

Artificial intelligence and tobacco use: A bibliometric analysis 1997–2026

Shengbo Wang¹, Jiahao Zhang², Liushen Chen²

ABSTRACT

INTRODUCTION Globally, tobacco use is a key public health issue. Exploring the impact of developing AI technologies on tobacco use and intervention will help promote the digital transformation of public health research. This study aims to map the trends, networks, and core themes of research at the intersection of artificial intelligence and tobacco use through visual analytics.

METHODS In this study, a bibliometric analysis was conducted on 335 articles on AI and tobacco use in the Web of Science core collection from 1997 to 2026. The search cutoff date was 8 April 2026. Descriptive and visual analysis were conducted on the publication trend, cooperation network, keyword co-occurrence, clustering, and burst of this literature.

RESULTS The speed of publication of literature in this field has increased significantly since 2018. In the cooperation network, national cooperation is dominated by the United States. Universities, rather than transnational institutions, significantly promote cooperation. The authors' cooperation network is more dispersed. The characteristics of interdisciplinary cooperation are obvious. The research hotspots focus on several aspects. These include machine learning prediction of smoking, dialogue AI and robot interventions, and NLP and social media monitoring. Other topics encompass the analysis of smoking-cessation behavior among youth groups, combined with AI, deep learning, and behavioral analysis; and the evaluation of tobacco characteristics, tobacco products, media, and health. Keyword hotspots and bursts reveal that the field has recently paid special attention to the intervention of cutting-edge technologies in smoking and smoking cessation behavior.

CONCLUSIONS Through data mining and visualization technology, this study reveals the overall evolution, interaction mode, and key fields of AI and tobacco use knowledge. The results provide an important framework for further research.

AFFILIATION

¹ Department of Sociology, The Chinese University of Hong Kong, Hong Kong, China

² Southwestern University of Finance and Economics, Research Institute of Social Development, Chengdu, China

CORRESPONDENCE TO

Jiahao Zhang, Southwestern University of Finance and Economics, No. 555 Liutai Avenue, Wenjiang District, Chengdu City, Sichuan Province, China

E-mail: 2240303z2001@smail.swufe.edu.cn

[smail.swufe.edu.cn](mailto:2240303z2001@smail.swufe.edu.cn)

ORCID iD: <https://orcid.org/0009-0003-8053-1046>

[org/0009-0003-8053-1046](https://orcid.org/0009-0003-8053-1046)

KEYWORDS

artificial intelligence, tobacco, bibliometric analysis

Received: 21 April 2026

Revised: 27 July 2026

Accepted: 3 August 2026

Tob. Induc. Dis. 2026;24(August):151

<https://doi.org/10.18332/tid/226603>

INTRODUCTION

Despite the declining trend in global tobacco use, the tobacco epidemic remains one of the most serious challenges in global public health in the 21st century¹. The World Health Organization (WHO) estimates that tobacco use causes more than 8 million deaths worldwide each year and causes more than trillions of dollars in economic losses, including direct healthcare expenditures². At the same time, the use of new tobacco products such as e-cigarettes is accelerating, and the use among youth groups is particularly prominent. There are tens of millions of young people aged 13–15 years around the world who are using e-cigarettes, and their average use rate is significantly higher than that of

adults³. In the face of the changing tobacco epidemic situation and the increasing burden of disease, how to improve the accessibility, accuracy, and scalability of tobacco control interventions through technological innovation has become an important issue in the field of global tobacco control⁴. In 2025, the WHO Regional Committee for the Western Pacific (WPRO), in its policy agenda, explicitly listed ‘the application of artificial intelligence in healthcare’ and ‘strengthening tobacco control’ as the core issues for the future of regional health, and advocated the use of artificial intelligence and other technical means to promote equitable and technology-driven health systems⁵.

In recent years, the rapid evolution of artificial intelligence technology has provided a new analytical framework and technical path for analyzing the health damage effect of tobacco use behavior and optimizing tobacco control strategies⁶. Technical branches such as machine learning, natural language processing, and generative artificial intelligence have gradually penetrated multiple key links of tobacco control interventions, covering behavior prediction, content monitoring, and personalized service delivery⁷. From the perspective of the public health demand side, there is a significant gap between the huge global tobacco use population and the supply capacity of smoking cessation services. AI-driven intervention tools represented by chatbots, with their low cost, high scalability and all-weather accessibility, have shown great potential in bridging this gap between supply and demand⁸. From the perspective of technological development, the integration of AI and tobacco control also reflects the development trend of digital public health^{9–11}.

Although the application of AI in tobacco control continues to expand, there are still key issues that need to be systematically addressed. Specifically, whether the current research hotspots match the public health needs, what is the evolution path between different research topics, and which directions are becoming emerging frontiers. Most of the existing reviews summarized the application and effect of AI in tobacco control from the content level^{6,7,12}, but the systematic analysis of this interdisciplinary field from a bibliometric perspective remains limited. The bibliometric method can reveal the research structure and development trend through visualization, and

has been widely used in the research of tobacco use and tobacco control¹³. However, the bibliometric analysis of this interdisciplinary field is still scarce. Therefore, applying this method to research on AI and tobacco use helps broaden the scholarly horizon and deepen our understanding of cutting-edge studies at the intersection of tobacco science and intelligent technologies.

Based on this, the study uses the bibliometric method and the Web of Science Core Collection database to systematically analyze the knowledge map of the research field on artificial intelligence and tobacco use. Specifically, this study aims to: 1) characterize the publication trends, core authors and institutions, major journals and high-impact literature in this field; 2) reveal research hotspots and knowledge structure through keyword co-occurrence and cluster analysis; and 3) combine burst detection and topic evolution analysis to identify potential frontier directions. The results will provide a systematic body of knowledge for researchers and policymakers in the field of tobacco control.

