Characterization and mapping of the scientific production indexed in Scopus (2013-2022) regarding tuberculosis in vulnerable populations

Cristina Torres-Pascual[1,2], Cristina Gordillo-Marroquín[2,3,4], Héctor Javier Sánchez-Pérez[2,3,4]

[1]Escola Universitària de la Salut i l'Esport. Universitat de Girona. España. [2]Grupos de Investigación para América y África Latinas (GRAAL). [3]El Colegio de La Frontera Sur (ECOSUR). Health Department. San Cristóbal de Las Casas. Chiapas. Mexico. [4]Red Mexicana de Investigación en Tuberculosis y otras Micobacteriosis AC (REMexTB).

Summary

Tuberculosis (TB) represents a serious public health issue, especially in populations with higher economic and social vulnerabilities. In this study, a characterization and bibliometric mapping of the scientific production between 2013 and 2022 regarding TB in vulnerable populations was performed, using Scopus (2013-2022) for the databases retrieval. Original articles and reviews published in any language and referring to TB in vulnerable populations were included. We analyzed production, authorship, collaborations, research topics and the most prolific journals. Metadata was analyzed with Bibliometrix/RStudio and VOSviewer tools. A total of 502 articles were retrieved. The author with more cited publications in 2022 was GB Migliori, from the Istituti Clinici Scientifici Maugeri IRCCS (80 citations). The Universidade de São Paulo (Brazil) was the institution with the highest number of works (18%). The United States led in scientific production and collaborations; the National Institutions of Health (USA) were the primary funding agency. The journal with the highest percentage of published articles, citations and h-index was the *International Journal of Tuberculosis and Lung Diseases*. Greater cooperation between countries, authors and institutions is essential, especially among those nations with the highest TB burden and low economic development.

Key words:
Tuberculosis. Bibliometrics. Scopus. Vulnerable populations. Poverty.

Caracterización y mapeo de la producción científica indexada en Scopus (2013-2022) sobre tuberculosis en poblaciones vulnerables

Resumen

La tuberculosis (TB) representa un grave problema de salud pública, especialmente en poblaciones con mayor vulnerabilidad económica y social. En este estudio se realizó una caracterización y mapeo bibliométrico de la producción científica entre 2013 y 2022 sobre TB en poblaciones vulnerables, utilizando Scopus (2013-2022) para la recuperación de bases de datos. Se incluyeron artículos originales y revisiones publicados en cualquier idioma referidos a la TB en poblaciones vulnerables. Se analizó la producción, la autoría, las colaboraciones, los temas de investigación y las revistas más prolíficas. Los metadatos fueron analizados con las herramientas Bibliometrix/RStudio y VOSviewer. El autor con más publicaciones citadas en 2022 fue GB Migliori, del Istituti Clinici Scientifici Maugeri IRCCS (80 citas). La Universidade de São Paulo (Brasil) fue la institución con el mayor número de trabajos (18%). Estados Unidos lideró la producción científica y las colaboraciones; los Institutos Nacionales de Salud de Estados Unidos fueron el principal organismo financiador. La revista con mayor porcentaje de artículos publicados, citas e índice-h fue la *International Journal of Tuberculosis and Lung Diseases*. Se concluye que es esencial una mayor cooperación entre países, autores e instituciones, especialmente de las naciones con mayor carga de tuberculosis y bajo desarrollo económico.

Palabras clave:
Tuberculosis. Bibliometría. Scopus. Poblaciones vulnerables. Pobreza.

Correspondence: Héctor Javier Sánchez-Pérez. E-mail: hsanchez@ecosur.mx
Cristina Gordillo-Marroquín. E-mail: cmgordillo@ecosur.mx

Introduction

Tuberculosis

Tuberculosis (TB) continues to be one of the world's most important infectious diseases and a serious threat to public health. Worldwide, an estimated 10.6 million people (range, 9.9-11.4 million) fell ill with TB in 2022 – increasing the number of people affected by TB (PATB) compared to what was reported in previous years (10.3 million in 2021 and 10.0 million in 2020). This led to the death of about 1.3 million people (range, 1.18-1.43 million)[1].

The World Health Organization (WHO) estimates that, globally, nearly half a million excess deaths occurred because of TB between 2020 and 2022 (in comparison to trends observed in 2019), due to delays in TB diagnosis and treatment caused by disruptions related to the COVID-19 pandemic. In other words, COVID-19 further aggravated underlying inequalities in healthcare for PATB, which has impacted with a reduction in access and delivery of health services, especially for the most vulnerable populations[1].

Vulnerable populations

About a quarter of the world's population is infected with TB (latent-TB, which is the asymptomatic and non-contagious form of TB), of which 5% to 10% will develop the disease at some point (active-TB, symptomatic and transmissible in its pulmonary form [PTB]), concentrating the highest rates of TB incidence and deaths in developing and low-income countries[2].

TB does not affect everyone equally; the risk of getting sick increases according to both the so-called social determinants of health (SDH), as well as certain individual characteristics that put people in a greater risk of developing the active form of TB, such as diabetes mellitus and HIV-AIDS (both conditions occurring with immunosuppression), malnutrition, alcoholism, tobacco use[2], poverty, legal status, and structural aspects such as education, gender, and ethnicity, among others[3-6].

In this sense, TB has been considered a social disease, due to its correlation with poverty, even before the discovery of its etiological agent (*Mycobacterium tuberculosis*) in 1882, whereas the high incidence rates of TB occurred among the working classes or in the poorest families[7].

From the SDH approach, populations that are considered vulnerable to TB can be defined as "people whose context leads to disadvantaged socioeconomic positions that put them at systematically higher risk for TB, with limited access to appropriate or high-quality TB care, thus with a higher likelihood of experiencing health inequalities, developing TB infection or progression to TB disease"[8]. In addition, these populations are usually living in situations or contexts of inequality, prejudice, marginalization, and barriers within their socioeconomic as well as cultural life[5].

The WHO has emphasized the need for socioeconomic interventions to reduce barriers of access to TB care and address the social determinants of TB[9]. Nevertheless, in addition to universal health coverage – to ensure that all PATB have access to care and treatment for TB (whether latent or active) – it is essential to develop multisectoral actions that address the aforementioned set of factors[1].

Hence, since most of the TB burden is sustained by specific vulnerable populations, having a deeper understanding of these groups will allow to boost research plans and approaches for these populations, to have better tools that allow to develop strategies and public policies that favor their care.

Bibliometrics

Bibliometric analysis is a tool used to assess the productivity and development of research works in a specific field, such as health sciences[10]. This tool focuses on the use of data obtained through publications of the area of study. The data is analyzed through quantitative and statistical research, in order to explore, summarize, visualize, and characterize a set of publications[11].

