

SCIENTIFIC RESEARCH ON BINARY LOGISTIC  
REGRESSION: A BIBLIOMETRIC ANALYSIS  
USING SCOPUS DATA (1974-2024)

INVESTIGACIÓN CIENTÍFICA SOBRE  
REGRESIÓN LOGÍSTICA BINARIA: UN ANÁLISIS  
BIBLIOMÉTRICO USANDO DATOS DE SCOPUS  
(1974-2024)

HUMBERTO LLINÁS<sup>1</sup> DANIELA NUÑEZ<sup>2</sup> JORGE TILANO<sup>3</sup> JORGE VILLALBA<sup>4</sup>

*Received: 13/Nov/2024; Accepted: 03/Dec/2025*

---

*Revista de Matemática: Teoría y Aplicaciones* is licensed under a Creative Commons  
Reconocimiento-NoComercial-Compartirigual 4.0 International License.  
Creado a partir de la obra en <http://www.revistas.ucr.ac.cr/index.php/matematica>



<sup>1</sup> Universidad del Norte, Departamento de Matemáticas y Estadística, Barranquilla, Colombia. Email: [hllinas@uninorte.edu.co](mailto:hllinas@uninorte.edu.co)

<sup>2</sup> Same as Humberto Llinás, E-mail: [nunezdm@uninorte.edu.co](mailto:nunezdm@uninorte.edu.co)

<sup>3</sup> Same as Humberto Llinás, E-mail: [jtilano@uninorte.edu.co](mailto:jtilano@uninorte.edu.co)

<sup>4</sup> Universidad Tecnológica de Bolívar, Facultad de Ciencias Básicas, Cartagena de Indias, Colombia. Email: [jvillalba@utb.edu.co](mailto:jvillalba@utb.edu.co)

### Abstract

This study conducts a comprehensive bibliometric analysis of binary logistic regression research spanning from 1974 to 2024. It examines 15,409 documents authored by 77,557 researchers across 137 countries and published in 4,669 different sources. The results reveal a strong and sustained annual growth rate of publications, with a marked surge after 2010, and a global shift toward interdisciplinary applications. China leads in volume of publications (10,553 articles), while the United Kingdom and the Netherlands demonstrate the highest citation impact (approx. 23.5 citations per article). Ethiopia emerges as a notable contributor from the Global South, with more than 8,000 publications. At the journal level, *PLOS One* and *BMC Public Health* stand out as the most prolific outlets, whereas the works of Hosmer and Lemeshow remain the most influential references in the field. Thematic mapping highlights clusters in public health, epidemiology, and artificial intelligence, underscoring the central role of binary logistic regression in both methodological and applied domains. These findings provide a panoramic view of the field, highlighting key contributors, trends, and opportunities for future research. The analysis was performed using R-4.4.1 and VOSviewer 1.6.20 software.

**Keywords:** binary logistic regression; bibliometric analysis; research trends; global research impact; R software; VOSviewer.

### Resumen

Este estudio lleva a cabo un análisis bibliométrico exhaustivo de la investigación sobre regresión logística binaria entre 1974 y 2024. Examina 15,409 documentos escritos por 77,557 investigadores de 137 países y publicados en 4,669 fuentes diferentes. Los resultados revelan una tasa de crecimiento anual fuerte y sostenida de las publicaciones, con un notable aumento después de 2010, y un cambio global hacia aplicaciones interdisciplinarias. China lidera el volumen de publicaciones (10,553 artículos), mientras que el Reino Unido y los Países Bajos muestran el mayor impacto en cuanto a citas (aproximadamente 23.5 citas por artículo). Etiopía emerge como un contribuyente notable del Sur Global, con más de 8,000 publicaciones. A nivel de revistas, *PLOS One* y *BMC Public Health* destacan como las publicaciones más prolíficas, mientras que los trabajos de Hosmer y Lemeshow siguen siendo las referencias más influyentes en el campo. El mapeo temático destaca los grupos en salud pública, epidemiología e inteligencia artificial, lo que subraya el papel central de la regresión logística binaria tanto en el ámbito metodológico como en el aplicado. Estos hallazgos ofrecen una visión panorámica del campo, destacando los principales contribuyentes, las tendencias y las oportunidades para futuras investigaciones. El análisis se realizó utilizando el software R-4.4.1 y VOSviewer 1.6.20.

**Palabras clave:** regresión logística binaria; análisis bibliométrico; tendencias de investigación; impacto de la investigación global; software R; VOSviewer.

**Mathematics Subject Classification:** Primary: 62-02; secondary: 62Q05.

## 1. INTRODUCTION

In recent years, binary logistic regression has become increasingly prominent worldwide, recognized for its effectiveness in classification models that aid decision-making across diverse disciplines. As a versatile statistical technique, it is widely utilized in various fields, including artificial intelligence and healthcare, demonstrating its broad applicability and impact. For instance, in clinical studies, researchers apply binary logistic regression to analyze risk factors associated with different pathologies [34, 35, 45]. In education research, this method is widely used to identify factors influencing academic performance, school dropout rates, and teacher training [10, 22, 59]. In credit risk analysis, logistic regression plays a crucial role in classifying clients based on their creditworthiness, aiding financial institutions in risk assessment [9, 27, 43, 56]. Recent studies further highlight its application in financial institutions for credit risk evaluation, as well as its use in analyzing the impact of education on poverty reduction [18, 47]. Additionally, in anthropology and forensic sciences, logistic regression models are employed for sex estimation and other biological classifications [37, 52, 53, 60]. In medical research, logistic regression aids in predictive modeling and decision-making, with applications implemented using various programming languages [40, 44, 46, 55, 58]. Its performance in consumer credit risk assessment rivals that of gradient-boosted machines and neural networks, offering better interpretability [48]. Additionally, logistic regression is effective in predicting loan approval likelihood based on applicant data [49].

Thus, binary logistic regression has proven to be a fundamental statistical tool across multiple domains, enabling evidence-based decision-making under uncertainty. Its ability to classify and predict outcomes based on multiple predictor variables underscores its versatility and significance in empirical research.

Despite its extensive use, no systematic bibliometric reviews have been conducted to map the intellectual trajectory, methodological innovations, and applications of binary logistic regression across various disciplines. Over the past decades, significant advances have improved its computational efficiency [6, 21, 26, 41], variable selection methods [4, 28, 30], and adaptability to hierarchical and longitudinal data structures [1, 36]. Moreover, its integration with artificial intelligence has led to new hybrid models for medical diagnostics [31], big data analysis [23, 57], and deep learning applications [7]. Logistic regression has also been instrumental in survey-based research, incorporating weighting adjustments for complex data [50, 51], and real-time fraud detection systems [2, 33, 38]. Given its continuous development, a bibliometric analysis is essential to systematically assess its evolution, identify research gaps, and explore emerging trends in its application across diverse fields of knowledge.

Bibliometric analysis, a critical document review technique, provides valuable insights into the literature on specific topics [20]. Over the past 50 years, the application of bibliometric methods has significantly increased, particularly in scientific

and applied fields [12]. Studies focusing on areas such as fuzzy logic research [32] and logistics [54] have revealed shifts in global research contributions and emerging trends. For instance, logistic regression is widely used in cancer research for tumor classification, prognosis, and risk assessment [24].

This paper aims to analyze bibliometric trends in binary logistic regression models, focusing on their evolution, application, and impact across various disciplines over recent years. It highlights how results are programmed and integrated to clarify the topic, gathering significant historical findings to identify patterns over time and achieve a comprehensive understanding [5, 12, 19].

This study addresses the following research questions: How have binary logistic regression models evolved in their application across scientific disciplines over the past 50 years? What emerging trends and patterns are evident in using logistic regression models within artificial intelligence and advanced computational techniques?

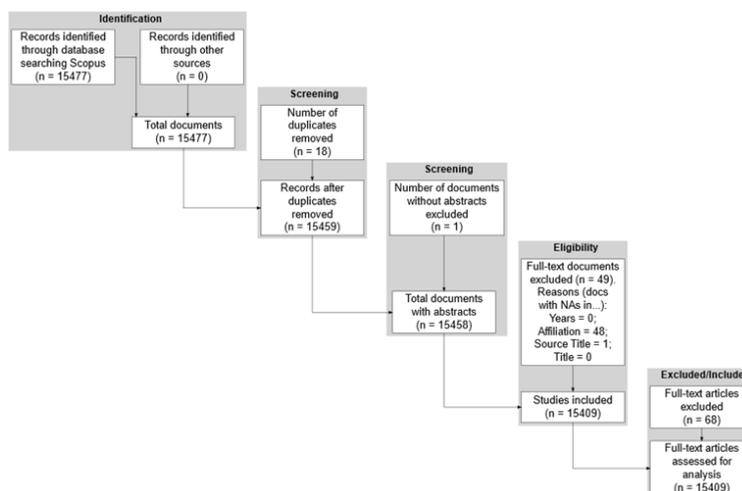
Key findings in this study include a description of sources and authors through indicators such as scientific production and its impact across different disciplines. The growth of scientific production is assessed using specific indicators [11, 13, 29] focusing on the number of publications by authors, institutions, countries, or subject areas [3]. Bibliometric analysis can map scientific output trends, as shown in studies on innovation indicators, which reveal an increasing number of publications over time [16]. Additionally, relative indicators accounting for variables such as population, GDP, or granted projects are crucial in evaluating scientific production [3].