METHODS

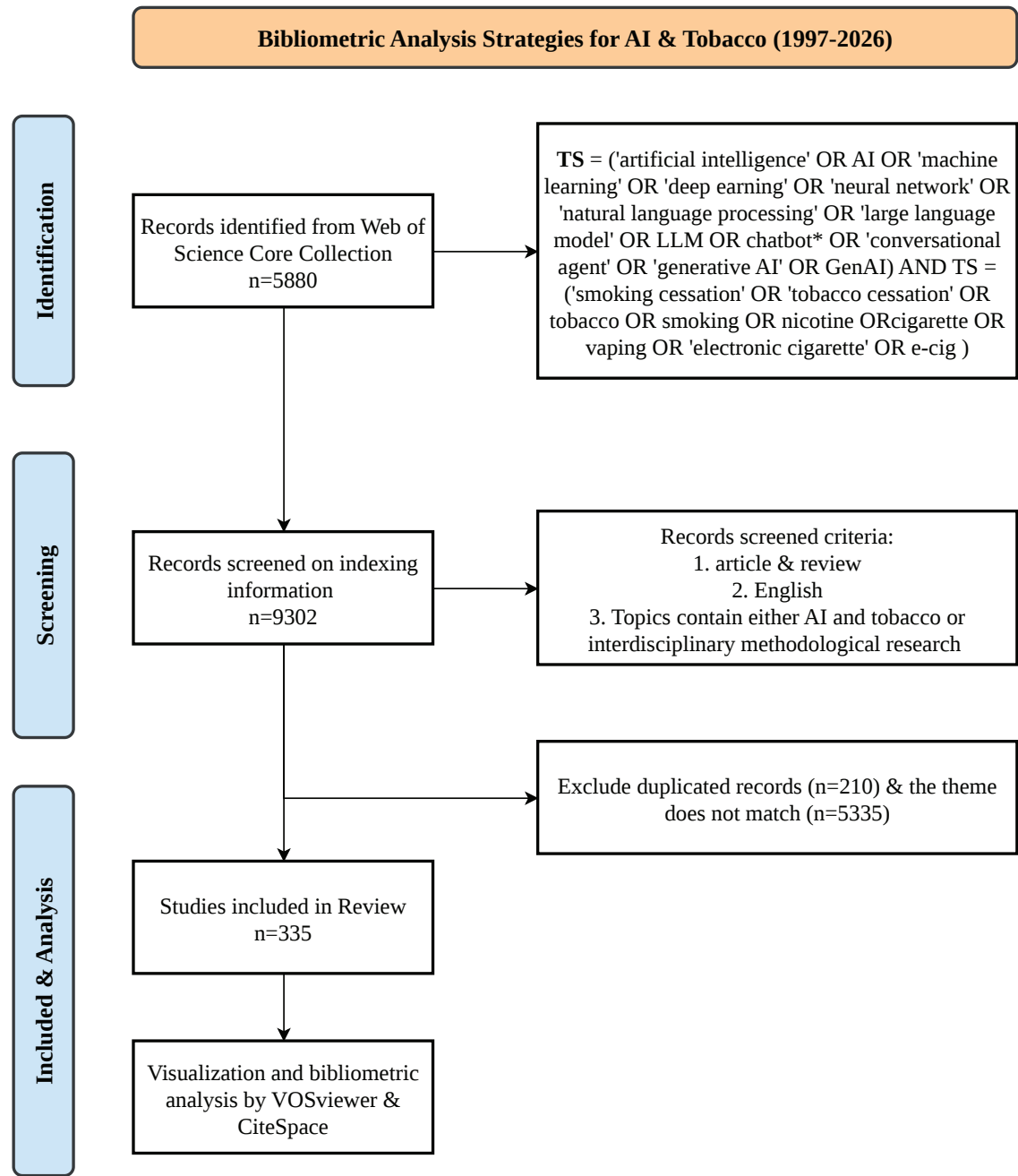
Data source and search strategy

This study uses the Web of Science core collection as the data source. The retrieval strategy employs topic field retrieval, and the retrieval formula is constructed using the Boolean logic combination of the artificial intelligence term group and tobacco use term group. The topic retrieval formula is shown in Figure 1. TS represents the topic field, which covers the title, abstract, and author keywords. The types of literature are limited to articles and reviews, the language is limited to English, and the search date is 8 April 2026. The initial literature totaled 5880. The screening process was independently conducted by two authors, followed by cross-checking and discussion. A third author resolved discrepancies. Included studies had to meet the following criteria: 1) English-language research articles or reviews; 2) Studies explicitly integrating AI as content or methodology within tobacco-related topics; and 3) Substantive methodological research involving interdisciplinary approaches. Exclusion criteria comprised: 1) Non-English publications or non-article/review formats; 2) Studies mentioning only

AI or tobacco use in isolation; 3) Studies using AI solely as a generic statistical tool or keyword without domain-specific application; and 4) Duplicate records. After removing 5545 irrelevant or duplicate records, the title information is exported in full-record and citation formats. Manual screening was based on the title, abstract, and keywords of the

literature to determine whether it explicitly involved the application of artificial intelligence technology to tobacco use-related behavior or policy research. It excluded studies were those that only mentioned tobacco use or used AI only as a general statistical tool. Finally, a total of 335 articles were included in the analysis. The flowchart is shown in Figure 1.

Figure 1. Flowchart of studies selection (1997–2026)



TS represents the topic field, covering title, abstract, and author keywords.