Background and objectives

The purpose of this study was to retrieve data from international publications regarding TB in vulnerable populations from 2013 to 2022 and to conduct the characterization and mapping of this scientific production. Although there are registers of bibliometric studies on TB research[10,12,13], these have not focused on vulnerable populations. Considering the high incidence and mortality rate of TB[1], due importance needs to be given to research in this type of populations[4].

Bibliometric analysis is a tool that has proven useful for examining the results of globally published scientific works in order to generate knowledge and establish better research strategies and opportunities for project collaborations that can support new approaches and policies to fill research gaps. For this reason, bibliometric studies are one of the main tools used for science policy decision-making to evaluate research performance and guide the allocation of funding[14].

Material and method

Study design

Retrospective descriptive study with a bibliometric approach.

Data source and research strategy

Data was obtained from the Scopus database, which was chosen because it is the largest academic database of citations and abstracts, with a larger number of active journals covering worldwide current and relevant research[15,16].

Data collection was conducted on October 20, 2023, and was restricted from January 1, 2013, to December 31, 2022. The

bibliographic search strategy was structured as follows: (TITLE-ABS-KEY ["health disparate minority vulnerable populations" OR "vulnerable populations" OR "vulnerable communities" OR "social determinants of health" OR "Vulnerable groups" OR "Social vulnerability" OR "Sensitive Population*" OR "Minority Population*"] AND TITLE-ABS-KEY [tuberculosis OR "mycobacterium tuberculosis"]).

Inclusion and exclusion criteria

Articles published in any language and those referring to TB in vulnerable populations (women, children, the elderly, people with another TB-associated disease such as diabetes, HIV or malnutrition, persons deprived of their liberty, migrants, refugees, people from developing countries and disadvantaged socioeconomic situations) were included. In data retrieval, publications were filtered to exclude summaries, conference papers, editorials, books, book chapters, and meeting abstracts.

Data tools and analysis

Initially, 1016 records were retrieved. After applying the exclusion criteria, 819 articles were obtained. Subsequently, 317 records were manually deleted because they did not correspond to the target field of study. When verifying the lack of duplication, the number of works included in the analysis was 502.

After obtaining the total sample for the study, fields were normalized for keywords, authors, affiliation institutions, and countries, e.g., Brasil – Brazil.

Data was downloaded in BibTex and CSV formats. Microsoft Excel® was used to handle and manage the data.

The following scientometric and bibliometric tools were used: Bibliometrix/RStudio and VOSviewer.

The Bibliometrix/RStudio application made it possible to obtain and organize data regarding authors, institutions, countries, and research fields, through the keywords of the most prolific authors and journals, as well as from impact factors, h-index (which indicates that a given article (h) has been cited at least 10 or more times)[17], or total citations[18]. In addition, VOSviewer was included to identify the links or relationship strengths in collaborations between countries[19].

With the data collected, the bibliometric study included two analysis categories: characterization and mapping of the scientific literature[20].

The characterization analysis identified:

- Growth and evolution of annual and cumulative production: number of works and annual growth rate, growth tendency and its determination adjustment.
- Authors analysis: number of authors, articles per author, authors' affiliations, co-authorship index (average number of signing authors) and co-authorship rate (percentage of articles signed by two or more authors), and authors' productivity. According to Lotka, authors are distributed in three levels of productivity: small or occasional producers (only one publication), medium producers (between 2 and 9 works) and large producers (with 10 or more articles)[21].
- Analysis of the authors' countries of affiliation: number of signatory countries, number of citations, international collaboration rate (percentage of articles signed by two or more countries), collaboration index (average of signatory countries), and correlation between collaborations and number of citations.
- Journal analysis: number of journals, articles per journal, and Bradford Core to identify those most focused on the field of study[22].
- Analysis of scientific visibility by: citations, h-index, and percentile of the Web of Science impact factor for 2022 of the most relevant journals.

On the other hand, mapping of the scientific literature was used to visualize the relational aspects of: a) the co-authorships of the 20 authors with the largest production based on the Louvain algorithm; b) the international collaborations corresponding to the 50 countries with the highest production based on the Spinglass algorithm; and c) the scientific link between research fields through the analysis of the conceptual structure that determines the main topics and trends that have been studied and explained, based on the co-occurrence of keywords from the authors, which allows analyzing the structure of a field of study[23,24].

The link network of co-authorships and international collaborations is shown through nodes and link lines. The size of each node represents the number of works produced in collaboration, and the thickness of the link lines show the intensity of the relationship that connects them. Nodes that correspond to the same cluster or group are identified by color.

For the analysis of conceptual structure, word cloud graphics, thematic distribution maps and factor analysis were used.

To observe the central point of the topic under study, the word cloud was used. The large font size and positioning of the words closer to the center indicate the highest co-occurrence of keywords and the most studied areas of a field. Words in smaller fonts suggest possible lines of research[25].

The thematic distribution map made it possible to categorize the main research topics through two dimensions: relevance and development. Relevance (degree of centrality or interaction with other graphic groups) establishes the relative importance of each topic in the research field and indicates the degree of correlation between different topics as well as density. The greater the number of relationships that a node has with others in the thematic network, the greater its centrality and importance, and if it is located within an essential position in the network. Development (degree of density) measures the internal strength of a cluster network that establishes the advancement and range of knowledge generated by the cohesion between the nodes, and also outlines its capacity to develop and sustain itself. The interaction of both dimensions establishes four quadrants of analysis:

- Motor topics: they comprise the research front that encompasses the topics with most relevance and development, and that are crucial for structuring a research topic.
- Basic and transversal topics: they show the topics with high relevance and stable development. They are vital for transdisciplinary research.
- Emerging or declining topics: they include those with low relevance and development, but with the potential to become basic or motor topics.
- Niche or specialized topics: they contain topics with a high degree of research development, but their relevance in the research field is not yet that important.

Both motor and basic topics are considered those that favor the advancement and consolidation of a field of knowledge[26].

Joint word analysis by multiple correspondence factor analysis interprets the development of a field as a function of the relative position of the keyword point. Taken as a whole, the terms closest to each other represent a large part of the article, but if they are distant from each other, they indicate that a small proportion of the article deals with these keywords together, and those that are positioned on the edges are less related to other research topics[27]. Joint word analysis by multiple correspondence factor analysis shows the popularity of a topic depending on how close a keyword is to the central point.

Ethics statement

Considering that this was a bibliographic analysis of published articles, no ethical approval or informed consent were required for this study. Neither were any authors contacted to obtain more information about their publications.

Results

Production

The 502 analyzed articles accumulated 7609 citations. The average number of citations per year was 3.73, while the average number of citations per publication was 19.42.

