Publication volume followed a clear exponential trajectory. Output rose modestly from 18 articles in 1996-2000 to 90 in 2006-2010, then accelerated sharply: the 2011-2015 period produced 155 articles, 2016-2020 nearly tripled that figure to 413, and 2021-2025 added another 60% to reach 662. Nearly half the entire corpus (46.92%) was published in the most recent five-year window alone. Peak annual output occurred in 2025 ($n = 157$), followed by 2023 ($n = 145$) and 2024 ($n = 133$). A slight dip in 2020 ($n = 96$, down from 102 in 2019) coincides with the onset of the COVID-19 pandemic, though recovery was rapid, with output reaching 114 by 2021. Figure 2 displays the annual publication trend from 1996 to 2025 with five-year period totals.

Figure 2.*Annual Publication Volume of Football Research in DergiPark (1996-2025)*

Source. The Authors.

Publication Language (RQ3): Language distribution for the assembled corpus is reported in Table 3. Turkish-language articles comprised nearly three quarters of the dataset (73.14%), while English-language articles accounted for a substantial 26.65%.

Table 3.*Distribution of Publications by Language*

Language	<i>n</i>	%
Turkish	1,032	73.14
English	376	26.65
Bilingual (TR & EN)	2	0.14
Russian	1	0.07
Total	1,411	100.00

Note. The Authors.

To map the internationalization of the field over time, the English-language share was disaggregated by five-year periods (Table 4). Figure 3 visualizes the relationship between total publication volume and English-language share across these periods.

Table 4*English-Language Publication Share by Period*

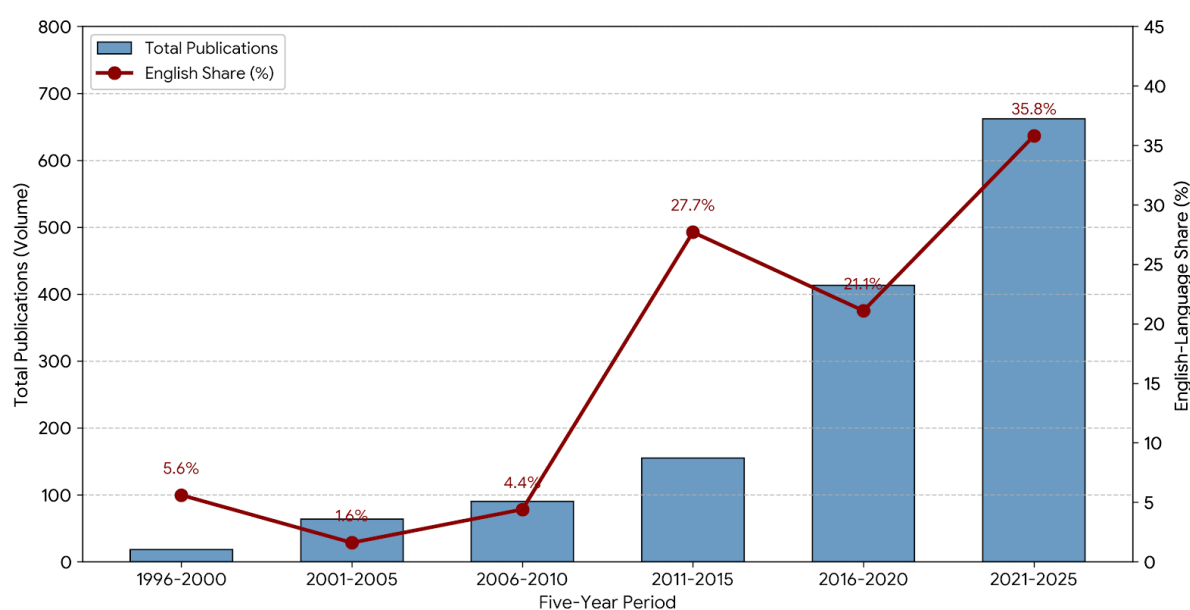
Period	Total	English	% English
1996-2000	18	1	5.6
2001-2005	64	1	1.6
2006-2010	90	4	4.4
2011-2015	155	43	27.7
2016-2020	413	87	21.1
2021-2025	662	237	35.8

Note. The Authors.