Data analysis methods

This study integrated bibliometric and knowledge map visualization methods^{14,15}. First, descriptive statistical tools were used to analyze publication trends, core author productivity, and journal distribution. Second, VOSviewer 1.6.2 was used to generate author collaboration and keyword co-occurrence networks and to perform keyword cluster analysis. Third, CiteSpace 6.3.3 was used for burst term detection (threshold set at 0.3) and thematic evolution path analysis, identifying research frontiers and their dynamic evolution characteristics. In terms of metric calculation methods, Betweenness Centrality (BC) was computed using shortest-path algorithms to identify pivotal nodes that bridge different research domains. Network modularity (Q) was calculated using the Louvain algorithm to assess the clarity of thematic clusters.

RESULTS

Publication trends and journal distribution

Figure 2 shows the temporal distribution of annual and cumulative publication counts in the field. The development of this interdisciplinary field can be divided into three main phases: before 2018, an embryonic period with low publication output (fewer than 10 articles per year); 2019–2022, a steady growth period with an average annual growth rate of 38.15%;

and 2023 to the present, a rapid expansion period, with annual publications exceeding 50 and reaching a historical peak of 96.

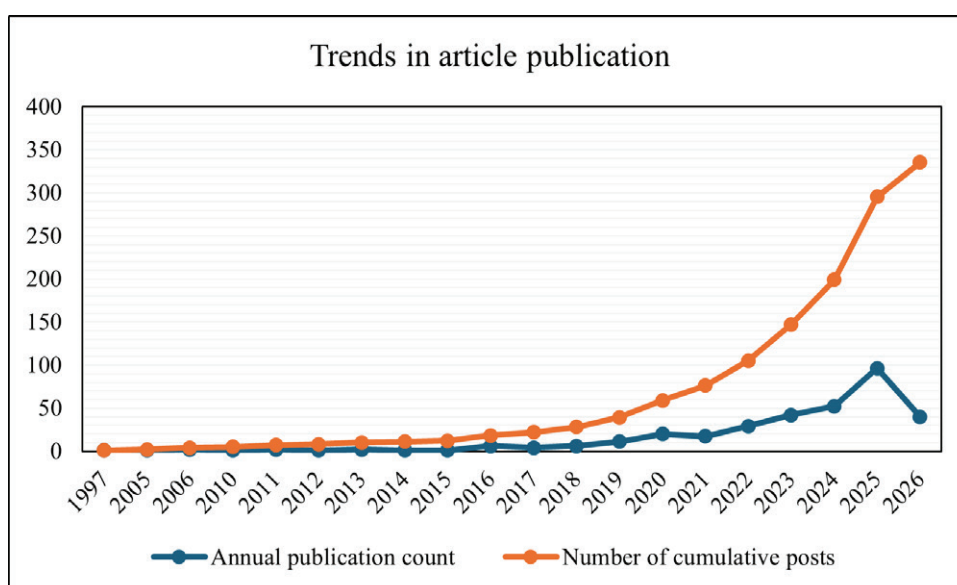
The literature included in the analysis was published in 198 journals. Table 1 lists the top 10 journals in terms of the number of published articles. The top three journals are Journal of Medical Internet Research (18 articles), Nicotine & Tobacco Research (18 articles), and PLOS One (11 articles), which together account for about 14%. These are the three most concentrated journals in this field. Among the journals included in SSCI, Nicotine & Tobacco Research, Digital Health, and other journals in the field of social science and public health account for a relatively high proportion; in terms of SCI journals, Journal of Medical Internet Research, PLOS One, and other journals in the field of medical informatics and computer science have outstanding performance. The distribution of journals shows clear dual-track characteristics: research results are published in journals of public health and behavioral science, as well as in journals of computer science and medical informatics.

Collaboration networks: countries, institutions, authors, and disciplines

Country collaboration network

The national-level scientific cooperation network is

Figure 2. Publishing trend on AI and tobacco use. The blue line represents the annual number of publications related to AI and tobacco use from 1997 to 2026, while the orange line shows the cumulative publication count each year. The overall upward trend is evident



shown in Figure 3. The United States ranked first with 143 articles, accounting for 42.69% of the total literature, and dominated the field; China (78 articles, 23.28%), Canada (26 articles, 7.76%), the

Table 1. Top 10 most prolific journals in AI and tobacco use research, 1997–2026

Journal	Number	Proportion (%)	Type	IF
Journal of Medical Internet Research	18	5.37	SCIE	6.0
Nicotine & Tobacco Research	18	5.37	SCIE/SSCI	3.0
PLOS One	11	3.28	SCIE	2.6
Digital Health	10	2.99	SCIE/SSCI	3.3
Drug and Alcohol Dependence	8	2.39	SCIE/SSCI	3.6
Scientific Reports	7	2.09	SCIE	3.9
Applied Sciences-Basel	6	1.79	SCIE	2.5
BMC Public Health	6	1.79	SCIE	3.6
Substance Use & Misuse	5	1.49	SCIE/SSCI	1.7
Tobacco control	5	1.49	SCIE/SSCI	4.7

SCIE: Science Citation Index Expanded. SSCI: Social Sciences Citation Index. IF: Impact Factor.

United Kingdom (20 articles, 5.97%), South Korea (17 articles, 5.07%) and India (17 articles, 5.07%) constituted the main research force. According to centrality indicators, the United States has the highest BC (0.38), followed by the United Kingdom (0.24) and China (0.13). The US–China, US–Canada, and US–UK cooperation networks are the main axes, but the overall density is low (0.062). High-income countries have an absolute advantage in the literature contribution in this field (about 83.28%), while the research participation of low- and middle-income countries is obviously insufficient¹⁶. The geographical distribution map of country cooperation (Figure 3) further shows that Europe and the United States have formed a relatively close cluster of cooperation. In contrast, the cooperation links among Asian, African, and Latin American countries are relatively sparse.