The annual growth rate between 2013 and 2022 was 21%. The last three years reached the highest number of published articles, accumulating 50% of all the analyzed production (Table 1). Both the annual and cumulative growth rates were exponential, with a determination adjustment of 0.9 and 1.0, respectively.

Table 1. Production and average annual citation of publications regarding tuberculosis in vulnerable populations (2013-2022).

Year	N	NA	MeanTCperArt	MeanTCperYear
2013	17	17	26.12	2.61
2014	23	40	31.17	3.46
2015	35	75	19.69	2.46
2016	28	103	22.79	3.26
2017	41	144	45.61	7.60
2018	60	204	11.13	2.23
2019	48	252	14.92	3.73
2020	80	332	11.53	3.84
2021	75	407	6.19	3.09
2022	95	502	5.06	5.06

N: number of articles published in the period studied; NA: number of accumulated articles; MeanTCperArt: average number of citations per article; MeanTCperYear: average number of citations per year.

Authorship, co-authorship, and institutions of affiliation

The analyzed production was signed by 2829 authors, of which 88% were occasional (with a single publication), 12% were medium producers (between 2 and 9 articles) and 0.2% were large producers (with 10 or more papers).

The large producers and more citations were RA Arcêncio (17; 90), ACV Ramos (14; 79) and TZ Berra (13; 253), all three from University of São Paulo (10; 204) (Figure 1).

Most of the authors with the largest production began publishing in 2017 (Figure 1). The author with the most cited publications in 2022 was GB Migliori from the Instituti Clinici Scientifici Maugeri IRCCS of Italy (80 citations), while the author with the most up-to-date publications was RA Arcêncio from the Universidade de São Paulo in Brazil, who published six articles in 2022.

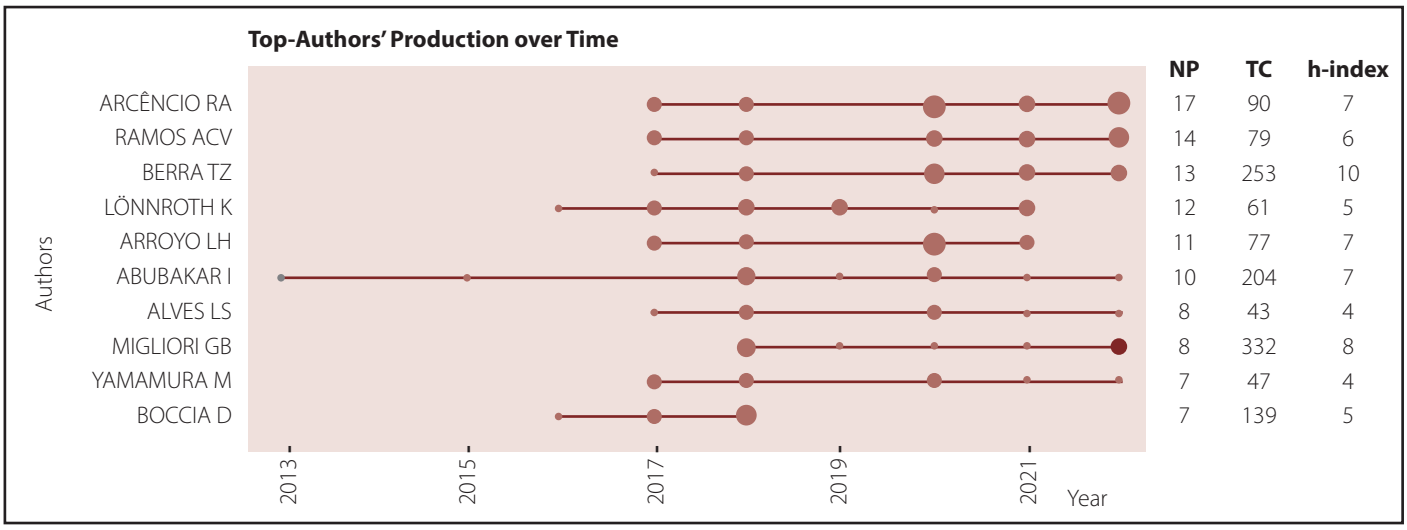
The co-authorship rate was 95% and the co-authorship index was 5.9. The average number of works per author was 0.17. Figure 2 shows how the authors who are identified as the most prolific tend to collaborate with each other.

The most prolific institutions (≥10 publications) accumulated 61% of the analyzed production. The Universidade de São Paulo of Brazil and the University College London of the United Kingdom were the institutions with the higher number of articles regarding the topic of study (18%), followed by the London School of Hygiene and Tropical Medicine and the Organisation Mondiale de la Santé (9%) (Table 2).

Geographic coverage and international collaboration

The geographical coverage of the publications corresponded to 101 countries. The main article generator was the United States, with 27% of the analyzed production and 3437 citations, along with the United Kingdom (19%; 2333). Next, countries belonging to the BRICS bloc are positioned: Brazil (15%; 980), India (10%; 1004), and South Africa (9%; 1004) (Table 3).

Figure 1. The 10 most productive authors over the period 2013-2022. Adapted from the figure obtained with Bibliometrix.



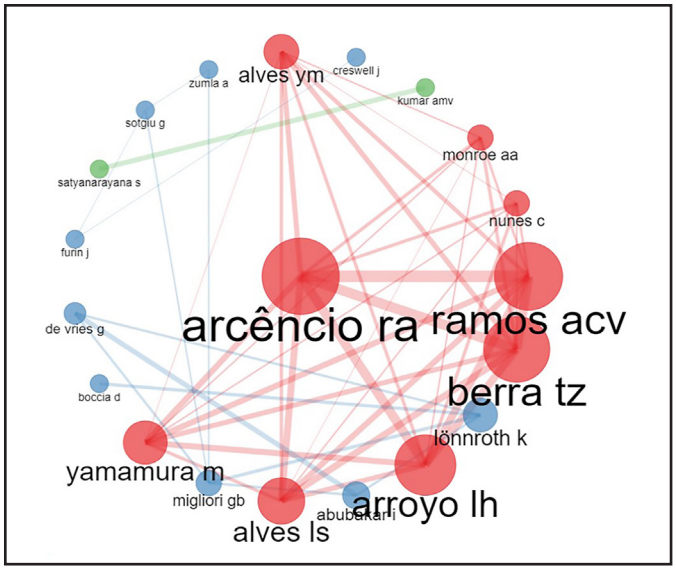
NP: number of articles published in the period studied; TC: Total Cited.

Table 2. Institutions of affiliation of authors with 10 or more publications on tuberculosis in vulnerable populations during the period 2013-2022.

