British Journal of Hospital Medicine

Br. J. Hosp. Med. (Lond) **2026**; 87(10): 54896
https://doi.org/10.31083/BJHM54896

Article

Artificial Intelligence in Valvular Heart Disease: A Bibliometric and Knowledge-Mapping Analysis Revealing a Three-Phase Evolution Toward Integrated Clinical Intelligence

Xuan Wang[1,†], Shiyuan Liu[2,†], Lixue Zhang[1], Guangpu Fan[1], Shuhe Yu[2], Xu Liang[3], Ziping Li[4,*], Zeye Liu[1,*]

[1]Department of Cardiac Surgery, Peking University People's Hospital, Peking University, 100044 Beijing, China
[2]School of Basic Medical Sciences, Peking University Health Science Center, 100191 Beijing, China
[3]School of Automation and Intelligence, Beijing Jiaotong University, 100044 Beijing, China
[4]Department of Structural Heart Disease, Fuwai Hospital & National Center for Cardiovascular Disease, State Key Laboratory of Cardiovascular Disease, Chinese Academy of Medical Sciences & Peking Union Medical College, National Clinical Research Center for Cardiovascular Diseases, 100037 Beijing, China
*Correspondence: ziping_li@yeah.net (Ziping Li); liuzeyeyouxiang@163.com (Zeye Liu)
[†]These authors contributed equally.
Academic Editor: Paschalis Karakasis
Submitted: 15 June 2026 Revised: 25 August 2026 Accepted: 15 September 2026 Published: 26 October 2026

Abstract

Aims/Background: The application of artificial intelligence (AI) in valvular heart disease (VHD) has advanced significantly in recent years. However, a systematic evaluation of this field's historical development, knowledge structure, and global research patterns has yet to be conducted. This paper aims to outline the global evolution of AI research in VHD utilizing bibliometric and knowledge-mapping methods. **Methods:** Publications related to AI and VHD from 1991 to 2025 were retrieved from the Web of Science Core Collection database to investigate research progress. CiteSpace software was employed for bibliometric and knowledge mapping analyses, assessing indicators such as publication trends, keyword networks, collaboration patterns, co-citation relationships, and emerging topics. **Results:** A total of 3291 pieces of literature were identified. Research productivity remained low until the year before 2004, followed by steady growth and a sharp increase after 2015, culminating in 461 articles by the end of 2025. Network analysis revealed three sequential stages: image-assisted applications, prediction modeling, and recent advancements in deep-learning-based integrated systems. The United States accounted for the largest research volume, while England and Germany served as core nodes in the international cooperation network. Co-citation analysis identified significant clinical guidelines and pioneering interventional studies as key intellectual pillars in this domain. **Conclusion:** Over the past 35 years, significant progress has been achieved, transitioning from computer systems used solely for image analysis to more integrated clinical applications. Currently, deep learning techniques are increasingly prevalent, accompanied by a focus on multimodal data integration and a stronger linkage between computational methods and clinical decision-making.

Keywords: artificial intelligence; valvular heart disease; bibliometric analysis; knowledge mapping

1. Introduction

Valvular heart disease (VHD) represents a significant and growing contributor to cardiovascular morbidity and mortality worldwide [1,2]. With an aging population, improved survival from other cardiovascular conditions, and enhanced detection of degenerative valve disease, the clinical burden of VHD is expected to increase substantially [1,3]. Contemporary management of VHD heavily relies on accurate anatomical characterization, quantitative imaging assessment, risk stratification, and timely selection of surgical or transcatheter interventions [2,4]. These features establish VHD as a data-rich and decision-intensive field.

Artificial intelligence (AI) has rapidly emerged as a transformative methodology in cardiovascular medicine [5,6]. Machine learning, deep learning, radiomics, computer-aided diagnosis, and multimodal modeling have increasingly been applied to echocardiography, computed tomography, electrocardiography, electronic health records (EHRs), and procedural imaging [5,6]. In VHD, AI has been utilized for automated image interpretation, valve segmentation, grading of disease severity, outcome prediction, procedural planning, and post-intervention surveillance [5,7,8]. The rapid expansion of transcatheter aortic valve replacement (TAVR), transcatheter edge-to-edge repair (TEER), and other structural heart interventions has further amplified the need for integrated analytical tools capable of processing complex imaging and clinical data [4].

Despite this growth, the global research landscape of AI applications in VHD remains insufficiently characterized. Existing discussions have largely focused on individual algorithms, imaging modalities, or clinical applications, while the broader intellectual structure, thematic evolution, collaborative networks, and emerging research fronts of this interdisciplinary field remain unclear [6,9]. It

Copyright: © 2026 The Author(s). Published by IMR Press.
This is an open access article under the CC BY 4.0 license.

Publisher's Note: IMR Press stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

remains unknown whether AI research in VHD has evolved through gradual linear expansion or distinct structural transitions driven by advances in data availability, computational methodology, and clinical demand.

Bibliometric and knowledge-mapping approaches offer a systematic framework for evaluating the development of a research field [10,11,12]. By integrating publication trends, keyword co-occurrence, co-citation networks, clustering structures, burst detection, and collaboration analysis, these methods reveal not only research productivity but also the organization and evolution of scientific knowledge. Such an approach is particularly useful for interdisciplinary domains like AI-driven VHD research, where methodological innovation and clinical translation are closely intertwined.

Hence, this study aimed to characterize the global landscape of AI applications in VHD from 1991 to 2025 using bibliometric and knowledge-mapping methods. Specific objectives included delineating publication trends, identifying major thematic clusters and emerging frontiers, mapping collaborative and intellectual networks, and constructing a conceptual framework to describe the evolutionary trajectory of AI in VHD.

2. Methods

2.1 Study Design

This study employed a bibliometric and knowledge-mapping analysis to characterize the global research landscape of AI applications in VHD from 1991 to 2025. An analytical framework was established to ensure transparency, reproducibility, and structural robustness, adhering to established scientometric standards commonly applied in digital health research [10,11,12]. The primary objective was to delineate the evolutionary trajectory, intellectual architecture, and collaborative structure of this rapidly expanding interdisciplinary field, rather than evaluating clinical efficacy.