As illustrated in Figure 3, the English-language share remained below 5% through 2010. However, the subsequent decade marked a period of rapid internationalization, with the English share rising sharply to 27.7% in 2011–2015. Following a slight structural dip to 21.1% in 2016–2020, the English publication rate surged to its highest recorded level at 35.8% in the 2021–2025 period. Of particular note, more than half of all English-language articles in the entire corpus ($n = 237$ of 376) were published in this most recent five-year window, which shows a strong, contemporary pivot toward global academic engagement.

Figure 3.

Evolution of English-Language Publication Share (1996-2025)



Note. The Authors.

Disciplinary Distribution (RQ4): DergiPark assigns subject categories to articles based on journal-level classification. Of the 1,411 articles in the corpus, 890 (63.1%) carried subject category data. Table 5 presents the 15 most frequent categories within this classified subset.

Table 5.

Distribution of Articles by Subject Category ($n = 890$ articles with data)

Subject Category	<i>n</i>	%
Sports Medicine	341	38.31
Sports Training	110	12.36
Sports Science and Exercise (Other)	66	7.42
Sports Activity Management	59	6.63
Health Care Administration	42	4.72
Communication and Media Studies	19	2.13

Sports Science and Exercise	18	2.02
Business Administration	18	2.02
Physical Activity and Health	18	2.02
Exercise Physiology	16	1.80
Physical Fitness	16	1.80
Tourism (Other)	16	1.80
Studies on Education	15	1.69
Exercise and Sports Psychology	13	1.46
Sport and Exercise Psychology	11	1.24
Other categories ($n = 95$)	199	22.36

Note. The Authors. Percentages are calculated against the 890 articles with subject category data. Articles without subject classification were concentrated in older publication years.

Sports Medicine dominated the classified corpus with 341 articles (38.31%), more than three times the count of the next largest category, Sports Training ($n = 110$, 12.36%). Together, the top four categories (Sports Medicine, Sports Training, Sports Science and Exercise, and Sports Activity Management) accounted for 64.72% of classified articles. A total of 110 unique subject categories appeared in the data, confirming football's reach into disciplines well beyond exercise science. Communication and Media Studies ($n = 19$), Business Administration ($n = 18$), Tourism ($n = 16$), and Education ($n = 15$) each contributed measurable shares, while Finance, Sociology, Law, and Engineering appeared in the lower tail.

From an educational perspective, this distribution is important because it shows that football research in DergiPark is not limited to biomedical approaches. Sports Training, Sports Activity Management, Exercise Science, and Studies on Education together point to a body of work that can be mobilized in physical education, sport pedagogy, teacher education, and the planning of school or university sport programs. Although Studies on Education appears as a small category ($n = 15$; 1.69%), its presence demonstrates an existing interface between football research and educational debates. The relatively small size of this category also indicates a research gap: pedagogical, curricular, inclusive, and policy-oriented approaches to football remain underdeveloped when compared with studies centered on performance and sports medicine.

When disaggregated by period, the disciplinary profile shifted visibly. Sports Medicine accounted for 60.7% of classified articles in 2016-2020 but dropped to 25.4% in 2021-2025, not because it shrank in absolute terms but because Sports Training, Sports Activity Management, and Exercise Science grew rapidly around it. Table 6 displays the top five subject categories across three periods.

Table 6.*Top Five Subject Categories by Period*

2006-2015 (n = 24)	n	%
Sports Medicine	9	37.5
Health Care Administration	5	20.8
Communication and Media Studies	2	8.3
Business Administration	2	8.3
Other	6	25.0
2016-2020 (n = 219)	n	%
Sports Medicine	133	60.7
Health Care Administration	28	12.8
Communication and Media Studies	11	5.0
Studies on Education	8	3.7
Economics	6	2.7
2021-2025 (n = 714)	n	%
Sports Medicine	181	25.4
Sports Training	110	15.4
Sports Science and Exercise (Other)	66	9.2
Sports Activity Management	59	8.3
Sports Science and Exercise	18	2.5

Note. The Authors.

Disciplinary diversification accelerated in the most recent period. While Sports Medicine remained the single largest category in 2021-2025, its relative share fell by more than half compared to 2016-2020, as Sports Training, Sports Activity Management, and Exercise Science emerged as substantial secondary categories.

Thematic and Text Mining Analysis (RQ5): Thematic patterns within the literature were examined through the natural language processing (NLP) of English-language article titles (n = 689). Because author-provided keyword fields can sometimes be inconsistent or absent, title text mining provides a highly accurate, objective map of the field's intellectual structure. Two complementary techniques were applied: bigram frequency analysis and word co-occurrence network analysis.