Institution collaboration network

The institutional-level cooperation network is shown in Figure 4. Among the top 10 research institutions, the University of California System (15 articles) and the China National Tobacco Corporation (12 articles) ranked first and second, respectively. University of Toronto, the Chinese Academy of Sciences, and the Center for Addiction & Mental Health–Canada ranked

Figure 3. National-level Scientific Cooperation Network (1997–2026). Nodes represent countries and connecting lines indicate cooperation between them. The size of the nodes and the thickness of the connecting lines are proportional to the countries' influence and the strength of their cooperation



third (10 articles). However, the cluster analysis of institutional cooperation networks shows that cooperation mainly occurs between institutions within the same country or region, and the direct cooperation of transnational institutions is relatively limited. The network modularity index ($Q=0.60$) indicates that the institutional cooperation network is highly modular, forming a core institutional group with clear agglomeration. In addition, some high-yielding institutions are comprehensive universities and their medical schools, and professional tobacco control research institutions (such as national tobacco control centers) are not prominent in terms of centrality and publication volume in the network.

Author collaboration network

The results of the author collaboration network analysis are shown in [Supplementary file](#) Figure 1. The field has a stable but small core group of authors. According to Price's Law, the minimum publication threshold for core authors is three articles. A total of 51 authors were identified as core authors, collectively contributing 187 articles (55.8% of the total), meeting

the 50% standard expected by Price's Law and indicating a stable core author group in the field. The top six high-output authors were P. Selby (7 articles); J. Rose (7 articles); N. Minian (5 articles); M. Ratto (5 articles); K. Mehra (4 articles); and L. Zawertailo (4 articles). The research topics of these authors are closely related, focusing on the application of chat robots and dialogic agents in smoking-cessation interventions. However, the overall density of the authors' cooperation network is only 0.01.

Interdisciplinary characteristics

Based on the Web of Science disciplinary classification system, the included documents involved 85 subject categories. [Supplementary file](#) Figure 2 (Disciplinary Collaboration Network) shows the distribution of publication volumes across disciplines. The top five disciplines by publication volume were: Public, Environmental & Occupational Health (73 articles, 21.79%); Health Care Sciences & Services (61 articles, 18.21%); Medical Informatics (60 articles, 17.91%); Substance Abuse (51 articles, 15.22%); and Psychiatry (29 articles, 8.66%). In terms of interdisciplinary

Figure 4. Institutional Collaboration Networks (1997–2026). Nodes represent institutions and connecting lines indicate collaborations between them. The size of the nodes is proportional to the journal's influence. The thickness of the connecting lines is proportional to the strength of collaboration

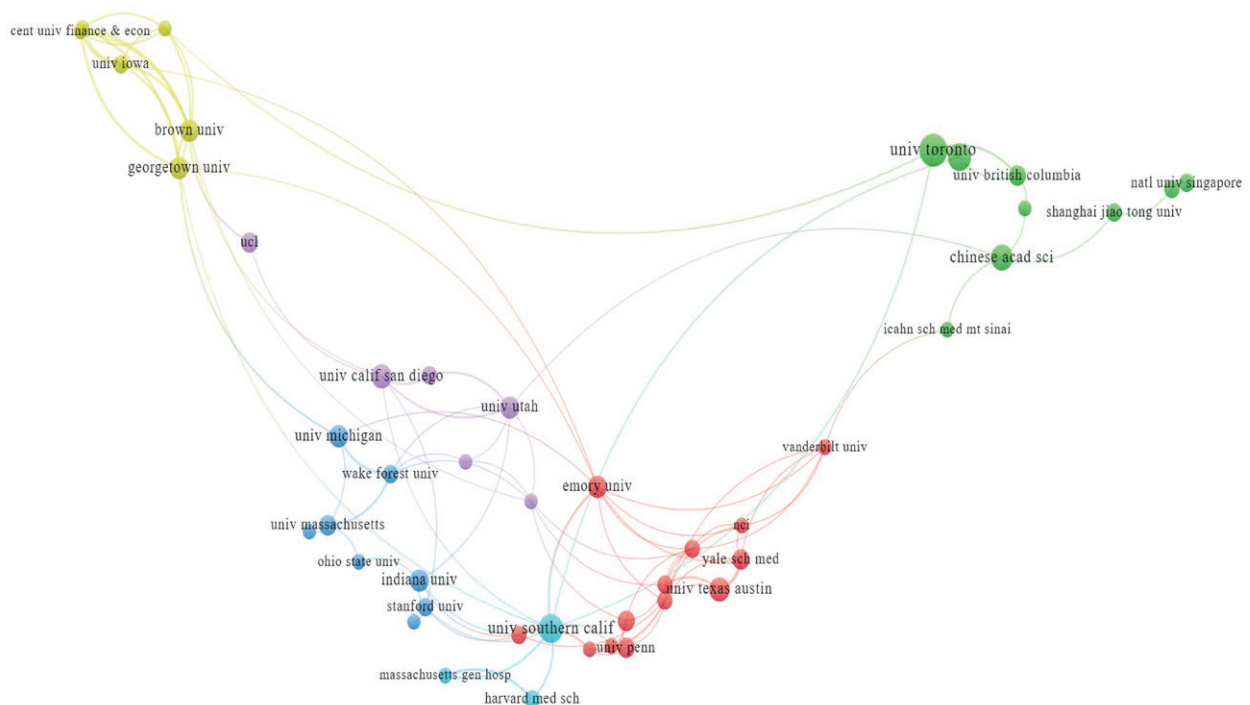


Table 2. First 20 high-frequency keywords, 1997–2026

Ranking	Keywords	Count	Ranking	Keywords	Count
1	machine learning	102	11	social media	22
2	smoking	70	12	nicotine	20
3	smoking cessation	51	13	chatbot	19
4	tobacco	30	14	classification	19
5	artificial intelligence	28	15	natural language processing	18
6	smokers	23	16	cigarette-smoking	17
7	United states	23	17	vaping	16
8	health	23	18	nicotine dependence	15
9	deep learning	23	19	smoking-cessation	15
10	cessation	22	20	risk	15

degrees, approximately 59.4% of the documents involved two or more subject categories, with an average of 1.87 categories per document.