Affiliations	Country	Articles
Universidade de São Paulo	Brazil	29
University College London	United Kingdom	29
London School of Hygiene and Tropical Medicine	United Kingdom	26
Organisation Mondiale de la Santé	Switzerland	26
Fundação Oswaldo Cruz	Brazil	19
Karolinska Institutet	Swede	19
Universidade Nova de Lisboa	Portugal	17
Harvard Medical School	USA	17
University of Cape Town	South Africa	17
Stellenbosch University	South Africa	15
Imperial College London	United Kingdom	13
University College London Hospitals NHS Foundation Trust	United Kingdom	13
The University of Sydney	Australia	12
KNCV Tuberculosis Foundation	Netherlands	12
Centers for Disease Control and Prevention	USA	12
South African Medical Research Council	South Africa	10
International Union Against Tuberculosis and Lung Disease	France	10
University of KwaZulu-Natal	South Africa	10

The international collaboration rate was 39%. The United States and Brazil showed a higher percentage of collaborative articles, both nationally and internationally (Table 2). Even though the United States, the United Kingdom and Brazil have a greater

Figure 2. Co-authorship network of the 20 authors with the highest production during the period 2013-2022 created with Bibliometrix.



number of collaborative works, those holding the best positions in the collaboration network according to the total link strength were the United Kingdom (261), the United States (198), France (123), and Switzerland (121) (Table 3). The collaboration network can be seen in Figure 3.

A Pearson correlation of 0.92 was obtained between the number of citations and the total link strength; and of 0.87 between the number of collaborative works and the number of citations.

Analysis of funding agencies

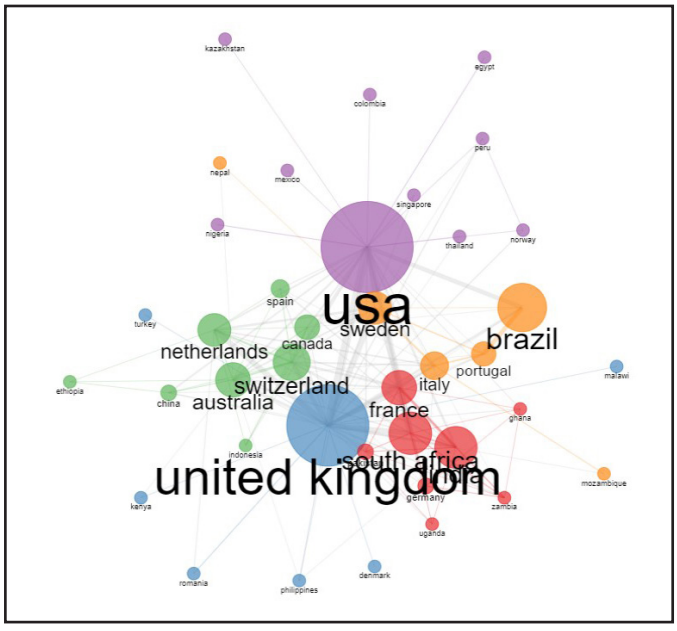
The funding agencies that have boosted research regarding TB in vulnerable populations, with more than 10 published works and concentrating 24% of all production, corresponded mainly

Table 3. Production and pattern of collaboration of countries with more than 10 articles during the period 2013-2022.

Country	Number of publications	Total citations	Total link strength	Collaborative articles	SCP	MCP
United States	136	3437	198	73	47	26
United kingdom	94	2333	261	39	12	27
Brazil	75	980	61	60	36	24
India	52	1004	77	34	28	6
South Africa	46	1155	104	21	9	12
Switzerland	35	1080	121	13	4	9
Australia	32	818	85	13	6	7
Canada	31	846	63	13	11	2
Sweden	27	458	90	7	0	7
Netherlands	27	658	94	6	0	6
France	26	1127	123	7	1	6
Italy	25	771	68	10	4	6
Portugal	25	350	61	6	3	3
Spain	20	180	47	9	7	2
Germany	16	214	44	9	5	4
China	16	383	32	8	6	2
Norway	11	483	31	4	1	3
Ethiopia	11	131	16	5	3	2
Peru	11	211	17	6	4	2

SCP: Intra-country publication; MCP: Inter-country publications.

Figure 3. Communities within international collaboration networks of the 25 countries with the highest production during the period 2013-2022.



to the United States and Brazil. The National Institutes of Health, which is the main United States federal agency to conduct and support medical research, is ranked first (5%). In second place (4%) is the Medical Research Council of the United Kingdom and in third place (3%) is the Conselho Nacional de Desenvolvimento Científico e Tecnológico, an organization linked to the Ministério da Ciência, Tecnologia e Inovação of Brazil (Table 4).

Documentary typology and keywords

According to the documentary typology, 78% of the production corresponds to original research, and 23% belongs to reviews.

Regarding the 502 analyzed articles, a total of 1040 keywords were noted. Keywords are an important measurement tool to understand the main content of research articles; for this purpose, different mappings were used to illustrate the thematic status of the works.

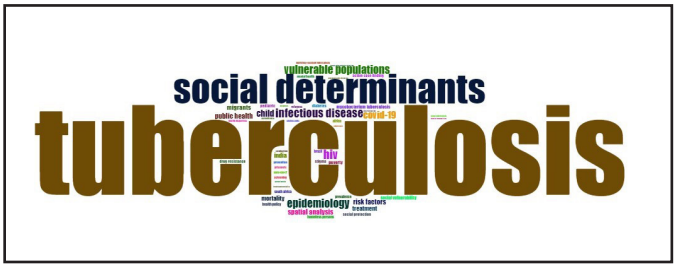
The keywords cloud showed a tendency to publish about areas linked to "tuberculosis", "social determinants", "infectious disease", "epidemiology", "children", "COVID-19" and "HIV" (Figure 4).

An analysis of the conceptual structure of the thematic map identified that well-developed and relevant motor topics for the construction of this scientific field are risk factors associated with drug resistance and socioeconomic characteristics (Figure 5). A second research field relates TB to migrant populations and, as a stigma, to mental health issues. Peripheral themes are developed within but have a marginal role; the vulnerable population of India was involved in the development of this scientific field. The basic and transversal topics, important for a stable development of the scientific field, but that still require more research are: pathologies such as HIV and COVID-19 related to TB, social determinants, poverty, and drug resistance. However, the least

Table 4. Funding agencies with 10 or more funded research in tuberculosis and vulnerable populations during the period 2013-2022.