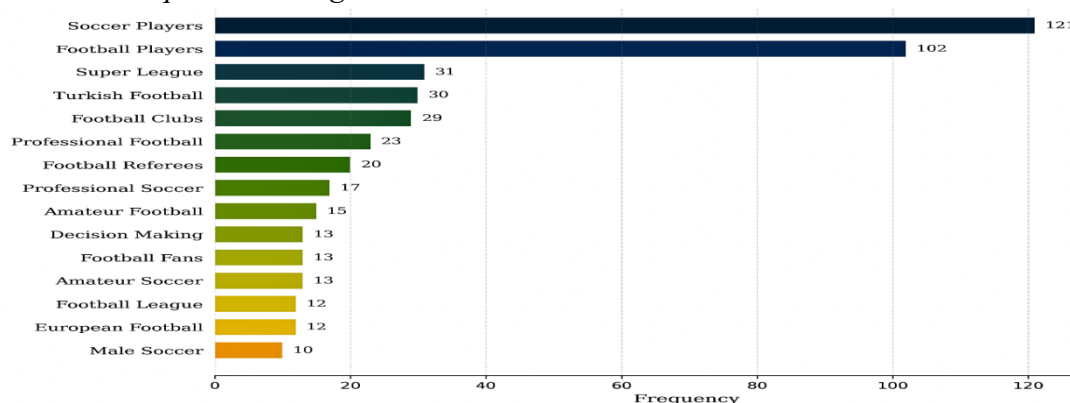
Bigram Analysis: Figure 4 presents the 15 most frequent substantive two-word phrases (bigrams) extracted from the article titles. Three patterns emerge from the bigram distribution. First, the field is overwhelmingly player-centric: “soccer players” and “football players” together account for 223 occurrences, dwarfing all other bigrams. Researchers in this corpus study athletes far more than they study the sport as a system or institution.

On the one hand, the prominence of terms such as “young,” “training,” “performance,” “agility,” and “speed” may support evidence-based planning for physical education and sport training contexts. On the other hand, the scarcity of terms explicitly associated with curriculum, teaching, learning, inclusion, assessment, or school contexts suggests that educational questions are still marginal in the title-level framing of football research. This finding reinforces the need for studies that connect football not only to athlete performance, but also to pedagogical mediation and educational management.

Organizational focus is heavily localized: “super league,” “Turkish football,” “football clubs,” and “football league” point to a research literature anchored in the domestic professional league context rather than in global or comparative frameworks. Third, a clear professional-amateur divide structures the literature, with professional football (23 + 17 = 40 combined) receiving roughly twice the attention of amateur football (15 + 13 = 28). “Male soccer” ($n = 10$) entered the top 15 while “female soccer” ($n = 8$) fell just below, indicating that gender-specific research remains a minor strand with near-equal visibility between male and female populations.

Figure 4.