The scope of this study includes a comprehensive bibliometric analysis of binary logistic regression models, drawing data from Scopus between 1974 and 2024, with a focus on their intersection with artificial intelligence and computational methods like R and Python. However, this study is limited by its reliance on Scopus-indexed publications, potentially excluding relevant research from other databases like Web of Science or PubMed. The focus on English-language documents may also overlook significant contributions in other languages. Lastly, while the analysis highlights trends and patterns, it does not delve into the technical depth of the methodologies used in each paper, concentrating instead on broader bibliometric indicators.

## 2. MATERIALS AND METHODS

### 2.1. Data Collection.

We followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines [39] to ensure a rigorous and transparent selection process for the articles included in our study, as illustrated in Figure 1.



**Figure 1:** PRISMA Flow diagram: detailed workflow of document selection and screening stages.

This process involved several key stages: identification, screening, eligibility, and inclusion. During the identification phase, we retrieved documents from Scopus using the following search string, applied to titles, abstracts, and keywords:

```
(logistic model* OR "logistic reg*" AND "dichotomic*" OR "dichotomous*"
OR "binary*") OR (logit model* OR "logit reg*" AND "dichotomic*"
OR "dichotomous*" OR "binary*") AND (AI* OR "artificial intelligence*"
OR "software R" OR "Python")
```

We selected Scopus as the primary database due to its broad coverage and diversity of peer-reviewed literature [14, 25]. The initial search identified 15,477 articles. In the screening stage, we removed duplicates ( $n = 18$ ) and records without abstracts ( $n = 1$ ). In the eligibility stage, 49 additional records were excluded due to incomplete metadata (missing affiliation = 48, missing source title = 1). This process resulted in 15,409 articles being included in the final dataset (Figure 1).

## 2.2. Bibliometric Analysis.

Bibliometric analysis has become a fundamental method for evaluating scientific literature, particularly gaining traction since the mid-1990s [12]. This computer-assisted technique enables researchers to quantitatively assess various dimensions of academic publications, including authorship, institutional affiliations, and citation patterns [25]. While bibliometric analysis offers valuable insights and visualizations, it should complement critical analysis and comprehensive literature reviews [14]. When used alongside other evaluation methods, bibliometric analysis provides a powerful tool for understanding research trends and impact.

In our study, we conducted a systematic analysis at three levels: country, source, and author. We used quantitative methods, such as citation analysis, to identify the most significant topics at each level. We then examined social structures (collaborations among authors and countries), conceptual structures (key themes and trends), and intellectual structures (the impact of specific works on scientific communities) to perform an in-depth analysis of knowledge structures within the field [8].

Bibliometric analysis typically involves two primary methods: performance analysis and science mapping [11, 15, 17]. Performance analysis assesses the impact and productivity of scientific contributors, such as researchers, institutions, and countries, using bibliographic data [11, 17]. Science mapping visualizes the structure and dynamics of scientific research, identifying subfields through techniques like co-word analysis and document co-citation analysis [11, 17]. For detailed information on these techniques, refer to Donthu et al. [11]. The selection of these techniques was based on their suitability for capturing both productivity and intellectual structures. Performance analysis enabled us to measure scientific impact through publication and citation metrics, while science mapping techniques such as co-word and co-citation analysis allowed us to identify conceptual clusters and collaboration patterns. This dual approach ensures a comprehensive evaluation of the field. All analyses were performed using the R-4.4.1 package `bibliometrix` [8] for quantitative bibliometric and scientometric analysis. Scientific mapping was conducted using VOSviewer 1.6.20, ensuring reproducibility and comparability with prior bibliometric studies. This combined approach provides a robust framework to capture both productivity indicators and intellectual structures, directly supporting the research questions guiding this study.

### 3. RESULTS

#### 3.1. Summary.

Table 1 summarizes the primary data used in the bibliometric analysis of binary logistic regression over a 50-year span (1974-2024). This comprehensive dataset includes 15,409 documents from 4,669 sources, encompassing journals, books, and other publication types.

The annual growth rate of document production in binary logistic regression research is 14.81%, reflecting a significant increase in publications over the years. The average number of citations per document is 12.38, indicating the substantial impact of these works. With an average document age of 4.71 years, the research in this field is relatively recent and dynamic. Author collaboration is extensive, with 65,272 contributors, including 415 single-authored documents, and 437 documents written by individual authors.

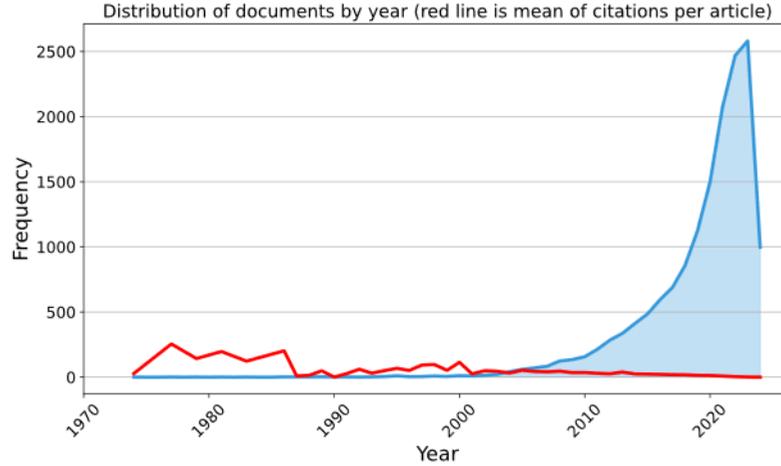
**Table 1:** Primary data for the bibliometric analysis of binary logistic regression publications.

Description	Result
<b>Main information about data</b>	
Timespan	1974:2024
Sources (Journals, Books, etc)	4,669
Documents	15,409
Annual growth rate %	14.81
Document average age	4.71
Average citations per doc	12.38
<b>Document contents:</b>	
Keywords Plus	35,988
Author's keywords	29,153
<b>Authors:</b>	
Authors	65,272
Authors of single-authored docs	415
<b>Authors collaboration</b>	
Single-authored docs	437
Co-authors per doc	6.31
International co-authorships %	20.53
<b>Document types:</b>	
Article	14,629
Book and Book chapter	55
Conference Paper	537
Note and Review	174

Most works are collaborative, with an average of 6.31 co-authors per document, and 20.53% of the publications involve international co-authorship, highlighting the global nature of this research. In terms of document types, articles dominate with 14,629 publications, followed by 50 book chapters, 537 conference papers, 13 notes, and 161 reviews, underscoring the predominance of articles as the main medium for disseminating research findings.

Figure 2 illustrates the distribution of documents published annually from the 1970s to 2024 and the average citations per article.

While publication numbers were low and stable during the 1970s and 1980s, a sharp increase began in the 2010s, peaking around 2020, reflecting heightened interest in the field. However, despite this surge, the average citations per article have remained consistently low, suggesting that the rise in publication volume has not led to a corresponding increase in impact. This may indicate that the perceived quality or relevance of research has not kept pace with the growth in quantity, or that citations are becoming more dispersed due to the increasing number of publications. To better understand this trend, it would be useful to analyze the types of journals or conferences where these papers are published, the relevance of the research to current trends, and how citations are distributed across different document types. Comparing this pattern with other fields could also help determine if this is unique to binary logistic regression or part of a broader trend in academia.



**Figure 2:** Trends in article publications and citation averages.

### 3.2. Sources.

Table 2 provides performance metrics for key scientific sources in binary logistic regression, covering the period from 1974 to 2024. It includes the number of publications, total citations, h-index, and the year of the first publication.

Table 2 highlights the performance of the top 30 scientific journals in the field of binary logistic regression, ranked by the total number of publications. “Plos One” (Rank 1) stands out as the most productive source, with 529 publications. It also leads in citations, with a remarkable total of 7,385, and has an impressive h-index of 40, underscoring its substantial influence in the field since its inception in 2012. This combination of high publication volume and citations makes “Plos One” a key player in the dissemination of research in this area.