Funding agencies	Country	Number of publications
National Institutes of Health	United States	26
Medical Research Council	United Kingdom	19
National Institute of Allergy and Infectious Diseases	United States	15
Conselho Nacional de Desenvolvimento Científico e Tecnológico	Brazil	14
Bill and Melinda Gates Foundation	United States	13
Coordenação de Aperfeiçoamento de Pessoal de Nível Superior	Brazil	12
Fogarty International Center	United States	11
Fundação de Amparo à Pesquisa do Estado de São Paulo	Brazil	10

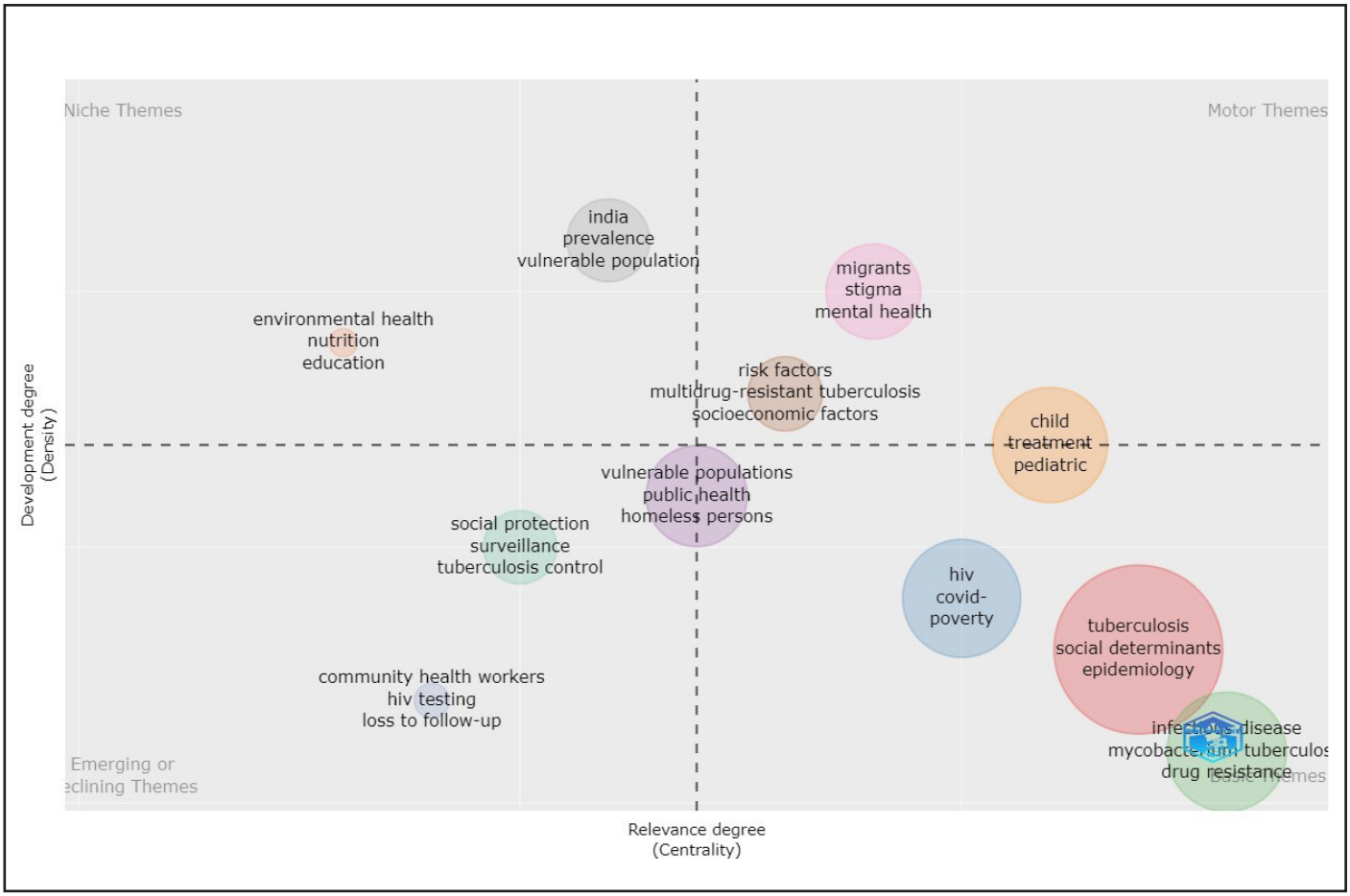
Figure 4. Keyword cloud obtained with Bibliometrix.



advanced topics were those related to the tuberculosis control and social protection. The treatment of TB in children appeared at the crossroads between motor, transversal, and basic topics.

Joint word analysis by multiple correspondence factor analysis identified four groups as areas of scientific influence (Figure 6). The first important group (the largest highlight), includes most of the relevant concepts in the field of study and is very consistent with topics related to the prevalence, in this case, of TB. The second group (the medium remarking), covered works on migration, with a consistency on malaria and TB in migrant populations.The third group (the smallest remarking), corres-

Figure 5. Conceptual structure of the thematic map with Bibliometrix.



ponds to the vulnerability of the population with TB, and, finally, the purple group corresponds to prevention and drug resistance.

Analysis of journals and language coverage

Identifying the top journals can help researchers to choose the most appropriate one to consult or to publish their work.

Articles on the field of study were published in 250 international journals. The dispersion obtained with Bibliometrix shows a nucleus with 14 journals and 170 articles that accumulated 2443 citations (Table 5). Next, was the finding of zone 1 with 72 journals and 168 works, and zone 2 with 164 journals and 164 articles. Most journals in the nucleus had a JIF percentile higher than 75%, which

Figure 6. Conceptual map by multiple correspondence factor analysis obtained with Bibliometrix.