Top 15 Most Frequent Title Bigrams



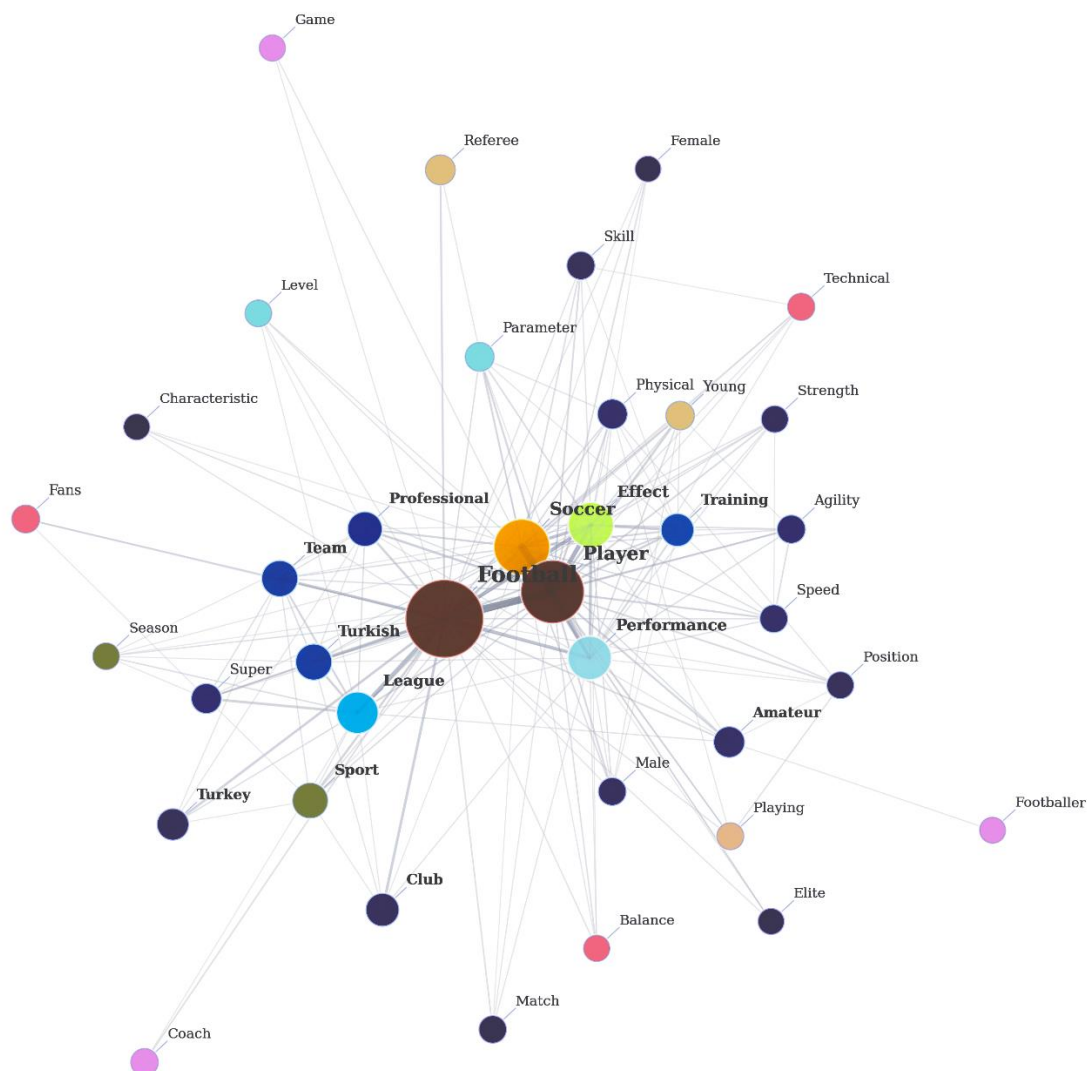
Note. The Authors.

Co-occurrence Network: Figure 5 presents a word co-occurrence network constructed from title terms. Node size reflects word frequency; edge width reflects co-occurrence strength; color intensity indicates weighted degree centrality. The network reveals two distinct research clusters. A dense central cluster links “performance,” “training,” “agility,” “speed,” “strength,” “physical,” and “young” around the core “football/soccer/player” node. Such a cluster maps the dominant research paradigm in the corpus: studies measuring the effect of training interventions on player physical performance. A sparser satellite cluster connects “league,” “club,”

“professional,” “amateur,” “team,” “fans,” and “Türkiye,” reflecting a secondary research stream concerned with sport administration and spectator behavior. Limited edge connectivity between these two clusters indicates that the physiological and social-scientific wings of Turkish football scholarship operate largely in parallel, with few studies bridging the divide.

Figure 5.

Title Word Co-occurrence Network (English-Language Titles, $n = 689$)



Note. The Authors. Nodes represent words appearing in 18 or more titles. Edges represent co-occurrence frequencies of 5 or higher. Turkish stopwords and methodological terms were excluded prior to analysis.

DISCUSSION

Eight research questions guided this bibliometric analysis of 1,411 football-related research articles published in DergiPark between 1967 and 2025. Following the distinction between performance analysis and science mapping proposed by Donthu et al. (2021) and

Öztürk et al. (2024), the findings describe a field that has grown rapidly in volume, expanded across disciplinary boundaries, and begun to internationalize its language profile, yet remains structurally constrained by a narrow citation economy, player-centric research framing, limited collaborative networks, and a still insufficient articulation with educational research. However, the central contribution of the study is not simply to quantify football publications, but to indicate how this scientific production may support physical education, sport pedagogy, teacher education, and institutional planning in sport-related educational settings.