“BMC Public Health” (Rank 2) follows closely with 253 publications and 4,402 citations, maintaining a significant presence in the field since 2002. Its h-index of 35 further emphasizes its continued relevance and impact. Similarly, the “International Journal of Environmental Research and Public Health” (Rank 3) has established itself as a crucial source since 2012, with 190 publications and 2,029 citations, supported by a respectable h-index of 22. Other notable sources include “Scientific Reports” (Rank 5) and “BMJ Open” (Rank 6), both of which have made substantial contributions to the field. “Scientific Reports”, with 145 publications and 1,131 citations, and “BMJ Open”, with 130 publications and 1,161 citations, both have h-indices that reflect their growing influence since their relatively recent introduction in 2016 and 2013, respectively. However, it is interesting to note that some sources, despite having a high number of publications, show relatively lower h-index or citation counts. For example, “Frontiers in Public Health” (Rank 4)

has 168 publications but a lower citation count of 931 and an h-index of 14, which may suggest that while it is a prolific source, its individual articles may not be as frequently cited or impactful as those in other journals. Similarly, “Frontiers in Neurology” (Rank 7) and “BMC Women’s Health” (Rank 10) also show a disparity between their publication count and citation impact, indicating potential areas for growth in terms of influence. One opportunity for further enhancing the impact of journals like “Frontiers in Public Health” (Rank 4) and “Frontiers in Neurology” (Rank 7) could be to focus on increasing the visibility and quality of individual articles. Strategies might include targeting more interdisciplinary research that could appeal to a broader audience or ensuring that published articles are shared and cited more widely through academic networks and collaborations. Moreover, journals with a high publication volume but relatively low h-indices, such as “BMC Women’s Health” (Rank 10) and “Frontiers in Psychology” (Rank 23), might benefit from initiatives aimed at improving the overall citation impact of their articles. This could involve encouraging submissions of more groundbreaking or highly innovative research, which is more likely to garner citations, or enhancing peer-review processes to ensure that only the most impactful studies are published.

### 3.3. Authors.

Table 3 presents the ten most prolific authors in the field of binary logistic regression research, spanning from 1974 to 2024. These authors are recognized for their contributions based on the number of publications, total citations, h-index, and the year they began publishing in this domain.

“Wang Y” (Rank 1) tops the list with 229 publications and 2,395 citations since beginning their contributions in 2009, boasting an impressive h-index of 22. This reflects both a high volume of research output and significant impact within the field. “Zhang Y” (Rank 2) follows closely with 202 publications, 2,186 citations, and an h-index of 22, highlighting their influential role in binary logistic regression research. “Li Y” (Rank 3), active since 2011, has made substantial contributions with 192 publications and 1,614 citations, resulting in a strong h-index of 21. “Liu Y” (Rank 4) has 174 publications, 1,247 citations, and an h-index of 20, demonstrating consistent academic influence since 2009.

Notable contributors also include “Li J” (Rank 5) and “Li X” (Rank 6), both with over 150 publications. “Li J” has 1,518 citations, while “Li X” has 1,513 citations, each with an h-index of 19, solidifying their positions as leaders in the field. “Liu X” (Rank 7), despite having fewer publications (102), impressively has 1,354 citations, indicating a high impact per publication.

To further enhance visibility and impact, authors like “Li Y” and “Liu X”, who have strong citation counts relative to their publication numbers, could focus on producing high-quality, groundbreaking research. Strategic dissemination and international collaboration could increase citations and broaden their influence globally. For those with a high volume of publications but slightly lower citation

**Table 2:** Top 30 scientific journals for binary logistic regression research based on 15,409 articles.

Rank	Source	Pub.	Cit.	h-index	Year
1	Plos One	529	7385	40	2012
2	BMC Public Health	253	4402	35	2002
3	International Journal of Environmental Research and Public Health	190	2029	22	2012
4	Frontiers in Public Health	168	931	14	2014
5	Scientific Reports	145	1131	16	2016
6	BMJ Open	130	1161	17	2013
7	Frontiers in Neurology	88	537	11	2015
8	BMC Psychiatry	87	824	15	2010
9	BMC Pregnancy and Childbirth	85	1621	25	2011
10	BMC Women's Health	74	534	10	2014
11	Medicine (United States)	71	1025	19	2015
12	Sustainability (Switzerland)	71	974	16	2016
13	BMC Health Services Research	70	855	15	2008
14	Journal of Clinical Medicine	63	387	12	2019
15	Nutrients	61	630	13	2016
16	BMC Geriatrics	59	877	17	2010
17	Frontiers in Psychiatry	56	881	12	2014
18	Journal of Affective Disorders	55	1007	18	2010
19	Frontiers in Endocrinology	54	436	11	2013
20	BMC Pediatrics	52	654	14	2010
21	BMC Infectious Diseases	48	603	14	2013
22	HIV/AIDS - Research and Palliative Care	46	364	11	2015
23	Frontiers in Psychology	44	584	12	2016
24	Biomed Research International	43	456	13	2013
25	BMC Oral Health	41	365	11	2007
26	BMC Research Notes	38	807	17	2010
27	Journal of Medical Internet Research	31	596	12	2016
28	Archives of Public Health	31	345	11	2014
29	Reproductive Health	30	475	12	2012
30	Journal of Clinical Nursing	30	344	11	2008

**Table 3:** Top ten most prolific authors in binary logistic regression based on 15,409 articles. Here, **Pub.** is the number of publications; **Cit.** is the total citations; **Year** is the publication start year.

Rank	Author	Pub.	Cit.	h-index	Year
1	Wang Y	229	2395	22	2009
2	Zhang Y	202	2186	22	2009
3	Li Y	192	1614	21	2011
4	Liu Y	174	1247	20	2009
5	Li J	164	1518	19	2008
6	Li X	158	1513	19	2012
7	Liu X	102	1354	19	2013
8	Wang X	146	1290	19	2012
9	Wang Z	94	903	19	2013
10	Zhang J	140	1681	19	2009

impact, such as “Li J” and “Li X”, engaging in interdisciplinary and cutting-edge research and targeting high-impact journals or conferences could boost their academic standing and h-index, enabling them to lead the field more decisively.

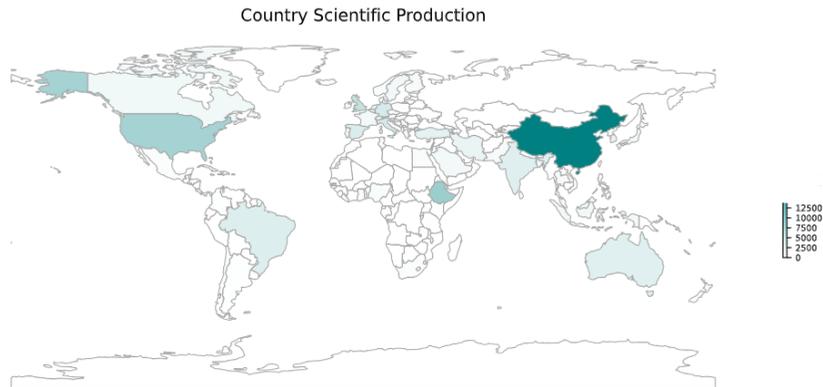
### 3.4. Countries.

Table 4 presents a bibliometric analysis of the 30 most productive countries in academic research from 1974 to 2024, highlighting total publications, total citations, and average citations per article.

**Table 4:** Top 30 countries: citation and publication metrics. Here, **Pub.** is the number of publications; **Cit.** is the total citations; **Year** is the publication start year.

Rank	Country	Pub.	Cit.	Average Article Citations
1	China	10553	23166	8.40
2	Ethiopia	8005	11713	7.60
3	USA	7265	19638	19.70
4	United Kingdom	4726	12323	23.60
5	Germany	3437	7802	16.70
6	Italy	3386	6231	19.10
7	Spain	3046	9765	21.90
8	Brazil	2728	3422	8.80
9	India	2592	4742	8.80
10	Australia	2496	7697	19.40
11	Turkey	2168	3043	7.60
12	Iran	1718	3103	10.10
13	Korea	1709	3929	14.00
14	Netherlands	1546	4442	23.50
15	Japan	1160	1596	9.40
16	Canada	1148	2419	14.20
17	Indonesia	1080	855	3.80
18	Sweden	1074	3423	15.60
19	Malaysia	1001	1476	8.50
20	Saudi Arabia	998	1213	7.70
21	France	972	1542	15.90
22	Bangladesh	889	1952	13.40
23	Finland	795	1369	12.10
24	Thailand	779	1209	6.50
25	Pakistan	728	807	8.20
26	Portugal	710	1314	12.60
27	Greece	667	1175	12.10
28	Belgium	621	1634	20.90
29	Nigeria	613	738	6.30
30	Austria	607	1332	18.50

China (Rank 1) leads with 10,553 publications and 23,166 citations, averaging 8.40 citations per article. Despite the high output, the lower average suggests a diluted impact per study. Ethiopia (Rank 2) follows with 8,005 publications and 11,713 citations, averaging 7.60 citations per article, indicating strong productivity but limited influence per paper. The United States (Rank 3) shows a strong performance with 7,265 publications and 19,638 citations, averaging 19.70 cita-