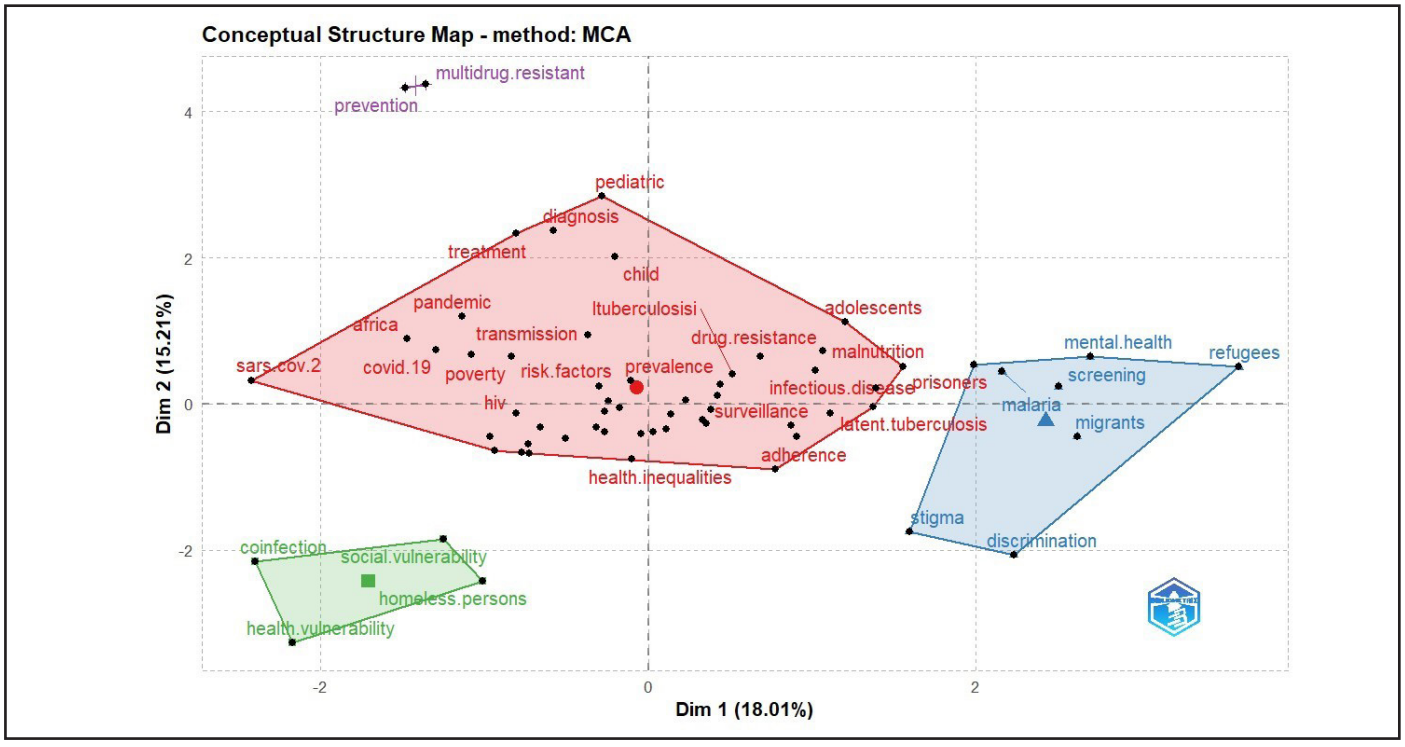


Table 5. Core journals according to Bradford dispersion.

Journal	NP	TC	h_index	JIF percentile (2022)
International Journal of Tuberculosis and Lung Disease	31	462	14	58.3
Plos One	23	339	10	78.73
BMJ Open	16	205	9	75.99
Indian Journal of Tuberculosis	15	35	2	-
International Journal of Environmental Research and Public Health	13	82	6	64.34
International Journal of Infectious Diseases	12	499	9	88.76
BMC Infectious Diseases	9	123	7	58.53
European Respiratory Journal	9	354	8	97.7
Pan American Journal of Public Health	8	76	5	45.8
Tropical Medicine and International Health	8	76	5	85.40
BMC Public Health	7	105	6	80.63
Pan African Medical Journal	7	60	4	11.88
Tropical Medicine and Infectious Disease	6	26	3	68.90
Journal of Infection in Developing Countries	6	5	1	25.19

NP: Number of publications during the period 2013-2022; JIF: Journal impact factor; TC: Total citations.

is why they were in the first quartile, giving them a great visibility. However, *The Lancet* – the journal with the highest number of citations (1126) – is not among the top journals or considered in the nucleus, having only two publications on the topic of study.

The main journal was the *International Journal of Tuberculosis and Lung Disease* due to its highest percentage of published articles, citations and h-index, although it is located in the second quartile. The second relevant journal was *Plos One*.

Regarding language, 95% of all the articles were published in English. The rest were published in Spanish (2%), Portuguese (1%), German (0.8%), as well as Chinese, French, Russian, and Swedish (0.2%).

Discussion

This study presents an overview of research trends regarding TB in vulnerable populations over the last decade, in order to provide information for future works based on scientific evidence.

The exponential growth of scientific production on TB in vulnerable populations, especially in recent years[10,12], reveals the interest of researchers in ending the TB epidemic. This trend in the scientific literature suggests an important increase of the research regarding this problem in the medium term[21]. This dynamic concurs with following the WHO's goal established for 2035, which is to reduce 95% of deaths and 90% of the incidence of TB compared to 2015[14].

However, to meet these goals, global funding needs to be increased. The United Nations only funded half of the planned budget for 2018-2022[1]. The three main sources of funding for the analyzed research came from the United States, the United Kingdom, and Brazil.

Regarding authors' productivity, it is confirmed that Lotka's law is fulfilled, which states that the bulk of articles published on a specific field matches with a very small number of authors who are specialized in that area of knowledge[21]. The result of more than 60% being occasional authors – that is, with a single publication – shows a lack of consolidation of the scientific literature on TB in vulnerable populations. Therefore, work must be done so that researchers continue along this line and, in this way, the literature regarding the analyzed topic can be consolidated, so that there is a higher percentage of specialized authors in the field of study[28].

Half of the most prolific authors (with 10 or more articles) are affiliated to the Universidade de São Paulo in Brazil, which is the institution with the largest number of published research works on TB in vulnerable populations.

Even though 95% of TB cases occur in developing countries[29], research is mainly led by non-burdened countries[12]. The United States and the United Kingdom, followed by India –a country with a high incidence of TB–[1] are among the leading countries in research on TB in vulnerable populations, as well as in the rest of the bibliometric studies on TB[14]. However, there is a difference in our results because Brazil ranks third in terms of vulnerable populations, while, according to other bibliometric studies, it ranks seventh in scientific production regarding TB[12].

Countries with a low TB burden, but a higher Gross Domestic Product (GDP) for scientific research, carry out policies for cross-border contributions that favor international collaboration[30]. Even though the United States is a country with a low TB burden, it is affected by migratory flows coming from the border with Mexico, which favors the transmission of infectious diseases such as TB[31]. This would favor the interest of the United States in researching the subject.

Of the 10 countries with the highest production, only three are regions with a high TB burden: India, Brazil, and South Africa. These, in turn, were the countries with five or more international collaborative investigations. In 2015, national TB research plans were established in several countries with high and medium TB burdens, such as Brazil, Russia, India, China, and South Africa, which correspond to the so-called BRICS bloc. These countries concentrate almost 50% of the global TB burden[32], and have become academic leaders of the BRICS National Tuberculosis Program, participating in the BRICS TB Research Network. However, the role of China and Russia in research on TB in vulnerable populations is more discrete, with them not being among the top 20 countries.

Brazil takes part in different structured global networks on TB, such as the Brazilian Rede TB or Global TB Network GTN[33], a fact that favors its potential to work in national and international collaborations[34].