Keyword co-occurrence analysis and research hotspots

Keyword co-occurrence network

[Supplementary file](#) Figure 3 shows the keyword co-occurrence network map, where node size represents keyword frequency, connections represent co-occurrence relationships, and node color represents cluster membership. The overall density of the network (the proportion of actual co-occurrence links to all possible keyword pairs) is 15.01%, and the total number of co-occurrence relationships is 2957. The top 20 high-frequency keywords are shown in Table 2. In addition to the core AI terms such as ‘machine learning’, ‘artificial intelligence’, ‘deep learning’, ‘chatbot’, and ‘natural language processing’. Tobacco-related terms such as ‘smoking’, ‘smoking cessation’, ‘tobacco’, ‘nicotine’, and ‘vaping’ also occupied a high frequency. It is worth noting that terms such as ‘United States’ (count=23), ‘health’ (count=23), ‘social media’ (count=22), and ‘cessation’ (count=22) appear more frequently.

Cluster analysis and hotspot themes

[Supplementary file](#) Figure 3 further presents the results of the keyword co-occurrence network clustering analysis, with a total of 7 main clusters identified. According to the results of cluster analysis and combined with high-frequency keywords, this

field can be summarized into the following seven research hotspots:

1. Machine learning-driven smoking behavior prediction and risk stratification (green clustering). This hotspot occupies the core position of the keyword network, including keywords ‘machine learning’, ‘smoking’, ‘prediction’, ‘exposure’, ‘risk’, ‘cigarette-smoking’, etc.
2. Natural language processing and social media monitoring (purple clustering). The cluster contains keywords such as ‘natural language processing’, ‘social media’, ‘Twitter’, ‘tobacco’, ‘sentiment analysis’, ‘data-mining’, ‘impact’, etc.
3. Application of deep learning and computer vision in smoking behavior detection (golden clustering). The cluster contains ‘deep learning’, ‘convolutional neural network’, ‘classification’, ‘behavior’, etc.
4. A series of problems, such as e-cigarettes for adolescents and youth groups (blue clustering). The cluster contained keywords such as ‘vaping’, ‘adolescents’, ‘youth’, ‘depression’, and ‘cigarette use’.
5. Conversational AI, chatbots, and smoking cessation (red clustering). The cluster contained keywords such as ‘chatbot’, ‘conversational agent’, ‘smoking cessation’, ‘artificial intelligence’, and ‘smoking’.
6. Health, media, and tobacco products (orange cluster). The cluster contains keywords such as ‘health’, ‘media’, ‘advertising and promotion’, ‘tobacco industry’, and ‘tobacco’.
7. Feature prediction of nicotine dependence and tobacco use (cyan clustering). The cluster contains

keywords such as ‘nicotine dependence’, ‘tobacco use’, ‘feature selection’, ‘predictive’, ‘models’, etc.

Keyword density heatmap

The keyword density heatmap is shown in [Supplementary file](#) Figure 4. The color in the figure changes from blue (low density) to green and then to yellow (high density), which intuitively presents the concentration of keywords in the research field. High-density areas were mainly clustered around words such as ‘machine learning’, ‘smoking’, ‘deep learning’, ‘smoking cessation’, ‘tobacco’, ‘health’, and ‘artificial intelligence’. Relatively speaking, the thermal strength of nodes such as ‘e-cigarettes’, ‘twitter’, and ‘mortality’ is relatively weak. It is worth noting that the heat of nodes such as ‘public health’, ‘large language model’, and ‘chatbot’ has shown a rapid upward trend in the near future^{6,10}. The heatmap of superposition time coloring shows a transition from blue to green to yellow, reflecting the dynamic migration of the technology frontier.

Keyword burst detection

A total of 30 keywords with significant burst strength were identified. The top 15 keywords by burst strength are shown in [Supplementary file](#) Figure 5, with red segments indicating burst periods. In terms of burst strength, early burst terms (approx. 2006–2019) mainly included ‘data mining’ (intensity=1.67) and ‘support vector machines’ (intensity=1.68), reflecting the field’s early stage characterized by basic algorithm applications. Mid-period burst terms (approx. 2019–2022) included ‘cigarette smoking’ (intensity=3.55), ‘association’ (intensity=2.71), ‘text mining’ (intensity=2.37), ‘social media’ (intensity=1.74), and ‘natural language processing’ (intensity=1.67), indicating a gradual expansion from traditional machine learning algorithms to text mining, NLP, and other techniques. Recent burst terms (2022–present) with the strongest burst strength include ‘youth’ (intensity=2.02), ‘model’ (intensity=2.02), ‘conversational agent’ (intensity=1.91), and ‘smokers’ (intensity=1.84).

DISCUSSION

Publication trends and field development characteristics

The bibliometric analysis of this study shows that the

literature output in the field of artificial intelligence and tobacco use research has evolved from slow accumulation to rapid expansion. Before 2018, the number of publications in this cross-cutting field was limited, and the topics were scattered. From 2019 to 2022, it entered a period of steady growth, and machine learning and natural language processing technologies began to systematically intervene in tobacco control issues. Since 2023, it has shown accelerating growth, and the introduction of generative AI and large language models has further expanded the scope of applications. This growth trajectory is consistent with the acceleration of global digital public health policy and the increasing emphasis on digital smoking cessation tools by the World Health Organization Framework Convention on Tobacco Control (WHO FCTC) in the time dimension^{17,18}.