**Figure 3:** Trends in article publications and citation averages.

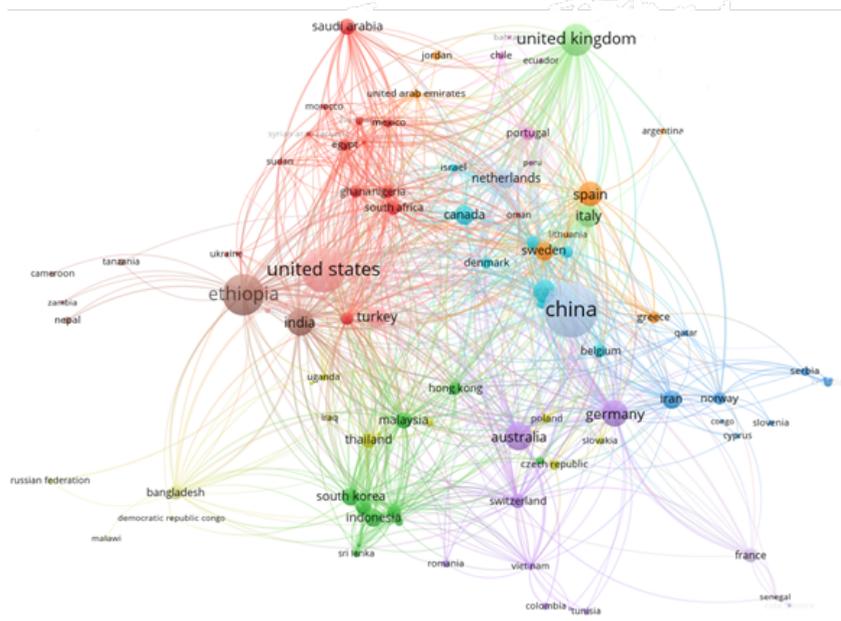
tions per article, reflecting its significant impact in the field. The United Kingdom (Rank 4) and Germany (Rank 5) also excel, with averages of 23.60 and 16.70 citations per article, respectively, indicating the high quality of their research. The UK's high average highlights its focus on producing influential studies. Italy (Rank 6) and Spain (Rank 7) also stand out, with averages of 19.10 and 21.90 citations per article, respectively. Spain's 3,046 publications and 9,765 citations emphasize its impact in the field. Australia (Rank 10) is notable with 2,496 publications and 7,697 citations, averaging 19.40 citations per article, indicating strong influence and quality.

Countries like China and Ethiopia could enhance their research impact by focusing on higher-quality studies and collaborations with leading researchers. Nations with lower average citations, such as Indonesia (Rank 17) and Nigeria (Rank 29), might improve by participating more in international collaborations and targeting higher-impact journals. The United Kingdom and Spain should continue to leverage their high citation impact by maintaining a focus on quality and exploring further global dissemination opportunities to sustain their influence in the field.

Table 4's data is visually represented in Figure 3, where darker shades on the map indicate higher levels of scientific production, clearly highlighting the global research contributions in binary logistic regression, with China and the U.S. leading in output.

### 3.5. Citation analysis.

Figure 4 represents a citation network research output by country. The visualization is based on the following parameters: Type of analysis - Citation, Unit of analysis - countries, with document weights and nine color clusters highlighting

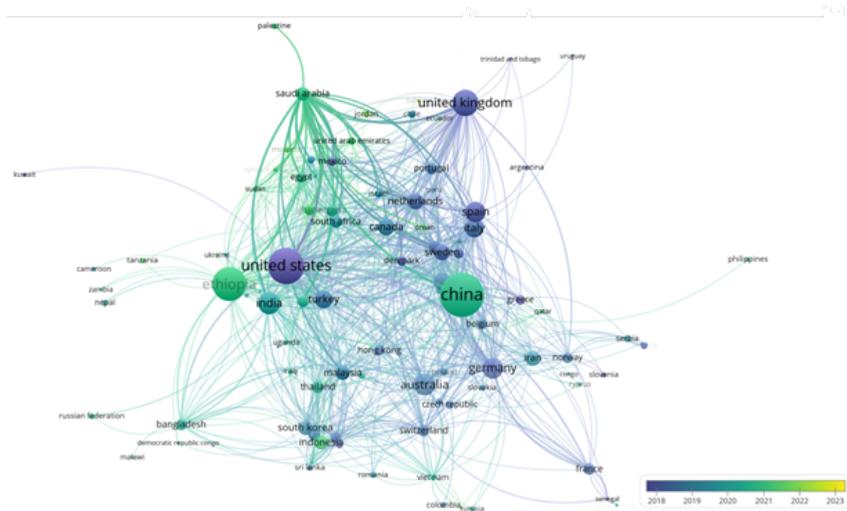


**Figure 4:** Citation network and distribution of research outputs by country.

the distribution and influence of various countries in the global research landscape related to binary logistic regression.

In terms of nodes and sizes, larger nodes represent countries with more substantial research output or greater influence in citations, with the United States, China, the United Kingdom, and Ethiopia standing out as key players. Smaller nodes, such as Norway or Bangladesh, indicate relatively lower contributions. Connections and lines show collaboration and citation relationships between countries. Thicker lines, such as those between the United States and China, suggest stronger citation links or collaborations, while denser regions reveal highly interconnected networks, particularly in clusters like the red cluster (United States, South Africa, Turkey) and blue cluster (China, Canada, Denmark).

Geographical and thematic clustering is also evident: countries in the same region, such as South Africa and Ghana or the U.K. and Italy, tend to be part of similar clusters (red and green, respectively), indicating regional collaborations. The clusters also highlight thematic connections in research outputs, with Africa (Ethiopia, Nepal) forming its network, potentially emphasizing regional interests in development and health, and Asia (South Korea, Malaysia, Indonesia) contributing to different aspects of logistic regression applications. This visualization provides a comprehensive overview of how binary logistic regression research is disseminated and interconnected on a global scale.



**Figure 5:** Trends in article publications and citation averages.

Figure 5 offers a dynamic view of the evolution of global research activity in binary logistic regression from 2018 to 2023. The color gradient, ranging from blue (2018) to yellow (2023), highlights the timeline of countries' contributions to this field. The United States, shown in darker blue, has been consistently active since 2018, maintaining a central position in the global research network. China, while emerging more prominently in recent years (closer to 2023), quickly established itself as a key contributor, forming strong citation and collaboration links with major players like the United States, Germany, and India.

Ethiopia, another notable player, also shows significant activity closer to 2023, reflecting an increasing involvement in the global research landscape. The United Kingdom, despite being active since 2018, exhibits slightly weaker connections than the United States and China but remains a crucial node in European collaborations. Overall, the visualization underscores China's rapid integration into the research network alongside established contributors like the United States, as well as the growing importance of Ethiopia in recent years. This temporal analysis reveals the expanding influence and evolving collaboration patterns in the study of binary logistic regression, with different countries rising to prominence over time.

### 3.6. Co-citation analysis.

Figure 6 presents a co-citation analysis based on cited authors in the field of binary logistic regression. This network visualization uses citation weights and displays 9 color-coded clusters, each representing distinct groups of authors frequently cited together. The red cluster, which includes highly cited authors such as Zhang Y., Wang Y., and Liu Y., forms the densest and most interconnected part of the net-

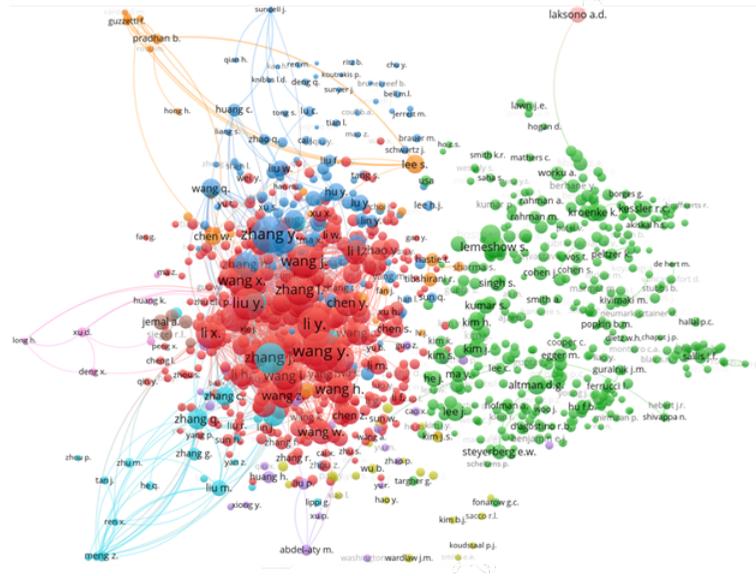


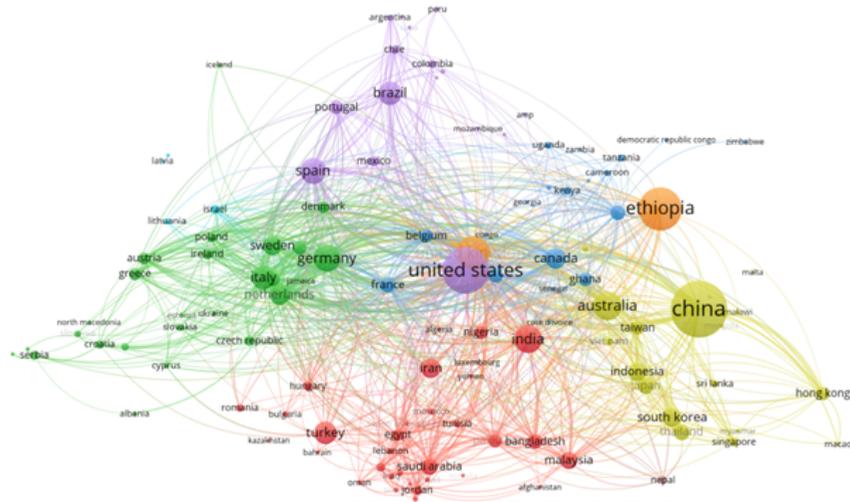
Figure 6: Citation distribution and co-citation network of cited authors.

work, reflecting a concentration of citations and a robust co-citation pattern. These authors appear to be central figures in the field, contributing significantly to research in logistic regression. The green cluster includes authors such as Lemeshow S., a notable figure in the development of statistical methods, particularly in logistic regression models. This cluster, while smaller in size compared to the red one, exhibits strong internal connections, indicating that these authors often appear together in the same citations.