According to our results, on the one hand, developing countries most affected by TB need to take responsibility for leading research in order to control the epidemic[10] and, on the other hand, they need to establish international collaborations[12].

Overall, scientists will have to make efforts in order to significantly increase international collaborations, since they are less than 50%. Therefore, not only will resources and knowledge be shared, but the studies will have greater visibility, as shown by the results, according to the high correlation between international collaborations and the number of citations[35].

Poverty, HIV, and drug resistance are topics associated to the global TB epidemic[36]. Just as the results of the bibliometric study by Garrido-Cardenas *et al.*[14] show, it is necessary to continue researching on HIV and drug-resistant TB.

HIV infections continue to be of great interest to scientists working with TB, since this is one of the main causes of death in these populations[37].

The resistance of TB to the drugs used for its treatment has become the center of attention of different research efforts[38], especially in vulnerable populations, due to the cost, the duration of its treatment, and the fact that MDR/RR-TB cases have a lower success rate of the anti-tuberculosis treatment than cases sensitive to said treatment (88% versus 63%, respectively)[1].

Historically, poverty has been one of the factors that most influences high TB rates[7]. Therefore, a thorough analysis of these populations should continue to be a priority topic for research,

especially considering the complexity of the SDH related to TB[8], where there are more cases, but also fewer resources to fight it.

As expected, practically the whole production regarding this topic is published in English, not only because it is the language of scientific communication, but also because more than half of the research comes from Anglo-Saxon countries. Therefore, it is necessary to boost publications in other languages so that their speakers have greater access to the information generated regarding the topic of study. In addition, the Web of Science mostly indexes journals in English.

In this sense, the reference journals for researchers, as shown in different bibliometric studies, are the *International Journal of Tuberculosis* and *Lung Disease and Plos One*[39].

Finally, regarding the possibility that the Covid-19 could have affected the results in the final stage, it can be noted that although in 2019 there was a decrease in the production of articles, in the period of greatest crisis of this pandemic, the number of papers published on TB and vulnerable populations increased again, probably because it is also an infectious respiratory disease. As for the average number of citations for this period, the decrease observed in 2021 and 2022 is normal because such a decrease always occurs in the last years because there has not been time to cite. In this regard, it would be very useful to analyze trends in the coming years.

Limitations

The following are the main limitations of the present study: 1. Scopus databases, just as other databases such as Web of Science, are not perfectly adapted to bibliometric analyses, since they usually return a certain amount of erroneous data which limit the conclusions that can be drawn from them[40]; 2. using a single database and an established time period does not show all scientific coverage on TB in vulnerable populations; 3. the scope of the review may be limited to the Scopus coverage, which is known to overrepresent journals in English and underrepresent those from the Global South – a consideration that implies specific differences for researchers based in low-income countries, where access to funding for research and for article processing charges can be a limitation[41]; 4. the number of citations may not reflect the quality of the research[10]; 5. the results from the different topics analyses are conditioned by the parameters used for their recovery; and, 6. the use of keywords may not have retrieved records focused on race, immigrant populations or socioeconomic levels, among others.

Conclusions

This bibliometric study describes the most notable trends in global research on TB in vulnerable populations from 2013 to 2022. There has been an increasing trend in scientific production over the last decade, especially in the last three years. High-income, low-burdened countries such as the United States and the United Kingdom have led global production, although countries with a high TB burden such as Brazil, India and South Africa were also among the top five countries for production of TB related works. Half of the most prolific authors were from Brazil. More frequent and deeper cooperation between countries, authors, and institutions, is essential, mainly among those with the highest TB burden and low economic development. Finally, we consider that it is necessary to carry out more international and regional bibliometric studies on TB in vulnerable populations with a more in-depth exploration, an also associate alternative metrics to analyze the social impact of scientific results.

Statements and Declarations

No funds, grants, or other support was received. This study was financed with the authors' own resources. The authors have no financial or proprietary interests in any material discussed in this article.

Authors's contribution

Conceptualization: Cristina Torres-Pascual (CTP), Cristina Gordillo-Marroquín (CGM), Héctor Javier Sánchez-Pérez (HJSP); Literature search: CTP; Data analysis: CTP, HJSP, CGM; Writing – original draft: CTP, CGM, HJSP; Writing – review & editing: HJSP, CGM, CTP. All authors read and approved the final manuscript.

Bibliography

1. WHO. Global tuberculosis report 2023. *Geneva: World Health Organization*. 2023. p. 1-75.
2. WHO. *Tuberculosis: Key facts* [Internet]. 2023 [cited 2023 Jun 25]. Available from: https://www.who.int/news-room/fact-sheets/detail/tuberculosis
3. WHO. WHO consolidated guidelines on tuberculosis. Module 2: Screening Systematic screening for tuberculosis disease. *WHO. Geneva;* 2021. p. 1-68.
4. Miggiano R, Rizzi M, Ferraris DM. *Mycobacterium tuberculosis* pathogenesis, infection prevention and treatment. *Pathogens.* 2020; 9(5):385.
5. Greenall M, Kunii O, Thomson K, Bangert R, Nathan O. Reaching vulnerable populations: Lessons from the Global Fund to fight AIDS, tuberculosis and Malaria. *Bull World Health Organ*. 2017;95(2):159-61.
6. Sánchez-Pérez HJ, Gordillo-Marroquín C, Vázquez-Marcelín J, Martín-Mateo M, Gómez-Velasco A. Sociodemographic factors associated with the success or failure of anti-tuberculosis treatment in the Chiapas Highlands, Mexico, 2019-2022. *PLoS One*. 2024;19(1): e0296924.
7. Paluzzi JE. A social disease/a social response: Lessons in tuberculosis from early 20th century Chile. *Soc Sci Med.* 2004;59(4):763-73.
8. Wu S, Litvinjenko S, Magwood O, Wei X. Defining tuberculosis vulnerability based on an adapted social determinants of health framework: a narrative review. *Glob Public Health.* 2023;18(1):2221729..

9. WHO. *The End TB Strategy* [Internet]. 2015 [cited 2023 Aug 19]. Available from: https://iris.who.int/bitstream/handle/10665/331326/WHO-HTM-TB-2015.19-eng.pdf?sequence=1

10. Ekenoğlu MY, Etiz P. A Scopus-Based Bibliometric Analysis of Global Tuberculosis Publications: 1849-2020. *Turk Thorac J*. 2022;23(3):246-56.

11. Igwaran A, Edoamodu CE. Bibliometric analysis on tuberculosis and tuberculosis-related research trends in Africa: A decade-long study. *Antibiotics (Basel)*. 2021;10(4):423.