2.2 Data Source and Retrieval Strategy

All bibliographic records were retrieved from the Web of Science Core Collection (WoSCC), a curated citation database renowned for its standardized indexing and comprehensive citation metadata, making it widely used in scientometric research [13]. The retrieval period extended from January 1, 1991, to December 31, 2025, allowing for a thorough assessment of the developmental trajectory of AI integration into VHD research. All records were downloaded on a single date to minimize potential bias associated with ongoing database updates.

A structured Boolean search strategy was designed using a dual-domain intersection framework to ensure both sensitivity and thematic specificity. The first domain (Group A) focused on AI-related methodologies, including machine learning, deep learning, neural networks, transformer-based architectures, multimodal learning, radiomics, and computer-aided diagnostic techniques. The second domain (Group B) encompassed VHD-related conditions and procedural terminology—such as aortic stenosis, mitral regurgitation, bicuspid aortic valve, tricuspid regurgitation, TAVR, TEER, valve repair, and valve replacement. TAVR and TEER were specifically included as they represent common interventional contexts for VHD management, ensuring the retrieval of studies where AI-driven imaging, risk stratification, and procedural planning are integrated into specific clinical workflows.

Within each domain, keywords were combined using the Boolean operator OR (union operation), while the two domains were interconnected using AND (intersection operation), guaranteeing that only publications containing at least one term from each group were included. This union–intersection strategy facilitated comprehensive coverage of both methodological and clinical dimensions, maintaining high relevance to interdisciplinary research at the intersection of AI and VHD. Detailed keyword classifications are presented in **Supplementary Table 1**.

The search field was restricted to "Topic" (covering title, abstract, and keywords), and document types were limited to original articles and reviews, with no language restrictions applied.

2.3 Study Selection and Data Processing

Following record retrieval, duplicate entries were removed using the built-in deduplication function of the Web of Science, and bibliographic metadata were exported in plain text format, including author names, institutional affiliations, abstracts, keywords, cited references, and citation data. To minimize fragmentation bias during network construction, standardization of terminology was performed. Synonymous expressions (e.g., transcatheter aortic valve implantation [TAVI] and TAVR), inconsistencies in abbreviations, and variations in country names were harmonized. Institutional naming variants were unified where feasible to enhance the accuracy of collaboration networks. These preprocessing procedures improved structural coherence and ensured reliable estimation of centrality and clustering metrics.

2.4 Bibliometric Analysis and Network Construction

Bibliometric analyses were conducted using CiteSpace (version 6.4.R1, Chaomei Chen, Philadelphia, USA), a Java-based scientometric platform designed to identify research fronts and visualize evolving knowledge structures [14,15]. Developed by Professor Chaomei Chen from Drexel University, Philadelphia, USA, CiteSpace was selected for its comprehensive integration of co-occurrence analysis, co-citation mapping, clustering algorithms, burst detection, and temporal visualization within a unified analytical environment.

Time slicing was established from 1991 to 2025 using one-year intervals to facilitate fine-grained longitudinal as-

sessment. Within each slice, the top 25 most frequent items were selected to construct the network, focusing on node types such as keywords, countries, institutions, and authors. Pathfinder was used as the pruning algorithm, applied to sliced networks, to eliminate redundant links and preserve meaningful structural relationships, thereby enhancing interpretability and reducing visual complexity.

To delineate the structural evolution of the field, temporal trends were examined using the timeline view and time-zone visualization provided by CiteSpace. While the overall knowledge network was constructed using the full 1991–2025 dataset to ensure stable topological metrics, the identification of phase-specific turning points was primarily guided by burst detection algorithms and the temporal distribution of keywords across the time slices. This dual approach enabled the capture of both the global intellectual architecture and sequential shifts in research focus.

2.5 Network Metrics and Cluster Validation

Network properties were evaluated using established scientometric indicators. Modularity (Q) quantified the degree of cluster separation, with values exceeding 0.3 indicative of meaningful community structure [16]; higher values correspond to stronger modular organization. Cluster homogeneity was assessed with the silhouette score (S), where values above 0.5 suggest acceptable internal consistency and those exceeding 0.7 denote high reliability [17].

Betweenness centrality was calculated to identify nodes functioning as structural bridges across clusters, highlighting their integrative role within the knowledge network. Nodes with a betweenness centrality greater than 0.1 are generally considered high-centrality nodes in bibliometric networks. Nodes with elevated centrality were interpreted as conceptual connectors linking methodological innovation and clinical application in AI-driven VHD research.

Keyword clustering employed the log-likelihood ratio (LLR) algorithm for its statistical discrimination capability and interpretability in sparse networks. Temporal evolution was assessed through timeline and time-zone visualizations to identify sequential thematic shifts. Emerging research fronts were detected using Kleinberg's burst detection algorithm, which identifies abrupt increases in keyword frequency and quantifies burst strength as an indicator of growth intensity [15].

2.6 Robustness and Sensitivity Analysis

Structural stability was examined via sensitivity analyses varying node selection thresholds across time slices. Specifically, we tested three commonly used thresholds: Top 15, Top 20, and Top 25. The results showed that the Q value varied by less than 0.05, and the number of major clusters remained constant at five major clusters. The S values remained consistently high across all three thresholds (S >0.9), further confirming the robustness of the cluster structure. Furthermore, the ranking of the top five high-centrality keywords remained unchanged across these parameter settings. These findings confirm that the overall network topology was robust to reasonable variations in node selection. Network density was calculated to assess sparsity and structural coherence; the observed low-density values aligned with typical interdisciplinary knowledge networks.

Burst interpretation was restricted to terms demonstrating sustained activity over consecutive years to minimize noise-driven fluctuations. Although bibliometric analysis does not assess study quality or causal inference, the use of a curated citation database, structured retrieval strategy, rigorous preprocessing, and validated network metrics enhances analytical reliability and reproducibility.

The observed low network density values, specifically 0.008 for the keyword co-occurrence network, 0.0337 for the country collaboration network, 0.0055 for the institutional network, and 0.0045. These low values are consistent with the intrinsically sparse structure of large-scale interdisciplinary knowledge networks and reflect a standard feature of bibliometric mapping rather than a flaw in network construction.