The connections and lines between clusters and individual authors illustrate the intensity of co-citation relationships. Thicker lines, such as those linking key authors within the red cluster, including Zhang Y., Wang Y., and Liu Y., and within the green cluster, featuring authors like Lemeshow S. and Altman D.G., suggest a high degree of co-citation, pointing to established networks of academic influence. In contrast, smaller clusters, like the purple group with authors such as Long J.S. and the yellow group with Kim J.O., represent niche areas or emerging collaborations, indicating newer or more specialized fields of research. Overall, this figure provides a clear representation of the citation dynamics and co-citation patterns among leading authors in the domain, revealing both dominant contributors and their respective academic networks.

**3.7. Co-authorship analysis.**

Figure 7 illustrates a co-authorship analysis where the unit of analysis is countries, showcasing the global collaboration network. The plot uses document weights and



**Figure 7:** Global distribution and co-authorship network by countries.

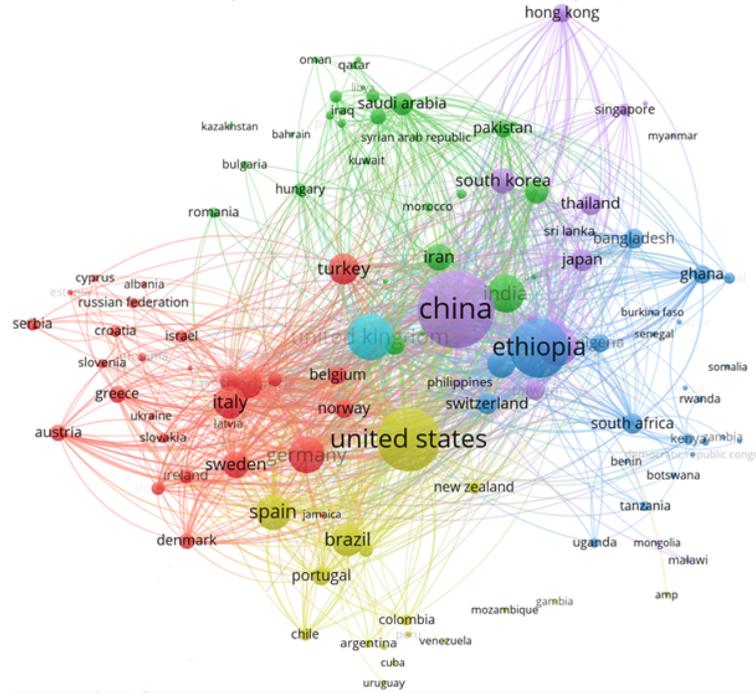
displays five distinct clusters, each represented by different colors, to highlight the collaborative relationships between countries in binary logistic regression research.

The United States is the most prominent node in the network, positioned centrally and connected to numerous countries across all clusters, indicating its extensive role in fostering international collaborations. China, another significant player, has formed strong co-authorship ties with countries in its cluster, including Taiwan, Indonesia, and South Korea, emphasizing its growing academic influence. Similarly, Ethiopia appears as a major node within its cluster, representing collaborations mainly with countries in Africa and select regions in Asia. The green cluster, consisting of countries such as Germany, Italy, and Sweden, reveals a dense European collaboration network. This cluster shows that European nations tend to co-author frequently with each other, further reinforcing the importance of regional collaboration. In the purple cluster, countries like Brazil, Spain, and Portugal dominate, showcasing the strength of collaborations within Latin America and Southern Europe. Connections and lines between countries represent the strength of co-authorship ties, with thicker lines, like those connecting the United States, Germany, and China, indicating stronger collaborative links. In contrast, thinner lines suggest less frequent collaboration or newer partnerships. This figure highlights both the global reach and regional concentrations of co-authorship networks in binary logistic regression research, showing how different countries contribute to and benefit from international research partnerships.

### 3.8. Bibliographic coupling analysis.

Figure 8 presents a bibliographic coupling analysis where the unit of analysis is countries, showcasing the global distribution of co-occurrences in academic pub-

lications. The plot employs occurrence weights and divides countries into seven clusters, each color-coded, to represent countries that cite similar sets of papers, indicating shared academic interests and influences.



**Figure 8:** Global distribution of occurrences and bibliographic coupling network.

The United States, as expected, appears as the largest node, indicating the highest number of occurrences and serving as a central hub within the network. China, Ethiopia, and India are also prominent nodes, particularly within their respective clusters, reflecting their growing influence in global academic discourse. The United Kingdom and Germany similarly occupy central positions, especially in relation to European and transatlantic collaborations. Each cluster highlights a region or group of countries with shared bibliographic coupling ties. The red cluster, including countries like Italy, Sweden, and Germany, indicates strong European collaboration networks, where frequent academic citations are shared among nations. The green cluster, featuring countries such as Saudi Arabia, Turkey, India, and Iran, represents academic relationships primarily in Asia and the Middle East. The blue cluster, with countries such as Ethiopia, Ghana, and South Africa, points to the strengthening bibliographic ties within Africa and with countries like China. Connections and lines between these clusters reflect the degree of bibliographic coupling, with thicker lines (such as those between the United States and China) suggesting stronger citation overlaps, while thinner lines indicate less frequent bibliographic coupling or emerging research relationships. In summary, the



often emphasizing model performance and accuracy. The yellow cluster incorporates terms such as body, smoking, and hypertension, suggesting research themes related to lifestyle factors and chronic diseases. These studies often focus on the biological and behavioral predictors of health outcomes.

#### 4. DISCUSSION

The bibliometric analysis reveals several strategic opportunities to advance binary logistic regression research. However, the study is constrained by its reliance on Scopus as the sole data source, which may exclude relevant research indexed in other databases. Encouraging international collaborations, particularly between established and emerging institutions, could enhance research diversity and innovation [11]. Expanding beyond journal articles to include more conference papers and book chapters would facilitate the sharing of preliminary findings and foster discussions that may not fit within traditional journals [42].

Additionally, integrating binary logistic regression with fields such as artificial intelligence and data science could drive growth and relevance in this area [48]. This expansion may involve developing new methodologies or adapting existing models to address challenges in these emerging fields. However, one challenge is the integration of such interdisciplinary approaches without losing the theoretical rigor of traditional statistical methods. Increasing the visibility and impact of this research through open access publishing and active participation in global research networks could further establish binary logistic regression as a key tool in both theoretical and applied research [46].

The analysis highlights strategies to enhance the impact of binary logistic regression research. One approach is to foster collaborative and interdisciplinary efforts, as the stable average citation rate suggests the need for more innovative studies that attract citations [11]. Promoting international partnerships can also produce more globally relevant work, increasing visibility and citation potential [20]. Enhancing dissemination through high-impact journals, conferences, social media, and academic networks can further boost visibility [14]. Exploring newer formats such as preprints or open-access repositories could also expand the reach of research outputs [11]. A limitation in current dissemination practices is the uneven access to such platforms across different regions, which could affect global reach. By focusing on quality and visibility, researchers can ensure their contributions are both significant and enduring in the field.

The findings reveal strategic opportunities for journals to boost their impact in binary logistic regression research. For journals with high output but lower citation impact, such as *Frontiers in Public Health* and *BMC Women's Health*, focusing on high-quality, innovative research is crucial [12, 13]. This may involve stricter peer reviews, targeted calls for papers in emerging fields, and better dissemination support through academic networks and social media [11]. However, a challenge for these journals may be attracting high-quality submissions while competing with

more established publications. Interdisciplinary research is also key. Journals with a niche audience could benefit from encouraging submissions that cross disciplinary boundaries, attracting a broader readership and more citations [14]. Promoting international collaborations among authors can further enhance global relevance and attract broader attention [29]. By prioritizing quality, interdisciplinary reach, and global relevance, these journals can significantly enhance their influence in binary logistic regression research.