Publication output followed an exponential curve, with five-year totals rising from 18 articles in 1996-2000 to 662 in 2021-2025. Nearly half the entire corpus was published in the most recent five-year window. Such acceleration did not occur uniformly: output remained modest through 2012, increased steadily between 2013 and 2015, and then surged after 2016. Two institutional factors likely converge to explain this pattern. First, DergiPark underwent a major infrastructure upgrade in the mid-2010s under Tübitak-Ulakbim, which lowered barriers to journal creation and article submission and attracted a growing number of sport science journals to the platform. Second, Turkish universities expanded their sport science faculties during the same period, producing a larger pool of researchers and graduate students generating publication output. Whether the growth reflects deeper intellectual engagement with football or simply more outlets for similar types of work is a question the present data alone cannot resolve. Block and Fisch (2020) caution that bibliometric research must distinguish between genuine field expansion and mere inflation of publication outlets, a distinction that requires content-level analysis beyond the scope of performance metrics alone. A longitudinal content analysis tracking topical shifts within the corpus over time would help clarify this. The slight dip in 2020 ($n = 96$, down from 102 in 2019) coincides with the onset of the COVID-19 pandemic, which disrupted data collection dependent on athlete access. Recovery was rapid, with output reaching 114 by 2021, suggesting that the pandemic produced a temporary interruption rather than a structural break in the field's growth.

Football research in DergiPark spreads across 343 journals, a figure that resists easy characterization. On one hand, the top five journals account for only 23.81% of all publications, indicating that no single outlet dominates the field. On the other hand, the long tail of 323 journals publishing fewer than 12 articles each accounts for nearly half the corpus (47.91%). Such a distribution is consistent with Bradford's Law of Scattering, which predicts that a small core of journals produces a disproportionate share of publications in any given field while a large periphery contributes sporadically (Donthu et al., 2021; Passas, 2024). The dispersal

carries a practical consequence for researchers: monitoring Turkish football scholarship requires searching broadly rather than tracking a small set of core journals. It also raises questions about editorial quality control, since journals publishing only a handful of football articles may lack specialist reviewers in sport science. *SPORMETRE*'s leading position ($n = 102$, 7.23%) reflects its longstanding role as a generalist sport science outlet, while the presence of journals from communication, law, economics, and health administration disciplines in the broader list confirms that football functions as a genuinely multidisciplinary research subject in the Turkish context.

The most significant finding concerns the gap between the field's breadth in volume and its concentration in citation impact. Although the corpus spans 110 disciplinary categories and 343 journals, the ten most-cited articles cluster into just three thematic streams: sports medicine and physical performance, fan behavior and spectator violence, and sport economics. Absent from the top-cited list are studies on coaching methodology, sport psychology, women's football, referee performance, or youth development, all of which appear in the corpus by volume but have not generated comparable citation impact. Such concentration raises a structural question. Are physiological studies and spectator violence surveys genuinely producing more impactful knowledge, or do they benefit from built-in citation advantages? Zupic and Čater (2015) note that citation analysis inherently favors older and more methodologically conventional studies, since newer or methodologically unconventional work has had less time and fewer disciplinary peers to accumulate citations. Clinical and survey-based designs also tend to cite each other more frequently, creating self-reinforcing citation clusters, while qualitative, historical, or policy-oriented studies cite across disciplinary boundaries, dispersing their citation footprint in ways that bibliometric tools may undercount. Whether the citation hierarchy reflects quality or methodology is a question future research should address, and whether emerging sub-fields like women's football and coaching science are building citation networks that have not yet reached threshold visibility.

Educational Implications for Physical Education and Sport Management

The results suggest that football research can be used as a strategic resource for educational decision-making. In physical education, the mapped literature may help teachers and curriculum designers identify evidence on training progression, youth development, motor learning, motivation, and participation. In higher education, the corpus can support research training in sport sciences by showing students how a field is structured, which methods

dominate, and where new pedagogical questions can be formulated. In educational management, the dispersion of publications across many journals and categories indicates the need for stronger institutional mechanisms to connect sport science, pedagogy, inclusion, and policy.