3. Results

A total of 3291 publications were identified between 1991 and 2025. Annual publication output remained below 20 articles from 1991 to 2003 (Fig. 1). From 2004 onward, the volume of publications increased steadily, with a marked acceleration post-2015, particularly after 2020. Annual output reached a peak of 461 publications in 2025, representing the highest value during the study period.

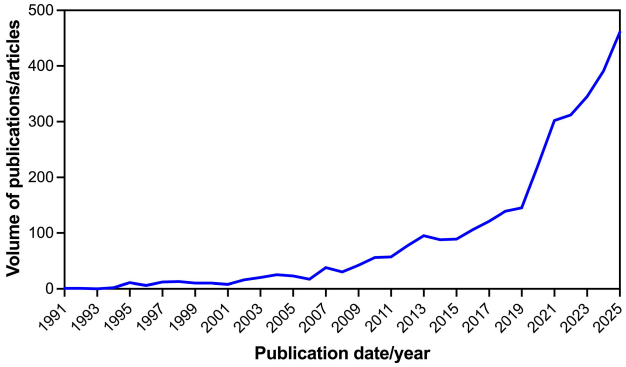


Fig. 1. Annual publication trends of artificial intelligence applications in valvular heart disease (1991–2025). The x-axis represents publication year (1991–2025), and the y-axis represents the annual number of publications. The blue line indicates the yearly publication output. The figure was generated by the authors using Microsoft Excel [Version 16.0.20228.20190, Microsoft Corporation, Redmond, WA, USA]. **Abbreviations**: AI, artificial intelligence; VHD, valvular heart disease.

Table 1. Centrality metrics of high-impact keywords (1991–2025).

Rank	Keyword	Occurrences (n)	Betweenness centrality	Interpretation
1	replacement	195	0.36	Core bridging term linking intervention studies
2	echocardiography	208	0.31	Key diagnostic modality across clusters
3	disease	233	0.26	General conceptual hub
4	diagnosis	174	0.25	Central clinical process node
5	mitral valve	97	0.23	Major disease-specific cluster
6	aortic regurgitation	43	0.23	Valve pathology linkage
7	stenosis	285	0.19	High-frequency disease phenotype
8	bicuspid aortic valve	80	0.19	Congenital valve research node
9	implantation	190	0.17	Procedural/therapeutic link
10	dysfunction	55	0.17	Outcome-related concept

Legend: Betweenness centrality quantifies the extent to which a keyword lies on the shortest paths between other nodes, highlighting its bridging role across different research clusters. Higher values indicate greater structural importance within the knowledge network. Occurrences represent the frequency of each keyword in the dataset.

The overall developmental pattern is summarized in the Graphical Abstract, which integrates publication trends, keyword networks, and knowledge structures into a three-phase framework.

The phase boundaries were determined based on the temporal distribution of burst terms and change-point detection of annual publication volume, specifically identifying 2004 and 2015 as significant structural turning points.

3.1 Phase I (1991–2003): Assistive Digitalisation Anchored in Imaging Quantification

The overall keyword co-occurrence network (1991–2025) exhibited a Q value of 0.8165 and an S value of 0.9337. Before characterizing the Phase I network, it is important to note that the overall keyword network (1991–2025) exhibited high betweenness centrality for terms such as “replacement” (0.36), “echocardiography” (0.31), and “disease” (0.26) (Table 1). However, when focusing specifically on the Phase I (1991–2003) period, our temporal analysis revealed that the early research landscape was primarily anchored by imaging and disease-related core keywords such as “mitral regurgitation”, “stenosis”, and “diagnosis”, which dominated the foundational phase of AI development in VHD.

Frequently occurring early keywords included “mitral regurgitation”, “stenosis”, “diagnosis”, and “disease”, all of which appeared in the earliest time slices of the network (Fig. 2). While AI-related terms began to emerge during this period, they remained in peripheral positions of the network.

Cluster analysis identified several prominent communities, including #0 mitral regurgitation, #7 mitral valve, and #9 transesophageal echocardiography (Fig. 3). Timeline visualization confirmed that these clusters first appeared during the 1990s and persisted in subsequent years (**Supplementary Fig. 1**).

Burst analysis indicated sustained activity for several imaging-related terms, including “transesophageal echocardiography” (1998–2017), “Doppler echocardiography” (1991–2014), and “doppler heart sounds” (2003–2009) (**Supplementary Table 2**). The network remained relatively dispersed, with limited connectivity among major clusters.

3.2 Phase II (2004–2014): Predictive Modelling and Algorithmic Structuring

Publication output increased from 25 to 88 publications per year during this period (Fig. 1). During Phase II (2004–2014), the temporal analysis revealed the emergence of predictive modeling and algorithm-related terminology, with key terms including “feature extraction” and “mitral valve prolapse”, while the early disease-oriented cluster “mitral regurgitation” has persisted from the initial phase to the present (Table 2).

The largest cluster was “mitral regurgitation” (mean year = 2001; 36 keywords), followed by clusters such as “takayasu arteritis” (mean year = 2003; 32 keywords), “feature extraction”, and “mitral valve prolapse” (Table 2). The early emergence of Takayasu arteritis may reflect the inclusion of studies on systemic inflammatory and vascular diseases that share imaging modalities with valvular heart disease (VHD) research, rather than studies directly focused on VHD itself. Given the broad scope of the search strategy, this finding should be interpreted with caution. Keywords within the “feature extraction” cluster included “convolutional neural networks”, “heart valves”, and “Doppler heart sounds”.

Timeline visualization demonstrated the sustained presence of earlier disease-oriented clusters, such as “mitral regurgitation”, alongside the emergence of additional computationally related clusters, such as “machine learning” and “feature extraction” (**Supplementary Fig. 1**). In the time-zone map (Fig. 4), disease-specific and diagnostic keywords predominantly appeared in earlier years, while terms such as “machine learning” emerged later within this phase.