The findings suggest strategies for authors in binary logistic regression to boost their academic impact. Authors with high output but lower citation rates could focus on interdisciplinary research and publish in higher-impact journals to reach broader audiences [15]. Engaging in international collaborations and academic networks can also increase visibility and citations [11]. Mentorship and collaboration among top authors is another key strategy. Established researchers could co-author with emerging scholars to elevate the profile of less-cited work while providing valuable mentorship [12]. This also helps junior researchers learn from seasoned experts, improving research quality. However, balancing mentorship with the demands of publishing high-quality research may pose a challenge. Additionally, focusing on emerging trends and incorporating new methodologies can help authors maintain leadership and advance the field [25].

The findings reveal strategic opportunities for countries to enhance their impact in binary logistic regression research. For China and Ethiopia, which have high publication volumes but lower citation rates, focusing on improving research quality and visibility is crucial [14]. International collaborations with countries like the United States and the United Kingdom, which have higher citation averages, could help increase the global reach and impact of their work [11, 13]. Countries with lower citation averages could also target innovative or emerging areas in binary logistic regression, where competition is lower, and potential impact is higher [8]. Investing in cutting-edge research and publishing in high-impact journals can enhance their academic standing [12]. For leading nations like the United States, the United Kingdom, and Spain, continuing to produce high-quality research while exploring interdisciplinary opportunities could sustain and elevate their influence in the field [25].

The visual analysis suggests strategies to enhance global collaboration in binary logistic regression research. For Ethiopia, expanding research networks through stronger international collaborations, particularly with leading institutions in the U.S. and China, could significantly boost visibility and impact. Increased participation in global research consortia and international conferences would also expose Ethiopian researchers to cutting-edge developments and collaboration opportunities. For the U.S. and China, maintaining and strengthening leadership positions is key. The U.S. should continue fostering interdisciplinary collaborations that address emerging challenges in binary logistic regression. China should focus on consolidating its growing network by deepening collaborations with both established and emerging research countries. This could involve joint research projects,

increased funding for partnerships, and promoting research findings through high-impact publications. These strategies will help both countries stay at the forefront of global research in this critical area, driving innovation and shaping future research directions.

#### LIMITATIONS

Although the discussion emphasizes opportunities for authors, journals, and countries, the results should be interpreted in light of certain methodological limitations. These constraints do not diminish the value of the findings but rather point to avenues for refining future bibliometric research. First, the analysis is based exclusively on documents indexed in Scopus, which may exclude relevant publications available in other databases such as Web of Science or PubMed. Second, the study focuses primarily on English language documents, potentially overlooking significant contributions published in other languages. Third, the bibliometric approach emphasizes quantitative indicators (e.g., publication counts, citations, co-authorship networks) rather than the detailed technical content of individual methodologies. Consequently, while the results provide a comprehensive overview of global research trends, they should be interpreted with these constraints in mind.

Despite these limitations, the findings offer a reliable and informative map of the intellectual landscape of binary logistic regression research, contributing to a better understanding of its evolution, disciplinary applications, and emerging intersections with artificial intelligence.

#### 5. CONCLUSIONS

The bibliometric analysis of binary logistic regression research provides deep insights into its evolution and current status. Over the past 50 years, the research volume has significantly increased, establishing this technique as a crucial tool across various disciplines. The trend towards co-authorship reflects the interdisciplinary and global nature of the field, highlighting the complexity of the problems addressed and the collective efforts to find solutions. The predominance of journal articles underscores the rigorous peer-review standards in this area. While certain countries lead in research output and impact, the high level of international collaboration demonstrates that emerging nations also play a vital role, enriching research diversity and fostering innovative applications. While this study has acknowledged several methodological constraints (discussed in detail in the Limitations section) the overall findings provide a reliable and comprehensive map of binary logistic regression research.

Publication trends in binary logistic regression research show a significant increase, especially in the last decade, indicating its growing relevance across scientific disciplines. However, the stable citation rate per article suggests a potential disconnect between the research volume and its impact, which may indicate that

while the field is expanding, the influence of individual studies is not advancing at the same rate. One limitation of this study is the exclusion of certain high-impact studies that may not be well-cited within the Scopus database, leading to potential biases in interpreting citation impacts. This pattern reflects a common situation in rapidly expanding fields, where more publications do not always lead to greater academic recognition. The data suggest the need for more focused and impactful studies that stand out in an increasingly crowded field.

The analysis of the co-occurrence of keywords, both author-specific and indexed, reveals the central role of demographic and clinical terms such as “human,” “female,” “male,” and “aged.” These keywords highlight a strong focus on health-related outcomes and the application of binary logistic regression in epidemiological studies, particularly in understanding risk factors within specific populations. The interdisciplinary nature of this research is evident, with a significant overlap of themes across various fields such as public health, epidemiology, and mental health. While the United States, China, and the United Kingdom lead in shaping these research themes, the involvement of emerging countries like Ethiopia shows a broader global participation in advancing this field. Nevertheless, the study’s focus on dominant keywords may overshadow emerging terms and concepts that are not yet widely recognized, potentially limiting the scope of the analysis. However, this thematic concentration also suggests that certain research areas might be under-represented, indicating an opportunity to explore new topics and methodologies to further advance the field.

The analysis of top journals and authors highlights those who excel in both productivity and impact. Journals like *Plos One* and *BMC Public Health* balance quantity with quality, making significant contributions to the field. In contrast, journals with high publication counts but lower citation impacts might need to focus more on quality to enhance their academic influence. Similarly, while top authors like *Wang Y* and *Zhang Y* lead in both output and citations, others may benefit from targeting more specialized or emerging areas to increase their impact. The bibliometric analysis of the top 30 countries reveals significant differences in productivity and impact. While China leads in publication volume, the United States and the United Kingdom demonstrate higher citation averages, suggesting a trade-off between quantity and quality. Countries like Spain and Australia maintain a balance, indicating strong and respected research output, whereas nations like Ethiopia and Brazil face challenges in achieving broader recognition despite high publication counts. This variation underscores the importance of producing quality research that resonates globally. Future studies should consider analyzing the content and quality of publications to better understand the reasons behind variations in citation impact among different countries.

The visual analysis of global research collaborations in binary logistic regression reveals significant differences among the networks of the United States, China, and Ethiopia. The United States stands out as the most influential, with strong connections to other leading research nations, reflecting its central role in global

scientific collaboration. China, although rapidly expanding its network, is still developing in comparison to the U.S., while Ethiopia, as an emerging contributor, shows potential for growth by expanding its collaborations with more established research nations. A limitation of this analysis is that it focuses primarily on publication quantity rather than the qualitative aspects of these collaborations, which could provide a more nuanced understanding of their impact. The analysis of bibliographic coupling further highlights the tightly connected global network, with distinct research themes explored by different country groups, pointing to a division of labor or specialization in the field.

In summary, while binary logistic regression research is mature with established directions, opportunities remain for exploring emerging topics, enhancing interdisciplinary collaborations, and improving the quality and impact of research outputs globally. Looking ahead, future research should integrate additional data sources, expand the linguistic scope beyond English, and deepen the qualitative understanding of collaboration networks. These directions will ensure that binary logistic regression continues to evolve as a core methodological tool across disciplines, particularly at the intersection of statistics and artificial intelligence.

#### AUTHOR CONTRIBUTIONS

Conceptualization by Humberto Llinás and Jorge Villalba; methodology by Humberto Llinás; research by Jorge Tilano and Daniela Nuñez; writing-original draft preparation by Humberto Llinás; writing-review and editing by Humberto Llinás and Jorge Villalba; visualization by Daniela Nuñez; supervision by Jorge Villalba; project administration by Humberto Llinás. All authors have read and agreed to the published version of the manuscript.

#### FUNDING

This research received no external funding.

#### ACKNOWLEDGMENTS

We acknowledge Universidad del Norte for their support in this research.

#### DATA AVAILABILITY STATEMENT

The data presented in this study are available on request from the corresponding author.

#### CONFLICTS OF INTEREST

The author declares no conflicts of interest.