Future Research Directions

Öztürk et al. (2024) emphasize that well-conducted bibliometric research should not only map a field but also identify knowledge gaps and outline future research agendas. Four directions follow from the present findings. Expanding methodological range is the most pressing priority. Title analysis strongly suggests experimental and survey-based approaches dominate the corpus. Qualitative studies exploring fan culture, media representation, institutional governance, and athlete experience would complement the existing quantitative base and open research questions that measurement-oriented designs cannot reach. Building cross-disciplinary collaborative networks is the second priority. Structural separation between the performance-training cluster and the management-sociology cluster in the co-occurrence network reflects a practical division in how football is studied. Research teams that include exercise scientists, education researchers, teachers, coaches, and social scientists could produce integrative work linking physiological load monitoring data with pedagogical decision-making, curriculum planning, or club and school sport management. Increasing international visibility could be another direction. English-language share has grown substantially, but nearly three quarters of the corpus remains accessible only to Turkish-reading audiences. Bilingual publication strategies, where authors submit to both Turkish and international journals or publish in English within DergiPark-indexed journals, offer a path toward connecting Turkish findings to global football scholarship without abandoning the domestic knowledge circulation that Turkish-language publication serves.

Limitations

Analyses were drawn exclusively from DergiPark. Turkish football research published in Web of Science, Scopus, or TR Dizin-indexed journals not hosted on DergiPark falls outside this corpus. The search strategy captured articles containing “futbol,” “football,” or “soccer” in their titles or keywords; articles addressing football without these terms were missed. DergiPark does not export author-assigned keywords, which prevented conventional co-word analysis and required title-based NLP as a substitute. Title words reflect how authors frame their studies but

are not equivalent to deliberate keyword selections. Subject category data was available for only 63.1% of articles, with gaps concentrated in older publication years. Citation data was obtained through Google Scholar rather than through systematic citation analysis, and the ten most-cited articles represent an indicative snapshot rather than an exhaustive ranking. Future bibliometric work on this corpus would benefit from multi-database coverage, article-level keyword scraping, systematic citation analysis, and methodological classification conducted by multiple coders. A further limitation is that the dataset was not originally built exclusively from educational descriptors; therefore, the educational implications presented here are interpretive and should be strengthened in future studies through searches combining football with terms such as physical education, sport pedagogy, curriculum, teacher education, school sport, and educational policy.

FINAL CONSIDERATIONS

The bibliometric analysis of 1,411 football-related research articles in DergiPark shows a field that has grown at an accelerating rate, with nearly half the corpus published between 2021 and 2025 and English-language share rising from under 5% before 2010 to 35.8% in the most recent period. Such growth has been accompanied by disciplinary diversification across 110 subject categories and 343 journals, yet the field's research focus remains concentrated: title-level analysis identifies a dominant player-centric, physiological-experimental paradigm structurally separated from a smaller management-sociology stream, and citation impact clusters in sports medicine and spectator violence studies.

For education, these findings indicate both potential and insufficiency. Football research already contains strands relevant to Sports Training, Exercise Science, Sport Activity Management, and Studies on Education, but it still needs stronger engagement with physical education, sport pedagogy, teacher education, curriculum, inclusion, and institutional governance. The main conclusion is that mapping football research is educationally relevant when it is used to identify how scientific production can inform pedagogical practice, professional formation, and the management of sport-related educational programs. Strengthening this educational orientation would make future football studies more useful not only for sport science specialists, but also for teachers, curriculum planners, university programs, and public or institutional managers concerned with sport as a field of learning.