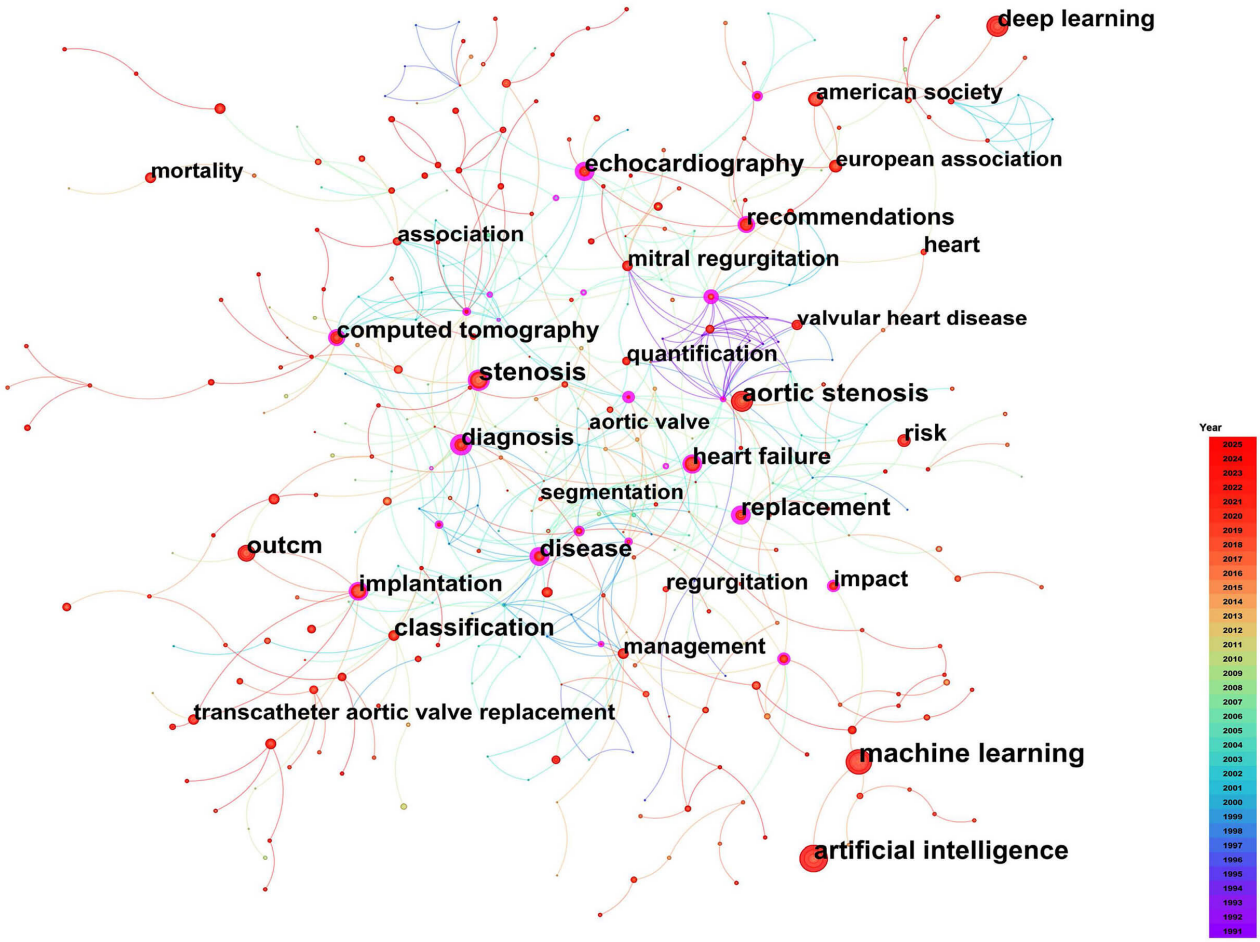


Fig. 2. Keyword co-occurrence network of artificial intelligence applications in valvular heart disease (1991–2025). Nodes represent keywords, with node size proportional to occurrence frequency. Links indicate co-occurrence relationships between keywords, with thicker links representing stronger co-occurrence strength. Node colours correspond to the year of first appearance according to the colour scale on the right. The figure was generated by the authors using CiteSpace. The abbreviation "Outcm" in the nodes represents "Outcome", which was automatically truncated by the network visualization software.

Compared to the preceding period, the network displayed increased connectivity, although major clusters remained clearly separated within the clustering structure (Fig. 3).

3.3 Phase III (2015–2025): Deep Learning and Integrated AI Ecosystems

Publication activity significantly increased during this phase, culminating in 461 publications in 2025 (Fig. 1).

In the keyword network, "machine learning" emerged as one of the largest nodes, demonstrating multiple connections with terms such as "artificial intelligence" (Fig. 2). Interconnections among major clusters were more extensive than in earlier periods.

Keyword burst analysis identified "artificial intelligence" (Strength = 28.0; 2023–2025) and "deep learning" (Strength = 15.28; 2023–2025) as the strongest recent bursts (**Supplementary Table 2**). In the time-zone visualization, both terms were observed within the most recent time slices (Fig. 4).

From a macro-perspective, the overall country-, institution-, and author-level collaboration networks for the entire study period exhibited increasing numbers of links and larger connected components (**Supplementary Figs. 2–6**), reflecting the global expansion of collaborative research over the past 35 years. Consistent with burst detection and timeline analyses, research interests have markedly expanded towards advanced computational frameworks such as deep learning and multimodal data integration during this recent period.

In summary, the transition toward deep learning has fostered a shift from single-task models to integrated, multimodal AI ecosystems, where diverse clinical data sources are combined to support comprehensive decision-making.