## REFERENCES

- [1] B. O. Ahinkorah, *Predictors of modern contraceptive use among adolescent girls and young women in Sub-Saharan Africa: a mixed effects multilevel analysis of data from 29 demographic and health surveys*. *Contraception and Reproductive Medicine* **5**(2020), no. 1, 32. DOI: [10.1186/s40834-020-00138-1](https://doi.org/10.1186/s40834-020-00138-1)
- [2] M. Ahsan, T. Y. Susanto, T. A. Virania, A. I. Jaya, *Credit card fraud detection using linear discriminant analysis (LDA), random forest, and binary logistic regression*. *Barekeng* **16**(2022), no. 4, 1337–1346. DOI: [10.30598/barekengvol16iss4pp1337-1346](https://doi.org/10.30598/barekengvol16iss4pp1337-1346)
- [3] R. Aleixandre-Benavent et al., *Bibliometría e indicadores de actividad científica (II). Indicadores de producción científica en pediatría*. *Acta Pediátrica Español* **75**(2017), no. 3-4, 44–50.
- [4] Z. Y. Algamal, M. H. Lee, *Applying penalized binary logistic regression with correlation based elastic net for variables selection*. *Journal of Modern Applied Statistical Methods* **14**(2015), no. 1, 168–179. DOI: [10.22237/jmasm/1430453640](https://doi.org/10.22237/jmasm/1430453640)
- [5] M. Alhajj et al., *Bibliometric analysis and evaluation of the Journal of Prosthetic Dentistry from 1970 to 2019*. *The Journal of Prosthetic Dentistry* **129**(2023), no. 2, 323–340. DOI: [10.1016/j.prosdent.2021.05.013](https://doi.org/10.1016/j.prosdent.2021.05.013)
- [6] A. M. Alharthi, M. H. Lee, Z. Y. Algamal, *Weighted L1-norm logistic regression for gene selection of microarray gene expression classification*. *International Journal on Advanced Science, Engineering and Information Technology* **10**(2020), no. 4, 1483–1488. DOI: [10.18517/ijaseit.10.4.10907](https://doi.org/10.18517/ijaseit.10.4.10907)
- [7] E. M. Ali, M. M. Ahmed, S. S. Wulff, *Detection of critical safety events on freeways in clear and rainy weather using SHRP2 naturalistic driving data: parametric and non-parametric techniques*. *Safety Science* **119**(2019), 141–149. DOI: [10.1016/j.ssci.2019.01.007](https://doi.org/10.1016/j.ssci.2019.01.007)
- [8] M. Aria, C. Cuccurullo, *bibliometrix: An R-tool for comprehensive science mapping analysis*. *Journal of informetrics* **11**(2017), no. 4, 959–975. DOI: [10.1016/j.joi.2017.08.007](https://doi.org/10.1016/j.joi.2017.08.007)
- [9] A. Bhattacharya, S. Biswas, A. Mandal, *Credit risk evaluation: a comprehensive study*. *Multimedia Tools and Applications* **82**(2022), 18217–18267. DOI: [10.1007/s11042-022-13952-3](https://doi.org/10.1007/s11042-022-13952-3)
- [10] W. Deng, J. Wang, *The effect of entrepreneurship education on the entrepreneurial intention of different college students: Gender, household registration, school type, and poverty status*. *Plos One* **18**(2023), no. 7, e0288825. DOI: [10.1371/journal.pone.0288825](https://doi.org/10.1371/journal.pone.0288825)
- [11] N. Donthu, S. Kumar, D. Mukherjee, W. M. Lim, *How to conduct a bibliometric analysis: An overview and guidelines*. *Journal of Business Research* **133**(2021), 285–296. DOI: [10.1016/j.jbusres.2021.04.070](https://doi.org/10.1016/j.jbusres.2021.04.070)

- [12] O. Ellegaard, J. A. Wallin, *The bibliometric analysis of scholarly production: How great is the impact?* English. *Scientometrics* **105**(2015), no. 3, 1809–1831. DOI: [10.1007/s11192-015-1645-z](https://doi.org/10.1007/s11192-015-1645-z)
- [13] Y.-n. Gan, D.-d. Li, N. Robinson, J.-p. Liu, *Practical guidance on bibliometric analysis and mapping knowledge domains methodology – A summary*. *European Journal of Integrative Medicine* **56**(2022), 102203. DOI: [10.1016/j.eujim.2022.102203](https://doi.org/10.1016/j.eujim.2022.102203)
- [14] C. García-Villar, J. García-Santos, *Bibliometric indicators to evaluate scientific activity*. *Radiología (English Edition)* **63**(2021), no. 3, 228–235. DOI: [10.1016/j.rxeng.2021.01.002](https://doi.org/10.1016/j.rxeng.2021.01.002)
- [15] M. Gaviria-Marin, J. M. Merigo, S. Popa, *Twenty years of the Journal of Knowledge Management: a bibliometric analysis*. *Journal of Knowledge Management* **22**(2018), no. 8, 1655–1687. DOI: [10.1108/JKM-10-2017-0497](https://doi.org/10.1108/JKM-10-2017-0497)
- [16] A. Gonzales-Aguilar, M.-J. Colmenero-Ruiz, F.-C. Paletta, L. Verlaet, *Loet Leydesdorff: bibliometric analysis and mapping of his scientific production*. *Profesional de la información* **32**(2023), no. 7. DOI: [10.3145/epi.2023.dic.09](https://doi.org/10.3145/epi.2023.dic.09)
- [17] E. Herrera-Viedma, M. A. Martínez, M. Herrera, *Bibliometric tools for discovering information in database*. 29th International Conference on Industrial Engineering and Other Applications of Applied Intelligent Systems. Springer. 2016, 193–203. DOI: [/10.1007/978-3-319-42007-3\\_17](https://doi.org/10.1007/978-3-319-42007-3_17)
- [18] K. Ingale, R. Paluri, L. Bonfigt, *Financial literacy and financial behaviour: a bibliometric analysis*. *Review of Behavioral Finance* **14**(2022), no. 1, 130–154. DOI: [10.1108/RBF-06-2020-0141](https://doi.org/10.1108/RBF-06-2020-0141)
- [19] N. Jayaratne, R. Zwahlen, *The evolution of dental journals from 2003 to 2012: a bibliometric analysis*. *PloS One* **10**(2015), no. 3, e0119503. DOI: [10.1371/journal.pone.0119503](https://doi.org/10.1371/journal.pone.0119503)
- [20] M. Khan et al., *Value of special issues in the Journal of Business Research: a bibliometric analysis*. *Journal of Business Research* **125**(2021), 295–313. DOI: [10.1016/j.jbusres.2020.12.015](https://doi.org/10.1016/j.jbusres.2020.12.015)
- [21] S. Kim et al., *What predicts people’s belief in COVID-19 misinformation? A retrospective study using a nationwide online survey among adults residing in the United States*. *BMC Public Health* **22**(2022), no. 1, 2114. DOI: [10.1186/s12889-022-14431-y](https://doi.org/10.1186/s12889-022-14431-y)
- [22] Á. Kocsis, G. Molnár, *Factors influencing academic performance and dropout rates in higher education*. *Oxford Review of Education* **51**(2024), no. 3, 414–432. DOI: [10.1080/03054985.2024.2316616](https://doi.org/10.1080/03054985.2024.2316616)
- [23] B. W. Koo, S. Guhathakurta, N. Botchwey, *How are neighborhood and street-level walkability factors associated with walking behaviors? A big data approach using street view images*. *Environment and Behavior* **54**(2022), no. 1, 211–241. DOI: [10.1177/00139165211014609](https://doi.org/10.1177/00139165211014609)

- [24] S. Kumar, V. Gota, *Logistic regression in cancer research: A narrative review of the concept, analysis, and interpretation*. *Cancer Research, Statistics, and Treatment* **6**(2023), no. 4, 573–578. DOI: [10.4103/crst.crst\\_293\\_23](https://doi.org/10.4103/crst.crst_293_23)
- [25] M. K. Lazarides, I.-Z. Lazaridou, N. Papanas, *Bibliometric analysis: bridging informatics with science*. *The international journal of lower extremity wounds* **24**(2025), no. 3, 515–517. DOI: [10.1177/15347346231153538](https://doi.org/10.1177/15347346231153538)
- [26] C. Leal Iga, M. T. Cedilla Salazar, *Regularización y mercado de suelo urbano en asentamientos irregulares: el caso Cima de La Loma, Monterrey, México*. *Contexto* **15**(2021), no. 23, 89–106. DOI: [10.29105/contexto15.23-346](https://doi.org/10.29105/contexto15.23-346)
- [27] J. Li, C. Xu, B. Feng, H. Zhao, *Credit risk prediction model for listed companies based on CNN-LSTM and attention mechanism*. *Electronics* **12**(2023), 1643. DOI: [10.3390/electronics12071643](https://doi.org/10.3390/electronics12071643)
- [28] Y. Li, F. Lu, Y. Yin, *Applying logistic LASSO regression for the diagnosis of atypical Crohn's disease*. *Scientific Reports* **12**(2022), no. 1, 11340. DOI: [10.1038/s41598-022-15609-5](https://doi.org/10.1038/s41598-022-15609-5)
- [29] W. M. Lim, S. Kumar, *Guidelines for interpreting the results of bibliometric analysis: A sensemaking approach*. *Global Business and Organizational Excellence* **43**(2024), no. 2, 17–26. DOI: [10.1002/joe.22229](https://doi.org/10.1002/joe.22229)
- [30] L. Lombardo, P. M. Mai, *Presenting logistic regression-based landslide susceptibility results*. *Engineering Geology* **244**(2018), 14–24. DOI: [10.1016/j.enggeo.2018.07.019](https://doi.org/10.1016/j.enggeo.2018.07.019)
- [31] C. D. Lopez et al., *Using machine learning methods to predict nonhome discharge after elective total shoulder arthroplasty*. *JSES International* **5**(2021), no. 4, 692–698. DOI: [10.1016/j.jseint.2021.02.011](https://doi.org/10.1016/j.jseint.2021.02.011)
- [32] J. López Guauque, A. Gil-Lafuente, *Fifty years of fuzzy research: A bibliometric analysis and a long-term comparative overview*. *Journal of Intelligent & Fuzzy Systems* **38**(2020), no. 5, 5413–5425. DOI: [10.3233/JIFS-179634](https://doi.org/10.3233/JIFS-179634)
- [33] T. Luan et al., *Dynamic risk analysis of flammable liquid road tanker based on fuzzy Bayesian network*. *Process Safety Progress* **42**(2023), no. 4, 737–751. DOI: [10.1002/prs.12508](https://doi.org/10.1002/prs.12508)
- [34] J. Ma et al., *Poor handling of continuous predictors in clinical prediction models using logistic regression: a systematic review*. *Journal of Clinical Epidemiology* **161**(2023), 140–151. DOI: [10.1016/j.jclinepi.2023.07.017](https://doi.org/10.1016/j.jclinepi.2023.07.017)
- [35] Q. Ma, *Recent applications and perspectives of logistic regression modelling in healthcare*. *Theoretical and Natural Science* **36**(2024), 185–190. DOI: [10.54254/2753-8818/36/20240614](https://doi.org/10.54254/2753-8818/36/20240614)
- [36] C. A. Magee, P. C. L. Heaven, *Big-five personality factors, obesity and 2-year weight gain in Australian adults*. *Journal of Research in Personality* **45**(2011), no. 3, 332–335. DOI: [10.1016/j.jrp.2011.02.009](https://doi.org/10.1016/j.jrp.2011.02.009)