In terms of collaboration networks, at the country level, the United States dominates the field in terms of publication volume and centrality, with China, the United Kingdom, and Canada constituting major research forces. The collaboration network shows a bilateral and multilateral pattern primarily structured around US–China, US–UK, and US–Canada axes, but the depth and breadth of cross-regional collaboration still have substantial room for improvement. At the institutional level, research capacity is concentrated mainly in universities, with limited cross-border collaboration. The author collaboration network similarly shows fragmentation, with a limited high-output author group and no stable core–periphery structure yet formed. This finding addresses the gap in existing AI- and tobacco-related review studies regarding the neglect of collaborative relationships among research topics^{6,19}. Notably, however, it diverges from broader tobacco research, which exhibits more frequent international collaboration and a distinctly global landscape²⁰.

From a disciplinary perspective, the field exhibits significant interdisciplinary characteristics. Public, Environmental & Occupational Health; Health Care Sciences & Services; Medical Informatics; Substance Abuse; and Psychiatry constitute the five pillar disciplines, each corresponding to different research orientations. Public health focuses on evaluating the population health effects of AI interventions²¹. Medical informatics and computer science concentrate

on building predictive models and decision support systems²²; and behavioral science and psychiatry examine the mechanisms of AI in individual behavior change⁹. However, the participation of sociology, communication studies, and policy science remains relatively limited, with insufficient attention to the social-structural dimensions of tobacco use behavior and to policy translation pathways. This aligns closely with the current state of tobacco research^{2,15} and underscores the need for broader engagement and integration from the social sciences.

Research hotspots and knowledge structure

Keywords in this field exhibit a high degree of correlation. Based on keyword co-occurrence networks and cluster analysis, we identified seven main research hotspots in the field, with a knowledge structure that exhibits a multi-layered progression from technical applications to public health practice. First, machine learning-driven prediction of smoking behavior and risk stratification constitutes the most intensive current research direction. Researchers widely use supervised learning algorithms such as Random Forest (RF), XGBoost, and Support Vector Machines (SVM) to predict smoking initiation, associated disease and mortality risks, and cessation success probabilities based on demographic characteristics, behavioral logs, and biomarker data^{6,23}. Second, natural language processing is extensively applied to monitor tobacco content on social media, using pre-trained language models such as BERT and Transformer for data and text mining to identify and analyze smoking behavior, its control, and health communication content^{24,25}. Third, personalized cessation interventions and message tailoring represent core applications of machine learning in behavioral support, with supervised learning dominating this subfield, covering message content selection, intervention timing optimization, and user feedback-driven adaptive adjustment²⁶. Fourth, image recognition and sequence modeling techniques based on deep learning architectures such as CNNs and RNNs are used for smoker behavior detection, health status assessment, and disease risk identification^{27,28}. Fifth, large language model-driven conversational intervention tools have become one of the most dynamic research directions, with LLM

chatbots demonstrating strong effectiveness and high user acceptance in smoking cessation assistance^{19,29}. Sixth, explainable AI and health equity focus on model decision transparency, the impact of algorithmic bias on health equity, and data privacy protection; the application of tools such as SHAP and LIME is still in its early stages³⁰. Seventh, AI techniques are used to predict, understand, and intervene in smoking behaviors and attitudes among special populations such as adolescents and young adults^{31,32}. The topics not only encompass the current mainstream research directions at the intersection of AI and tobacco studies^{9,28} but also bridge traditional technological approaches, such as SMS-based intervention program development¹³, with tobacco research.

Research frontiers and future directions

Burst detection and thematic evolution analysis reveal three main frontier directions. First, large language models and multimodal deep learning are driving a paradigm shift in smoking cessation interventions from 'reactive response' to 'real-time responsiveness'. Second, generative AI offers significant persuasive advantages in health communication, but it also carries the risk of misuse by the tobacco industry for marketing and promotion. This dual nature makes it an emerging governance issue that urgently requires attention. Third, the disciplinary gaps observed are also reflected in the burst term analysis, indicating that AI research remains largely confined to individual-level behavior, with insufficient attention to socioeconomic structural factors^{6,33}.

These frontier trends also reflect several key shortcomings in current research. Specifically, the limited diversity of data in current studies restricts the external validity of models. Second, weak interdisciplinary integration means that social determinants are not adequately incorporated. Third, the lack of interpretability and an ethical governance framework restricts practical application, while the systematic path from evidence to policy transformation under the FCTC framework remains lacking. Future research should focus on building a global open data platform, deepening the integration of social and behavioral science, promoting the development of interpretable AI applications, and establishing technical evaluation standards and ethical review

mechanisms under the FCTC framework. The use of AI technology in the tobacco industry poses a new challenge to the implementation of FCTC Article 5.3, and preventing the industry from using AI to erode public health policies should be a priority research topic.

Limitations

There are some limitations in this study. Firstly, the data source is limited to the Web of Science core collection, and other related databases such as Scopus and PubMed are not included, which may omit some literature in the fields of computer science and regional public health. Secondly, the bibliometric indicators mainly reflect the scale of research output and the characteristics of the network structure, and have limited evaluative power for the scientific contribution and policy influence of individual research. Finally, as only English-language literature was included in the screening process, the findings of this study may be subject to language bias, overlooking relevant knowledge production from non-English-speaking regions.

CONCLUSIONS

This study uses bibliometric and knowledge map analysis methods to systematically reveal the knowledge map and research patterns at the intersection of artificial intelligence and tobacco use. The study found that the number of publications in this field has continued to grow since 2018, forming a multi-research hotspot represented by machine learning prediction, social media content analysis, and conversational AI intervention. In general, interdisciplinary research at the intersection of artificial intelligence and tobacco control is in a period of rapid growth driven by technology. However, if the frontier algorithms are to be translated into practical public health outputs for tobacco control, it is still necessary to bridge multiple gaps across disciplinary boundaries, data bias, and policy transformation.