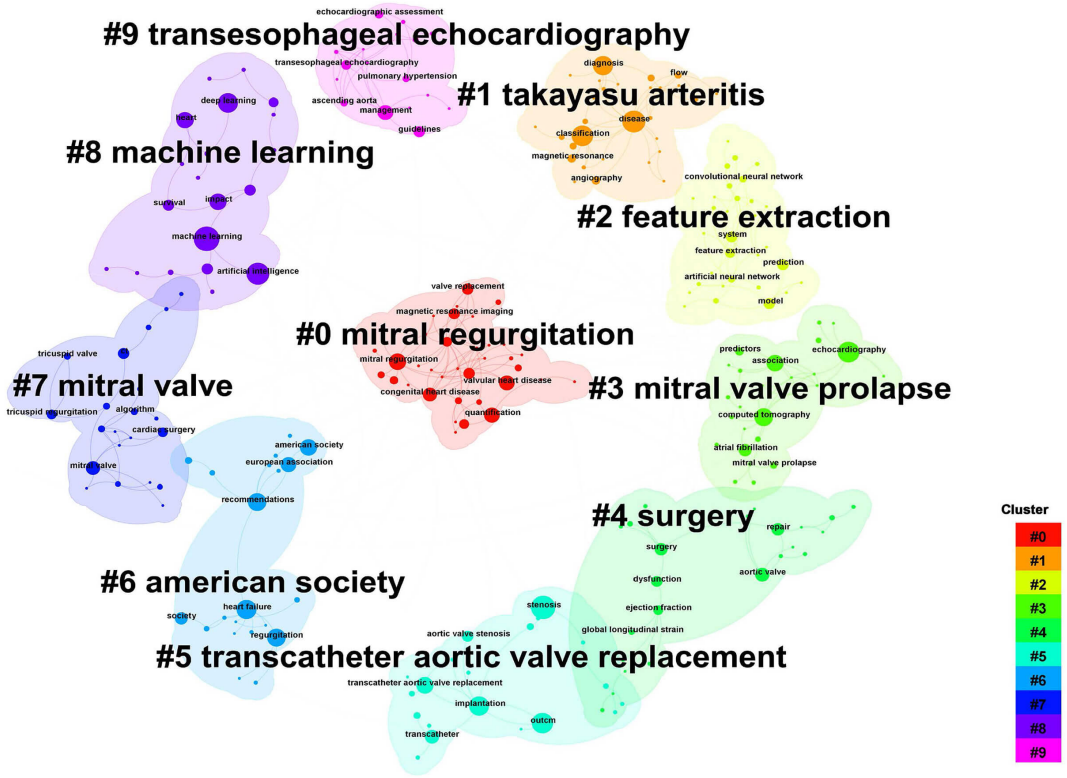


Fig. 3. Keyword clustering map of artificial intelligence research in valvular heart disease based on the log-likelihood ratio algorithm (1991–2025). Clusters were generated using the LLR algorithm. Each cluster is represented by a distinct colour and labelled numerically (e.g., #0–#9). Nodes represent keywords, with node size proportional to occurrence frequency, and links indicate co-occurrence relationships between keywords. The figure was generated by the authors using CiteSpace. The abbreviation "Outcm" in the nodes represents "Outcome", which was automatically truncated by the network visualization software. **Abbreviations**: AI, artificial intelligence; LLR, log-likelihood ratio; TAVR, transcatheter aortic valve replacement; VHD, valvular heart disease.

Table 2. Major keyword clusters identified in the co-occurrence network (1991–2025).

Rank	Cluster label	Representative keywords (Score, *p*-value)	Mean Year	No. of Keywords
1	mitral regurgitation	mitral regurgitation (60.47, 1.0×10^{-4}); magnetic resonance imaging (40.11, 1.0×10^{-4}); aortic regurgitation (40.11, 1.0×10^{-4}); valvular heart disease (38.48, 1.0×10^{-4}); quantification (30.02, 1.0×10^{-4})	2001	36
2	takayasu arteritis	takayasu arteritis (50.00, 1.0×10^{-4}); disease (41.22, 1.0×10^{-4}); diagnosis (31.81, 1.0×10^{-4}); arteries (26.84, 1.0×10^{-4}); MRI (25.31, 1.0×10^{-4})	2003	32
3	feature extraction	feature extraction (72.76, 1.0×10^{-4}); heart valves (39.98, 1.0×10^{-4}); Doppler heart sounds (34.36, 1.0×10^{-4}); fault diagnosis (28.74, 1.0×10^{-4}); convolutional neural networks (27.89, 1.0×10^{-4})	2010	31
4	mitral valve prolapse	mitral valve prolapse (31.87, 1.0×10^{-4}); computed tomography (27.99, 1.0×10^{-4}); cardiac magnetic resonance imaging (21.20, 1.0×10^{-4}); atrial fibrillation (20.69, 1.0×10^{-4}); mitral annular disjunction (15.80, 1.0×10^{-4})	2011	26
5	surgery	surgery (28.82, 1.0×10^{-4}); aortic valve (28.57, 1.0×10^{-4}); repair (23.99, 1.0×10^{-4}); 3D echocardiography (14.19, 0.001); transcatheter aortic valve implantation (13.04, 0.001)	2012	23

Legend: Mean year represents the average publication year of keywords within each cluster. The number of keywords indicates the size of each cluster. The Score and *p*-value in the "Representative Keywords" column indicate the log-likelihood ratio (LLR) test statistics and significance levels for each cluster label, respectively. A higher score indicates a more representative keyword for the cluster.
Abbreviation: MRI, magnetic resonance imaging.



Fig. 4. Time-zone visualization of keyword evolution in artificial intelligence research in valvular heart disease (1991–2025). The horizontal axis represents publication years from 1991 to 2025, and nodes are positioned according to the first year of keyword appearance. Each node represents a keyword, with size proportional to frequency. Links indicate co-occurrence relationships between keywords across time slices. Node colors correspond to temporal information, ranging from purple (earlier years) to red (more recent years), as shown in the color bar. The figure was generated by the authors using CiteSpace. The abbreviation "Outcm" in the nodes represents "Outcome", which was automatically truncated by the network visualization software. **Abbreviation**: TAVR, transcatheter aortic valve replacement.

3.4 Global Collaboration and Institutional Consolidation

Country-level analysis identified 101 countries connected by 170 collaborative links (Density = 0.0337) (**Supplementary Fig. 2**). The United States produced the largest number of publications (n = 1241), followed by China (n = 472) and Germany (n = 355). England (0.42) and Germany (0.36) exhibited the highest betweenness centrality values (Table 3). While the United States leads in total publication volume, its lower betweenness centrality may indicate a more self-contained domestic research network. In contrast, the UK and Germany function as pivotal structural brokers in global collaborative efforts, connecting diverse research communities.

Institutional collaboration analysis included 707 institutions connected by 1374 links (Density = 0.0055) (**Supplementary Fig. 3**). Mayo Clinic (n = 119), Columbia University (n = 91), and Harvard Medical School (n = 79) ranked as the most productive institutions. Duke University demonstrated the highest collaboration strength (degree = 34) (Table 4). Time-zone visualization indicated increasing institutional participation over time, particularly after 2010 (**Supplementary Fig. 4**).