- [37] A. T. Manjatika, J. G. Davimes, P. Mazengeny, *Estimation of sex using dimensions around the metatarsal diaphyseal nutrient foramen: Application of discriminant function analysis and logistic regression models*. *Legal Medicine* **68**(2024), 102417. DOI: [10.1016/j.legalmed.2024.102417](https://doi.org/10.1016/j.legalmed.2024.102417)
- [38] E. O. Oyerogba, *Forensic auditing mechanism and fraud detection: the case of Nigerian public sector*. *Journal of Accounting in Emerging Economies* **11**(2021), no. 5, 752–775. DOI: [10.1108/JAEE-04-2020-0072](https://doi.org/10.1108/JAEE-04-2020-0072)
- [39] M. J. Page et al., *The PRISMA 2020 statement: an updated guideline for reporting systematic reviews*. *BMJ* **372**(2021). DOI: [10.1136/bmj.n71](https://doi.org/10.1136/bmj.n71)
- [40] N. R. Panda, J. K. Pati, J. N. Mohanty, R. Bhuyan, *A review on logistic regression in medical research*. *National Journal of Community Medicine* **13**(2022), no. 04, 265–270.
- [41] N. G. Polson, J. G. Scott, *Data augmentation for non-Gaussian regression models using variance-mean mixtures*. *Biometrika* **100**(2013), no. 2, 459–471. DOI: [10.1093/biomet/ass081](https://doi.org/10.1093/biomet/ass081)
- [42] R. Prancutè, *Web of Science (WoS) and Scopus: the titans of bibliographic information in today's academic world*. *Publications* **9**(2021), no. 1, 12. DOI: [10.3390/publications9010012](https://doi.org/10.3390/publications9010012)
- [43] Z. Runchi, L. Xue, W. Qin, *An ensemble credit scoring model based on logistic regression with heterogeneous balancing and weighting effects*. *Expert Systems with Applications* **212**(2022), 118732. DOI: [10.1016/j.eswa.2022.118732](https://doi.org/10.1016/j.eswa.2022.118732)
- [44] A. Shaikh et al., *Bibliometric analysis on the adoption of artificial intelligence applications in the e-health sector*. *Digital Health* **9**(2023), 205520762211492. DOI: [10.1177/20552076221149296](https://doi.org/10.1177/20552076221149296)
- [45] Y. X. Song et al., *Comparison of logistic regression and machine learning methods for predicting postoperative delirium in elderly patients: a retrospective study*. *CNS Neuroscience & Therapeutics* **29**(2023), no. 1, 158–167. DOI: [10.1111/cns.13991](https://doi.org/10.1111/cns.13991)
- [46] N. Suresh, B. Thomas, J. Joseph, *Bibliometric Analysis and Visualization of Scientific Literature on Heart Disease Classification Using a Logistic Regression Model*. *Cureus* **16**(2024), no. 6, e63337. DOI: [10.7759/cureus.63337](https://doi.org/10.7759/cureus.63337)
- [47] A. Syed, H. Bawazir, *Recent trends in business financial risk – A bibliometric analysis*. *Cogent Economics & Finance* **9**(2021), no. 1, 1913877. DOI: [10.1080/23322039.2021.1913877](https://doi.org/10.1080/23322039.2021.1913877)
- [48] D. M. Tom, *Logistic regression teams up with artificial intelligence to beat neural network and gradient boosted machine*. 2021. DOI: [10.31224/osf.io/vz496](https://doi.org/10.31224/osf.io/vz496)
- [49] A. Vaidya, *Predictive and probabilistic approach using logistic regression: Application to prediction of loan approval*. 2017 8th International Conference on Computing, Communication and Networking Technologies. 2017, 1–6. DOI: [10.1109/ICCCNT.2017.8203946](https://doi.org/10.1109/ICCCNT.2017.8203946)

- [50] S. Valenzuela, N. M. Somma, A. Scherman, A. Arriagada, *Social media in Latin America: deepening or bridging gaps in protest participation?* Online Information Review **40**(2016), no. 5, 695–711. DOI: [10.1108/OIR-11-2015-0347](https://doi.org/10.1108/OIR-11-2015-0347)
- [51] M. G. Vaughn et al., *Toward a criminal justice epidemiology: behavioral and physical health of probationers and parolees in the United States*. Journal of Criminal Justice **40**(2012), no. 3, 165–173. DOI: [10.1016/j.jcrimjus.2012.03.001](https://doi.org/10.1016/j.jcrimjus.2012.03.001)
- [52] J. Venema, D. Peula, J. Irurita, P. Mesejo, *Employing deep learning for sex estimation of adult individuals using 2D images of the humerus*. Neural Computing and Applications **35**(2023), no. 8, 5987–5998. DOI: [10.1007/s00521-022-07981-0](https://doi.org/10.1007/s00521-022-07981-0)
- [53] L. Waltenberger et al., *Lateral angle: A landmark-based method for the sex estimation in human cremated remains and application to an Austrian prehistoric sample*. American journal of biological anthropology **184**(2024), no. 1, e24874. DOI: [10.1002/ajpa.24874](https://doi.org/10.1002/ajpa.24874)
- [54] C. Wang, M. K. Lim, A. Lyons, *Twenty years of the International Journal of Logistics Research and Applications: a bibliometric overview*. International Journal of Logistics **22**(2018), no. 3, 304–323. DOI: [10.1080/13675567.2018.1526262](https://doi.org/10.1080/13675567.2018.1526262)
- [55] H. Wang et al., *Bibliometric Analysis on the Progress of Chronic Heart Failure*. Current Problems in Cardiology **47**(2022), no. 9, 101213. DOI: [10.1016/j.cpcardiol.2022.101213](https://doi.org/10.1016/j.cpcardiol.2022.101213)
- [56] A. Yildiz, *Determining the factors for individual credit approval by applying logistic regression and hierarchical logistic regression*. Int. Journal of Management Studies and Social Science Research **5**(2023), no. 6, 58–67. DOI: [10.56293/ijmsssr.2023.4705](https://doi.org/10.56293/ijmsssr.2023.4705)
- [57] J. Yoon, S. Lee, *Spatio-temporal patterns in pedestrian crashes and their determining factors: application of a space-time cube analysis model*. Accident Analysis and Prevention **161**(2021), 106291. DOI: [10.1016/j.aap.2021.106291](https://doi.org/10.1016/j.aap.2021.106291)
- [58] Y. Zhou et al., *Coronary Heart Disease and Depression or Anxiety: A Bibliometric Analysis*. Frontiers in Psychology **12**(2021), 669000. DOI: [10.3389/fpsyg.2021.669000](https://doi.org/10.3389/fpsyg.2021.669000)
- [59] S. Zihan, S.-H. Sung, D.-M. Park, B.-K. Park, *All-year dropout prediction modeling and analysis for university students*. Applied Sciences **13**(2023), no. 2, 1143. DOI: [10.3390/app13021143](https://doi.org/10.3390/app13021143)
- [60] M. Zuckerman et al., *Biocultural perspectives on bioarchaeological and paleopathological evidence of past pandemics*. American Journal of Biological Anthropology **182**(2023), no. 4, 557–582. DOI: [10.1002/ajpa.24647](https://doi.org/10.1002/ajpa.24647)