REFERENCES

- Algin, A. (2024). Bibliometric analysis of research on “gamification in nursing” via visual mapping technique. *Cadernos de Educação, Tecnologia e Sociedade*, 17(se5), 219–229. <https://doi.org/10.14571/brajets.v17.nse5.219-229>
- Alkan, İ., Algin, A., Toros, T., & Serin, E. (2025). Social inclusion and sports: Strategies to promote participation of people with disabilities (last 10 years). *Cuestiones de Fisioterapia*, 54(3), 4849–4858. <https://doi.org/10.48047/jbfsy872>
- Armour, K. (2011). *Sport pedagogy: An introduction for teaching and coaching*. Pearson Education.
- Ayar, H., & Koç, M. C. (2018). An analysis of graduate theses in the field of recreation in Türkiye between 1980–2018. *Sportif Bakış: Spor ve Eğitim Bilimleri Dergisi*, 5(1), 96–107.
- Aydın, B., & Aksöz, O. E. (2019). Bibliometric profile of graduate theses published in the field of destination. *Journal of Tourism and Gastronomy Studies*, 7(1), 615–636. <https://doi.org/10.21325/jotags.2019.381>
- Bailey, R., Armour, K., Kirk, D., Jess, M., Pickup, I., Sandford, R., & BERA Physical Education and Sport Pedagogy Special Interest Group. (2009). The educational benefits claimed for physical education and school sport: An academic review. *Research Papers in Education*, 24(1), 1–27. <https://doi.org/10.1080/02671520701809817>
- Bao, J. (2002). Tourism geography as the subject of doctoral dissertations in China, 1989–2000. *Tourism Geographies*, 4(2), 148–152. <https://doi.org/10.1080/14616680210124918>
- Block, J. H., & Fisch, C. (2020). Eight tips and questions for your bibliographic study in business and management research. *Management Review Quarterly*, 70, 307–312. <https://doi.org/10.1007/s11301-020-00188-4>
- Cevizkaya, G., İlsay, S., & Avcıkurt, C. (2014). Bibliometric profile of studies on disabilities in tourism literature (2000–2013). *Gazi Üniversitesi Turizm Fakültesi Dergisi*, 2, 101–108.
- Çiçek, D., & Kozak, N. (2012). Bibliometric profile of peer-reviewed articles published in *Anatolia: Turizm Araştırmaları Dergisi*. *Türk Kütüphaneciliği*, 26(4), 734–756.
- Diodato, V. (1994). *Dictionary of bibliometrics*. Haworth Press.
- Donthu, N., Kumar, S., Mukherjee, D., Pandey, N., & Lim, W. M. (2021). How to conduct a bibliometric analysis: An overview and guidelines. *Journal of Business Research*, 133, 285–296. <https://doi.org/10.1016/j.jbusres.2021.04.070>
- Jacobs, D. (2001). A bibliometric study of the publication patterns of scientists in South Africa 1992–96, with particular reference to status and funding. *Information Research*, 6(3).
- Kaya, N., & Algin, A. (2022). Kamu hastanelerinde teknik etkinlik: Bir meta-regresyon analizi. *Eskişehir Osmangazi Üniversitesi İktisadi ve İdari Bilimler Dergisi*, 17(3), 810–821. <https://doi.org/10.17153/oguiibf.1094736>

- Kirk, D. (2010). *Physical education futures*. Routledge.
- Köseoğlu, M. A., Rahimi, R., Okumuş, F., & Liu, J. (2016). Bibliometric studies in tourism. *Annals of Tourism Research*, 61, 180–198. <https://doi.org/10.1016/j.annals.2016.10.006>
- Meyer-Arendt, K. J. (2000). Commentary: Tourism geography as the subject of North American doctoral dissertations and master's theses, 1951–98. *Tourism Geographies*, 2, 140–156. <https://doi.org/10.1080/14616680050027860>
- Moral-Muñoz, J. A., Herrera-Viedma, E., Santisteban-Espejo, A., & Cobo, M. J. (2020). Software tools for conducting bibliometric analysis in science: An up-to-date review. *Profesional de la Información*, 29(1), 1–20. <https://doi.org/10.3145/epi.2020.ene.03>
- Mukherjee, D., Lim, W. M., Kumar, S., & Donthu, N. (2022). Guidelines for advancing theory and practice through bibliometric research. *Journal of Business Research*, 148, 101–115. <https://doi.org/10.1016/j.jbusres.2022.04.042>
- Okubo, Y. (1997). *Bibliometric indicators and analysis of research systems: Methods and examples*. OECD Science, Technology and Industry Working Papers, (1997/01).
- Özel, Ç. H., & Kozak, N. (2012). Bibliometric profile of tourism marketing (2000–2010) and a citation analysis study. *Türk Kütüphaneciliği*, 26(4), 715–733.
- Öztürk, O., Kocaman, R., & Kanbach, D. K. (2024). How to design bibliometric research: An overview and a framework proposal. *Review of Managerial Science*, 18, 3333–3361. <https://doi.org/10.1007/s11846-024-00738-0>
- Passas, I. (2024). Bibliometric analysis: The main steps. *Encyclopedia*, 4(2), 1014–1025. <https://doi.org/10.3390/encyclopedia4020065>
- Polat, S., & Bingöl, H. (2022). Sport management as a profession and its ethical values. In S. Dal & S. Biçer Baikoğlu (Eds.), *Academic studies in sport sciences 20* (p. 22). Duvar Yayınları.
- Pritchard, A. (1969). Statistical bibliography or bibliometrics? *Journal of Documentation*, 25, 348–349.
- Raisig, L. M. (1962). Statistical bibliography in the health sciences. *Bulletin of the Medical Library Association*, 50(3), 450–461.
- Roy, B. S., & Basak, M. (2013). Journal of Documentation: A bibliometric study. *Library Philosophy and Practice*, 1, 1–10.
- Şahin, S., & Acun, A. (2015). Bibliometric profile of the field of tourist guidance. *Balıkesir Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 18(34), 213–234. <https://doi.org/10.31795/baunsobed.645358>
- Şakar, G. D., & Cerit, A. G. (2013). Turkish marketing literature in international indexes: Bibliometric analyses and a qualitative study. *Atatürk Üniversitesi İktisadi ve İdari Bilimler Dergisi*, 27(4), 37–62.