Author collaboration analysis identified 1110 authors connected by 1894 links (Density = 0.0031)

Table 3. Top 10 most productive countries in the scientific collaboration network.

Rank	Country	First publication year	Publications (n)	Betweenness centrality
1	United States	1991	1241	0.25
2	China	2005	472	0.00
3	Germany	1995	355	0.36
4	Italy	1999	260	0.00
5	England	2002	243	0.42
6	Canada	1999	243	0.07
7	France	1997	189	0.15
8	Netherlands	1997	175	0.09
9	Switzerland	1998	111	0.00
10	Spain	2003	110	0.12

Legend: Betweenness centrality indicates the relative importance of each country within the collaboration network. The first publication year denotes the earliest recorded contribution of each country to the field.

(**Supplementary Fig. 5**). Leon MB and Hahn RT each contributed 29 publications and occupied central positions within the collaboration network (**Supplementary**

Table 4. Top 10 most productive institutions in the collaboration network.

Rank	Institution	First publication year	Publications (n)	Collaboration strength
1	Mayo Clinic	2003	119	25
2	Columbia University	2005	91	24
3	Harvard Medical School	2016	79	14
4	Cleveland Clinic	2008	74	15
5	Duke University	2003	55	34
6	Emory University	2002	53	25
7	Stanford University	2003	52	28
8	University of Pennsylvania	1999	52	19
9	Icahn School of Medicine at Mount Sinai	2014	49	26
10	Cedars-Sinai Medical Center	2000	47	30

Legend: Collaboration strength indicates the number of co-authored links associated with each institution in the network. The first publication year denotes the earliest recorded contribution of each institution to the field.

Table 3). The core author group, represented by the most highly productive scholars, began to take shape after 2010 (**Supplementary Fig. 6**).

3.5 Intellectual Foundations and Evidence Anchoring

The co-citation network comprised 788 nodes and 1394 links (Density = 0.0045) (Fig. 5). The most frequently co-cited references included Otto CM et al. [2] (2021, *Circulation*; 92 citations), Mack MJ et al. [4] (2019, *New England Journal of Medicine*; 83 citations), and Vahanian A et al. [18] (2022, *European Heart Journal*; 65 citations).

4. Discussion

In this bibliometric and knowledge-mapping study, 3291 relevant publications were included, representing over three decades of AI-related research on VHD. These publications provide comprehensive insights into the progression of this research over time. However, growth was not characterized by steady increases; rather, it occurred through multiple distinct periods, each with its unique features. Analysis of publication trends, keyword evolution, and network structures revealed three significant stages: the initial focus on image-assisted applications, followed by predictive modeling, and culminating in today's increasingly sophisticated AI-driven clinical systems.

Previous studies have collectively illustrated a phased development of AI in cardiovascular medicine, beginning with imaging-assisted analysis and progressing to predictive modeling utilizing structured clinical data, ultimately leading to the current stage characterized by deep learning-driven multimodal integration and end-to-end learning. Initially, AI primarily assisted clinicians in performing echocardiographic analyses and assessing structural heart disease through imaging techniques, including image segmentation and quantitative evaluation [7,19,20]. This demonstrated AI's supportive role in clinical practice. Subsequently, as EHRs and standardized datasets became more accessible, researchers began to implement machine learning techniques for risk prediction and outcome assessment, although applications remained limited to specific tasks [21]. Recent advancements in deep learning technology and theory have facilitated significant progress, enabling AI to achieve improved capabilities in multimodal integration and various clinical scenarios [6,22,23].

Limited research has been conducted on AI evolution, primarily focusing on publication counts [24] or keyword frequency [22], which view AI development as a straightforward, step-by-step progression. In reality, AI's evolution consists of several leaps rather than a continuous process. A significant turning point emerged around 2015, followed by another major change after 2020, correlating with deep learning's success in medical image processing [22] and multimodal analysis [25]. The transition from Phase II to Phase III has witnessed dramatic changes, not only in model accuracy but also in the profound reorganization of knowledge and enhanced connections between computational models and clinical entities. Thus, AI evolution encompasses model improvement and shifts in knowledge organization. This suggests that traditional bibliometric approaches relying on frequency measures may overlook significant developments. The methodology employed, which includes network topology, community structure detection, and burst detection, captures underlying discontinuities, offering insights into the shift from modular patterns to systemic integration within the field.

The broader cardiovascular field serves as the foundational background, while our study further focuses specifically on the application of AI in VHD. The three-phase evolution observed in VHD closely mirrors this general cardiovascular trajectory, indicating that AI development in VHD is synchronized with the broader discipline.

The patterns observed across these three phases likely result from interactions among data supply, technological advances, and clinical needs, aligning with major milestones in the history of cardiovascular AI evolution. During Phase I (1991–2003), many imaging-relevant terms emerged, yet few AI-specific concepts were present, suggesting that computation was perceived mainly as a tool