REFERENCES

1. Tobacco and nicotine. World Health Organization. June 26, 2026. Accessed August 3, 2026. <https://www.who.int/news-room/fact-sheets/detail/tobacco>
2. Makate M, Whetton S, Tait RJ, et al. Tobacco cost of illness studies: a systematic review. *Nicotine Tob Res.* 2020;22(4):458-465. doi:[10.1093/ntr/ntz038](https://doi.org/10.1093/ntr/ntz038)
3. World Health Organization. WHO global report on trends in prevalence of tobacco use 2000–2024 and projections 2025–2030; 2025. Accessed August 3, 2026. <https://iris.who.int/server/api/core/bitstreams/2eca3aea-b772-4272-a2ae-6fa26f3f9cd5/content>
4. Naughton F, Nadkarni A. The contribution of digital treatment to efforts to reduce global tobacco use. *N Engl J Med.* 2025;393(15):1449-1452. doi:[10.1056/NEJMp2500683](https://doi.org/10.1056/NEJMp2500683)
5. WHO Western Pacific Member States take action on key health priorities. World Health Organization Regional Office for the Western Pacific. October 24, 2025. Accessed August 3, 2026. <https://www.who.int/westernpacific/news/item/24-10-2025-who-western-pacific-member-states-take-action-on-key-health-priorities>
6. Fu R, Kundu A, Mitsakakis N, et al. Machine learning applications in tobacco research: a scoping review. *Tob Control.* 2023;32(1):99-109. doi:[10.1136/tobaccocontrol-2020-056438](https://doi.org/10.1136/tobaccocontrol-2020-056438)
7. Ng YJX, Cheong SHR, Ang WW, Lau Y, Lau ST. Types, functions, and mechanisms of machine learning for personalizing smoking cessation interventions: a systematic scoping review. *Artif Intell Med.* 2025;170:103274. doi:[10.1016/j.artmed.2025.103274](https://doi.org/10.1016/j.artmed.2025.103274)
8. Bendotti H, Lawler S, Chan GCK, Gartner C, Ireland D, Marshall HM. Conversational artificial intelligence interventions to support smoking cessation: a systematic review and meta-analysis. *Digit Health.* 2023;9:20552076231211634. doi:[10.1177/20552076231211634](https://doi.org/10.1177/20552076231211634)
9. Abrams LC, Yousefi A, Wysota CN, Wu TC, Broniatowski DA. Assessing the adherence of ChatGPT chatbots to public health guidelines for smoking cessation: content analysis. *J Med Internet Res.* 2025;27:e66896. doi:[10.2196/66896](https://doi.org/10.2196/66896)
10. Kong G, Ouellette RR, Murthy D. Generative artificial intelligence and social media: insights for tobacco control. *Tob Control.* 2025;34(6):846-848. doi:[10.1136/tc-2024-058813](https://doi.org/10.1136/tc-2024-058813)
11. Li Z, Wang X, Yao X, Chen Y. Do China's e-cigarette control policies work? A decade-long analysis of public discourse using an AI-integrated mixed-methods approach. *Tob Induc Dis.* 2025;23(September):143. doi:[10.18332/tid/208810](https://doi.org/10.18332/tid/208810)
12. Olawade DB, Aienobe-Asekharen CA. Artificial intelligence in tobacco control: a systematic scoping review of applications, challenges, and ethical implications. *Int J Med Inform.* 2025;202:105987. doi:[10.1016/j.ijmedinf.2025.105987](https://doi.org/10.1016/j.ijmedinf.2025.105987)
13. Sweileh WM. Technology-Based interventions for tobacco smoking prevention and treatment: a 20-year bibliometric analysis (2003-2022). *Subst Abuse Treat Prev Policy.* 2024;19(1):13. doi:[10.1186/s13011-024-00595-w](https://doi.org/10.1186/s13011-024-00595-w)
14. Qi S, Hua F, Xu S, Zhou Z, Liu F. Trends of global health literacy research (1995-2020): analysis of mapping knowledge domains based on citation data mining. *PLoS One.* 2021;16(8):e0254988. doi:[10.1371/journal.pone.0254988](https://doi.org/10.1371/journal.pone.0254988)