- Sanchez, A. D., Del Rio, M. D. L. C., & Garcia, J. A. (2017). Bibliometric analysis of publications on wine tourism in the databases Scopus and WoS. *European Research on Management and Business Economics*, 23(1), 8–15.
- Schrader, A. M. (1981). Teaching bibliometrics. *Library Trends*, 30(6), 151–172.
- Sengupta, I. N. (1992). Bibliometrics, informetrics, scientometrics and librametrics: An overview. *Libri*, 42(2), 75–98. <https://doi.org/10.1515/libr.1992.42.2.75>
- Serin, E., Algin, A., & Toros, T. (2025). The relationship between exercise and sport and epigenetics. *Cuestiones de Fisioterapia*, 54(3), 4859–4871. <https://doi.org/10.48047/5eznhz12>
- Sökmen, C., & Özkanlı, O. (2018). Development of gastronomy tourism literature: An analysis of articles published in *Journal of Tourism and Gastronomy Studies*. *Journal of Tourism and Gastronomy Studies*, 6(2), 99–127.
- Sünnetçioğlu, A., Yalçınkaya, P., Olcay, M., & Mercan, Ş. O. (2017). Bibliometric profile of gastronomy-related theses in tourism. *Journal of Tourism and Gastronomy Studies*, 5(2), 345–354.
- Tayfun, A., Küçükergin, F. N., Aysen, E., Eren, A., & Özekici, Y. K. (2016). A bibliometric analysis of graduate theses in tourism. *Gazi Üniversitesi Turizm Fakültesi Dergisi*, 1(1), 50–69.
- Tazegül, Ü. (2014). Investigating the effect of sport on personality. *Journal of Academic Social Science Studies*, 25, 537–544.
- Toros, A. A. (2025). The development and importance of gamification in nurse managers. *International Journal of Sports Technology and Science*, 3(1), 195–206.
- Toros, A. A., & Şekeroğlu, M. Ö. (2026). Teknolojinin ve küreselleşmenin kesişiminde klinik otorite ve bakım paradigması: Hemşirelik ve sağlık tarihine sistemik bir bakış. *Journal of History School*, 19(80), 186–197. <https://doi.org/10.29228/joh.88740>
- UNESCO. (2015). *Quality physical education (QPE): Guidelines for policy-makers*. UNESCO Publishing.
- Yetiş, A. Ş., & Çokal, Z. (2018). Bibliometric profile of articles published in the field of winter tourism. *Journal of Recreation and Tourism Research*, 5(1), 38–52.
- Zencir, E., & Kozak, N. (2012, April 12–15). *Bibliometric profile of tourism articles published in social sciences institute journals (2000–2010)* [Conference paper]. 6th Graduate Tourism Students Research Congress, Antalya, Türkiye.
- Zupic, I., & Čater, T. (2015). Bibliometric methods in management and organization. *Organizational Research Methods*, 18(3), 429–472. <https://doi.org/10.1177/1094428114562629>

CRediT Author Statement

- ☐ **Acknowledgments:** Not applicable.
 - ☐ **Funding:** The development of this research did not receive financial support from any funding agency.
 - ☐ **Conflicts of interest:** None.
 - ☐ **Ethical approval:** Not applicable.
 - ☐ **Data and material availability:** The data sources are publicly available on the respective websites.
 - ☐ **Authors' contributions:** Turhan Toros: Conceptualization, supervision, theoretical framework, manuscript preparation. Emre Serin: Literature review, methodological design, data analysis. Arif Mert Ozkan: Pedagogical expertise, critical review, manuscript revision. Mehmet Bayansalduz: Data collection, analysis, visualization of educational materials. Ali Emre Eren: Data collection, analysis, visualization of educational materials. Mihriay Musa: Pedagogical expertise, theoretical framework. Erkan Gülgösteren: Visualization of educational materials.
-

Processing and editing: Editora Ibero-Americana de Educação
Proofreading, formatting, normalization and translation