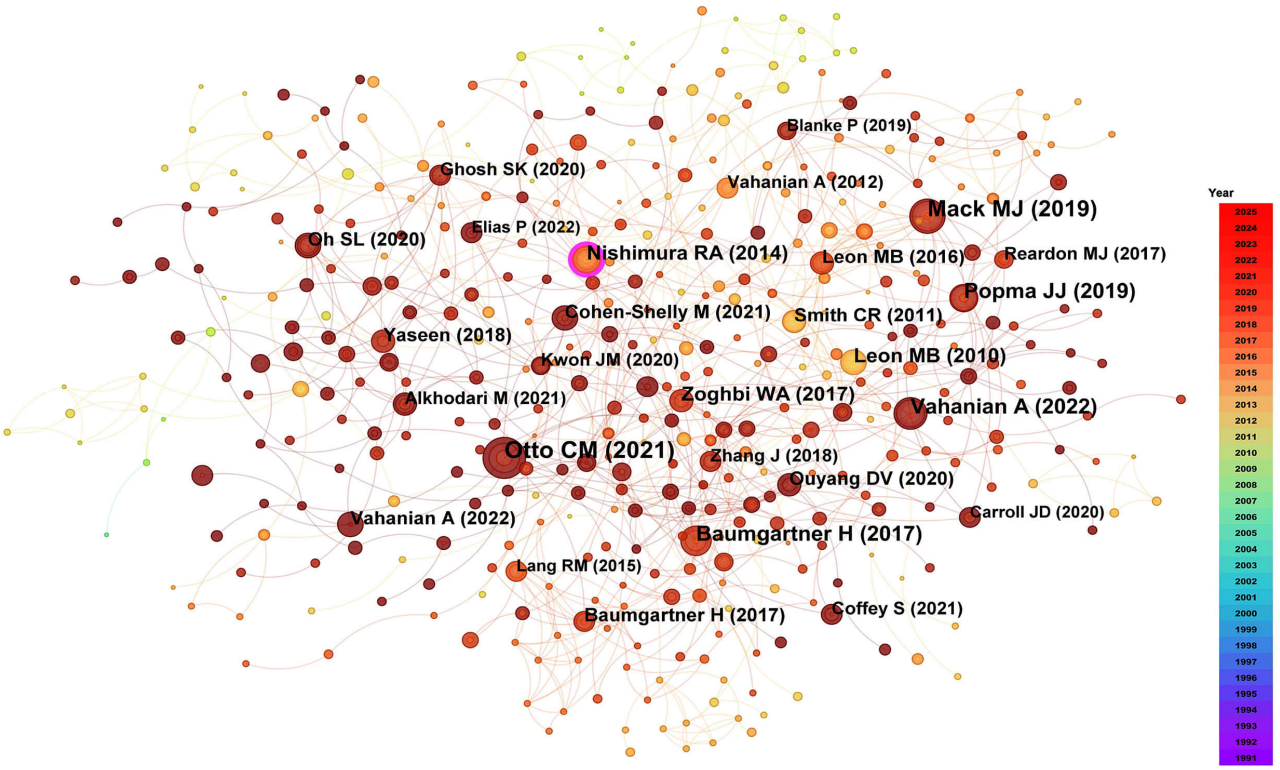


Fig. 5. Co-citation network of references in artificial intelligence research in valvular heart disease (1991–2025). Each node represents a cited reference, with node size proportional to citation frequency. Links indicate co-citation relationships between references. Nodes with multiple concentric rings indicate the temporal distribution of citations of the reference. The thickness of each ring represents the citation frequency in the corresponding year, with node colour reflecting the temporal progression from earlier (purple/blue) to more recent (red/orange) years. The figure was generated by the authors using CiteSpace.

to assist in manual interpretation tasks, such as measuring image features or performing quantitative analyses (e.g., 2D/3D echocardiographic quantification, ventricular function evaluation). This limitation stemmed from a lack of large structured datasets and relatively low computing power, coupled with algorithms primarily belonging to rule-based or statistical types in the early stages [26,27]. Consequently, computer-aided diagnosis (CAD) applications could only perform tasks like boundary detection, threshold division, and basic model recognition, lacking genuine learning capabilities; thus, AI could not operate independently but rather served an additive role [26,27].

In Phase II (2004–2014), the continuous increase in available clinical data resources—such as the digitalization of clinical data, including EHRs and structured databases—significantly promoted predictive modeling, coinciding with rapid advancements in machine learning algorithms. Various techniques, including support vector machine (SVM), random forest (RF), and earlier forms of neural networks, were applied to cardiovascular risk prediction, disease stratification, and outcome assessment [28]. Moreover, image analysis began to incorporate feature engineering methods alongside early radiomics concepts, laying the groundwork for future deep learning applications. Despite the emergence of various innovative strategies during this phase, the knowledge network remained semi-modular, indicating that most models were still task-specific and struggled to generalize across tasks and modalities. Similar to prior developments, AI-based systems during this phase were classified as narrow AI, making it challenging to integrate them into complex clinical workflows [24]. Thus, Phase II represents a transitional stage, marked by the increasing sophistication of AI-based methods, while complete system integration remained unachieved.

Phase III (2015 to present) marked the rise of deep learning, particularly in the realm of medical image recognition following the launch of ImageNet [29]. Notable research in this phase includes automated echocardiographic view classification and functional interpretation [30], estimation of left ventricular ejection fraction from cine magnetic resonance imaging (MRI) videos [8], and prediction of structural heart disease occurrences using electrocardiograms [31]. These studies demonstrate significant methodological advancements, transitioning from feature engineering-based approaches to end-to-end learning. Additionally, multimodal learning—integrating various types of inputs, such as imaging results, clinical records, and biological signals—and transformer-based frameworks have

further enhanced the analysis of complex clinical cases [32]. Consequently, our study revealed increased interconnections among nodes, indicating greater system integration compared to the previous modular structures. The impact of technological innovations extends beyond methodologies to influence clinical practice. The bibliometric trend illustrates a broadening scope of AI research across the clinical landscape, addressing areas from risk stratification to interventional planning. While direct evidence of routine clinical deployment is beyond this bibliometric analysis, the observed trends strongly suggest an increasing potential for AI to permeate multiple stages of patient care, from screening to therapeutic decision-making. In summary, deep learning has improved performance in general cardiovascular applications, and the bibliometric trends in our study suggest that it is progressively strengthening the linkage between computational methods and VHD-specific clinical process flows, positioning AI in the management of VHD as an emerging and integral component of a future comprehensive clinical intelligence system.

VHD offers a more suitable environment than other cardiovascular diseases (CVDs) for AI application due to four key reasons [2,4]: (1) VHD intrinsically involves medical imaging for diagnosis, severity classification, and treatment decisions, primarily relying on echocardiography, computed tomography (CT), and fluoroscopy. This data is high-dimensional, structurally rich, and rich in graphical information, aligning well with the strengths of deep learning, especially convolutional neural networks and transformer models [5]. (2) The targets in VHD, such as valve area, regurgitant volume, and transvalvular gradient, are relatively straightforward. In contrast, coronary artery disease presents complex diffuse lesions and pathological interactions that are not easily detectable. Hence, VHD provides clearer objectives for supervised learning. (3) Transcatheter interventional procedures (e.g., TAVR, TEER) involve specific operational processes where decisions depend heavily on measurement results from imaging modalities and risk stratification. In these scenarios, AI can augment decision-making, assist in patient selection, and aid in surgical planning [2,4]. (4) Evidence of standardization is particularly strong in the VHD field, especially within guideline-driven management contexts, suggesting a favorable environment for integrating AI into routine practice. Combining these factors, VHD may transition more rapidly than other CVDs from merely serving as an analytical tool to becoming an integral component of the clinical system.