15. Qi S, Hua F, Zhou Z, Shek DTL. Trends of positive youth development publications (1995–2020): a scientometric review. *Appl Res Qual Life*. 2022;17(1):421–446. doi:[10.1007/s11482-020-09878-3](https://doi.org/10.1007/s11482-020-09878-3)
16. Haddad CN, Mahler DG, Diaz-Bonilla C, Hill R, Lakner C, Lara Ibarra G. The World Bank's new inequality indicator: the number of countries with high inequality. World Bank Group; June, 2024. Policy research working paper 10796. Accessed August 3, 2026. <https://documents1.worldbank.org/curated/en/099549506102441825/pdf/IDU-bd155bac-6d78-43af-8833-f87564a9d6c8.pdf>
17. Oliveira da Silva AL, Bialous S. Global landscape of nicotine and tobacco products additives policies. *Tob Control*. doi:[10.1136/tc-2024-058622](https://doi.org/10.1136/tc-2024-058622)
18. Goel S, Walia D, Rajkhowa P, Goel C. WHO framework convention on tobacco control articles 9 and 10 implementation for sustainable tobacco control-a scoping review protocol. *Syst Rev*. 2025;14(1):73. doi:[10.1186/s13643-025-02822-2](https://doi.org/10.1186/s13643-025-02822-2)
19. He L, Balaji D, Wiers RW, Antheunis ML, Krahmer E. Effectiveness and acceptability of conversational agents for smoking cessation: a systematic review and meta-analysis. *Nicotine Tob Res*. 2023;25(7):1241–1250. doi:[10.1093/ntr/ntac281](https://doi.org/10.1093/ntr/ntac281)
20. Richa S, Praveen S, Albairiqi AA, et al. Global trends and emerging frontiers in smoking and smokeless tobacco research: a bibliometric analysis over the past decade. *Healthcare (Basel)*. 2025;13(11):1224. doi:[10.3390/healthcare13111224](https://doi.org/10.3390/healthcare13111224)
21. Yousefi F, Naye F, Ouellet S, et al. Artificial intelligence in health promotion and disease reduction: rapid review. *J Med Internet Res*. 2025;27:e70381. doi:[10.2196/70381](https://doi.org/10.2196/70381)
22. Eloranta S, Boman M. Predictive models for clinical decision making: deep dives in practical machine learning. *J Intern Med*. 2022;292(2):278–295. doi:[10.1111/joim.13483](https://doi.org/10.1111/joim.13483)
23. Kim N, Loh WY, McCarthy DE. Machine learning models of tobacco susceptibility and current use among adolescents from 97 countries in the Global Youth Tobacco survey, 2013–2017. *PLOS Glob Public Health*. 2021;1(12):e0000060. doi:[10.1371/journal.pgph.0000060](https://doi.org/10.1371/journal.pgph.0000060)
24. Hu M, Benson R, Chen AT, Zhu SH, Conway M. Determining the prevalence of cannabis, tobacco, and vaping device mentions in online communities using natural language processing. *Drug Alcohol Depend*. 2021;228:109016. doi:[10.1016/j.drugalcdep.2021.109016](https://doi.org/10.1016/j.drugalcdep.2021.109016)
25. Küçükali H, Erdoğan MS. AI for tobacco control: identifying tobacco-promoting social media content using large language models. *Nicotine Tob Res*. 2025;27(6):988–996. doi:[10.1093/ntr/ntae276](https://doi.org/10.1093/ntr/ntae276)
26. Perski O, Kale D, Leppin C, et al. Supervised machine learning to predict smoking lapses from ecological momentary assessments and sensor data: implications for just-in-time adaptive intervention development. *PLOS Digit Health*. 2024;3(8):e0000594. doi:[10.1371/journal.pdig.0000594](https://doi.org/10.1371/journal.pdig.0000594)
27. Karlsson A, Ellonen A, Irjala H, et al. Impact of deep learning-determined smoking status on mortality of cancer patients: never too late to quit. *ESMO Open*. 2021;6(3):100175. doi:[10.1016/j.esmoop.2021.100175](https://doi.org/10.1016/j.esmoop.2021.100175)
28. Khan A, Elhassan MAM, Khan S, Deng H. Deep learning-based smoker classification and detection: an overview and evaluation. *Expert Syst Appl*. 2025;267:126208. doi:[10.1016/j.eswa.2024.126208](https://doi.org/10.1016/j.eswa.2024.126208)
29. Liu Y, Calle P, Vadakekut M, et al. AI-Enabled personalized smoking cessation intervention with the Aipaca chatbot: mixed methods feasibility study. *JMIR Form Res*. 2025;9:e73319. doi:[10.2196/73319](https://doi.org/10.2196/73319)
30. Singhal A, Neveditsin N, Tanveer H, Mago V. Toward fairness, accountability, transparency, and ethics in AI for social media and health care: scoping review. *JMIR Med Inform*. 2024;12:e50048. doi:[10.2196/50048](https://doi.org/10.2196/50048)
31. Leung J, Sun T, Stjepanovic D, et al. Generative Artificial Intelligence with youth codesign to create vaping awareness advertisements. *JAMA Netw Open*. 2025;8(7):e2514040. doi:[10.1001/jamanetworkopen.2025.14040](https://doi.org/10.1001/jamanetworkopen.2025.14040)
32. Sheeran P, Kenny A, Bermudez A, et al. Artificial Intelligence simulation of adolescents' responses to vaping-prevention messages. *JAMA Pediatr*. 2024;178(5):504–506. doi:[10.1001/jamapediatrics.2024.0274](https://doi.org/10.1001/jamapediatrics.2024.0274)
33. Rahman MS, Shiddik AB. Utilizing Artificial Intelligence to predict and analyze socioeconomic, environmental, and healthcare factors driving tuberculosis globally. *Sci Rep*. 2025;15(1):13619. doi:[10.1038/s41598-025-96973-w](https://doi.org/10.1038/s41598-025-96973-w)

CONFLICTS OF INTEREST

The authors have completed and submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest and none was reported.

FUNDING

There was no source of funding for this research.

ETHICAL APPROVAL AND INFORMED CONSENT

Ethical approval and informed consent were not required for this study.

DATA AVAILABILITY

The data supporting this research are available from the authors on reasonable request.

AUTHORS' CONTRIBUTIONS

SW: conceptualization, data curation, visualization. JZ: methodology, validation, resources. LC: investigation, software. All authors: writing of original draft, and writing, reviewing and editing of the manuscript. All authors read and approved the final version of the manuscript.

PROVENANCE AND PEER REVIEW

Not commissioned; externally peer reviewed.

SUPPLEMENTARY FILE DISCLAIMER

The content has been provided by the author(s) and has not been reviewed, verified, or endorsed by European Publishing. It may not have undergone peer review. The views, opinions, and recommendations expressed are solely those of the author(s) and do not necessarily reflect the position of European Publishing. European Publishing accepts no responsibility or liability for any consequences arising from the use of, or reliance on, this content.