12. Abdullah M, Humayun A, Imran M, Bashir M, Malik A. A bibliometric analysis of global research performance on tuberculosis (2011-2020): Time for a global approach to support high-burden countries. *J Family Community Med*. 2022;29(2):117-24.

13. Castor K, Mota FB, da Silva RM, Cabral BP, Maciel EL, de Almeida IN, *et al.* Mapping the tuberculosis scientific landscape among BRICS countries: a bibliometric and network analysis. *Mem Inst Oswaldo Cruz* [Internet]. 2020;115. Available from: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0074-02762020000100308&tlng=en

14. Garrido-Cardenas JA, de Lamo-Sevilla C, Cabezas-Fernández MT, Manzano-Agugliaro F, Martínez-Lirola M. Global tuberculosis research and its future prospects. *Tuberculosis (Edinb)*. 2020; 121(October 2019):121:101917.

15. To WM. A Bibliometric Analysis of World Issues-Social, Political, Economic, and Environmental Dimensions. *World*. 2022;3(3):619-38.

16. Baas J, Schotten M, Plume A, Côté G, Karimi R. Scopus as a curated, high-quality bibliometric data source for academic research in quantitative science studies. *Quant Sci Stud.* 2020;1(1):377-86.

17. Gálvez-Toro A, Amezcua M. El factor h de Hirsch: the h-index: Una actualización sobre los métodos de evaluación de los autores y sus aportaciones en publicaciones científicas. *Index de Enfermería*. 2006;15(55):38-43.

18. Aria M, Cuccurullo C. bibliometrix: An R-tool for comprehensive science mapping analysis. *J Informetr*. 2017;11(4):959-75.

19. van Eck NJ, Waltman L. Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics*. 2010;84(2):523-38.

20. Donthu N, Kumar S, Mukherjee D, Pandey N, Lim WM. How to conduct a bibliometric analysis: An overview and guidelines. *J Bus Res*. 2021;133:285-96.

21. López PJ, Terrada M. Los indicadores bibliométricos y la evaluación de la actividad médico-científica (III). Los indicadores de producción, circulación y dispersión, consumo de la información y repercusión. *Med Clin (Barc)*. 1992;98:142-8.

22. Sembay M, Luiz Pinto A, De Macedo DDJ, Moreiro-González JA. Aplicación de la Ley de Bradford a la investigación sobre Open Government *An. Documentación* [Internet]. 1 de enero de 2020 [cited 12 de junio de 2023];23(1). Available from: https://revistas.um.es/analesdoc/article/view/326771.

23. Chen X, Chen J, Wu D, Xie Y, Li J. Mapping the Research Trends by Co-word Analysis Based on Keywords from Funded Project. *Procedia Comput Sci*. 2016;91:547-55.

24. de la Hoz-Correa A, Muñoz-Leiva F, Bakucz M. Past themes and future trends in medical tourism research: A co-word analysis. *Tourism Management*. 2018;65:200-11.

25. Mulay P, Joshi R, Chaudhari A. Distributed incremental clustering algorithms: A bibliometric and word-cloud review analysis. *Sci Technol Libr.* 2020;39(3):289-306.

26. Esfahani HJ, Tavasoli K, Jabbarzadeh A. Big data and social media: A scientometrics analysis. *Int J Data Netw Sci*. 2019;3(3):145-64.

27. Sharma P, Singh R, Tamang M, Singh AK, Singh AK. Journal of teaching in travel & tourism: a bibliometric analysis. *J Teach Trave Tour* 2021;21(2):155-76.

28. Ramos JM, Masiá M, Padilla S, García-Pachón E, Gutiérrez F. Producción española sobre tuberculosis a través del MEDLINE (1997–2006). *Arch Bronconeumol*. 2009;45(6):271-8.

29. Atun R, Raviglione M, Marais B, Zumla A. Tuberculosis control is crucial to achieve the MDGs. *Lancet*. 2010;376(9745):940-1.

30. Dara M, Sulis G, Centis R, D'Ambrosio L, De Vries G, Douglas P, *et al.* Cross-border collaboration for improved tuberculosis prevention and care: policies, tools and experiences. *Int J Tuberc Lung Dis*. 2017;21(7):727-36.

31. Bidegain EA. Movilidad y condiciones de vida de pacientes en tratamiento de tuberculosis pulmonar. *Alteridades*. 2022;32(63):119-33.

32. Raviglione M, Uplekar M, Weil D, Kasaeva T. Tuberculosis makes it onto the international political agenda for health…finally. *Lancet Glob Heal*. 2018;6(1):e20-1.

33. Silva DR, Rendon A, Alffenaar JW, Chakaya JM, Sotgiu G, Esposito S, *et al.* Global TB network: Working together to eliminate tuberculosis. *J Bras Pneumol*. 2018;44(5):347-9.

34. Sotgiu G, Sulis G, Matteelli A. Tuberculosis-a World Health Organization Perspective. *Microbiology Spectrum* [Internet]. 2017 [cited 2023 Oct 18];5(1). Available from: https://doi.org/10.1128/microbiolspec.TNMI7-0036-2016

35. Ni P, An X. Relationship between international collaboration papers and their citations from an economic perspective. *Scientometrics*. 2018;116(2):863-77.

36. Gupta-Wright A, Corbett EL, van Oosterhout JJ, Wilson D, Grint D, Alufandika-Moyo M, *et al.* Rapid urine-based screening for tuberculosis in HIV-positive patients admitted to hospital in Africa (STAMP): a pragmatic, multicentre, parallel-group, double-blind, randomised controlled trial. *Lancet*. 2018;392(10144):292-301.

37. Gandhi NR, Moll A, Sturm AW, Pawinski R, Govender T, Lalloo U, *et al.* Extensively drug-resistant tuberculosis as a cause of death in patients co-infected with tuberculosis and HIV in a rural area of South Africa. *Lancet*. 2006;368(9547):1575-80.

38. Ahmad N, Ahuja SD, Akkerman OW, Alffenaar JWC, Anderson LF, Baghaei P, *et al.* Treatment correlates of successful outcomes in pulmonary multidrug-resistant tuberculosis: an individual patient data meta-analysis. *Lancet*. 2018;392(10150):821-34.

39. Islam MA, Kundu S, Hanis TM, Hajissa K, Musa KI. A Global Bibliometric Analysis on Antibiotic-Resistant Active Pulmonary Tuberculosis over the Last 25 Years (1996-2020). *Antibiotics (Basel)*. 2022;11(8):1012.

40. Gaur A, Kumar M. A systematic approach to conducting review studies: An assessment of content analysis in 25 years of IB research. *J World Bus*. 2018;53(2):280-9.

41. Borrego Á. Article processing charges for open access journal publishing: A review. *Learned Publishing*. 2023;36(3):359-78.