4.1 Limitations

This study conducted a bibliometric analysis following the previously established methodological framework of bibliometrics and knowledge maps [10,11,12]. However, several limitations warrant careful consideration. Firstly, bibliometric studies inevitably encounter citation-based biases, including preferential citation of positive findings, frequent self-citations, and citation inflation associated with journal impact factors and author prominence [33,34,35]. Consequently, certain research directions may attract disproportionate scientific attention compared to others, skewing their visibility within the overall knowledge structure, while negative evidence and null findings often remain overlooked. Secondly, a temporal discrepancy exists between recent publication events and their subsequent citation records, resulting in the underrepresentation of novel studies in co-citation networks and burst detection analyses [36]. As a result, emergent research areas, particularly those involving the latest advancements in AI techniques such as foundation models and multimodal large-scale architectures, may not be adequately reflected in the knowledge map. Thirdly, this study relied solely on a single database (WoSCC), which, while recognized as reliable and authoritative, may introduce database selection bias. WoSCC tends to prioritize indexing journals with higher impact and articles published in English, potentially neglecting locally produced research and conference papers or technical reports disseminated via preprint servers and computer science archives, especially in fields like AI, where theoretical innovations often precede their acceptance by academic journals [37]. It should also be noted that the search strategy included generalized medical image analysis terms, which may have introduced some literature that is not strictly AI-specific. This may slightly broaden the scope of the retrieved literature beyond the strict VHD-AI intersection. Fourthly, the formation of knowledge networks heavily depends on specific techniques employed for keyword normalization, clustering methods, and filtering thresholds [15,16]. Despite numerous measures taken to minimize semantic variations among concepts and verify results through sensitivity tests, there remains a possibility of generating different network topologies and obtaining varying cluster categorizations under different conditions. While the clustering results presented herein are statistically supported by indices such as modularity value and average silhouette coefficient, caution is warranted when interpreting these classifications, given that they inevitably entail some degree of subjective judgment [16]. Fifthly, this paper utilizes bibliometric indicators such as total output volume and cumulative citation counts to assess researchers' productivity and influence. However, this approach may overlook other significant elements contributing to scientific contributions, such as methodological innovations, practical applications, and interdisciplinary collaboration. Lastly, the bibliometric approaches employed in this study serve primarily descriptive functions and do not evaluate research quality. They do not facilitate rigorous evaluations of causal relationships between interventions and outcomes or allow researchers to draw conclusions regarding clinical efficacy. Given these considerations, the tripartite evolution model described here should be seen as a structural construct and an abstract representation of the

overarching process rather than a direct reflection of technical performance levels or medical application values.

4.2 Future Perspectives

Future studies could explore ways to integrate these diverse aspects to construct composite indicators that reflect the multifaceted dimensions of scholarly influence [38]. Further research could integrate bibliometric analyses with systematic reviews or real-world verification experiments to enhance the explainability of the conclusions drawn here and improve the practicality of the resulting recommendations. Given the rapid advancements in AI technologies, the inherent publication lag associated with bibliometric mapping may lead to an underrepresentation of the most recent cutting-edge developments. Future studies might address this temporal gap by incorporating systematic reviews of recent preprints. Despite the potential shortcomings mentioned above, multiple lines of supporting evidence derived from analyses of diverse data types—such as annual publication trends, network connectivity features, similarities among clusters, and bursts in citation occurrences—confirm a strong agreement regarding the existence of observed patterns of change over time and validate the proposed conceptual explanations.

5. Conclusion

AI in VHD has experienced a structured, stage-wise transformation, evolving from imaging-based assistance to predictive modeling and ultimately to integrated clinical intelligence systems. This progression signifies a fundamental shift in the role of AI, transitioning from an auxiliary analytical tool to a potentially integrated component of future clinical systems. The framework proposed in this study lays the groundwork for understanding this transition and may inform future research, collaboration, and implementation strategies within AI-driven cardiovascular medicine.

Key Points

• This study provides the first systematic bibliometric and knowledge-mapping analysis of AI research in VHD over a 35-year period (1991–2025).

• The evolution of AI in VHD follows a distinct three-phase trajectory: from imaging-assisted quantification to predictive modeling using machine learning, and finally to integrated multimodal systems based on deep learning.

• Key research hotspots have shifted from early disease-oriented terms to algorithmic and procedural clusters, converging recently on AI, deep learning, and integrated clinical systems.

• VHD presents a promising and ideal scenario for AI implementation due to its strong reliance on imaging, well-defined anatomical targets, standardized interventional workflows, and guideline-driven management.

Availability of Data and Materials

The bibliographic data used in this study were obtained from the Web of Science Core Collection are publicly available. All data generated or analyzed during this study are available from the corresponding authors upon reasonable request. The official website link for the Web of Science is https://www.webofscience.com.

Author Contributions

ZYL and ZPL contributed to the study design, conceptualization of the manuscript, supervision, and manuscript preparation. SYL and XW conducted the formal analysis and drafted the manuscript. LXZ and GPF were involved in drafting the manuscript and contributed to the data acquisition and analysis. SHY and XL contributed to the interpretation of data. All authors have been involved in revising it critically for important intellectual content. All authors read and approved the final manuscript. All authors have participated sufficiently in the work to take public responsibility for appropriate portions of the content and agreed to be accountable for all aspects of the work in ensuring that questions related to its accuracy or integrity.

Ethics Approval and Consent to Participate

Not applicable.

Acknowledgment

The authors sincerely thank all researchers and clinicians whose work contributed to the foundation of this bibliometric analysis, as well as the developers of the Web of Science Core Collection and CiteSpace software for providing essential data and analytical tools. We also appreciate the support from our colleagues at Peking University People's Hospital, Peking University Health Science Center, Beijing Jiaotong University, and Fuwai Hospital.

Funding

This evaluation study was supported by the following grants: Peking University People's Hospital Scientific Research Development Funds (RDEB2024-13); Noncommunicable Chronic Diseases-National Science and Technology Major Project (2024ZD0539200); Liu Zeye Startup Fund for Introduced Talents (21960000513); and College Students' Innovative Experiment Project (DC20250152). The funders covered the costs associated with data retrieval and software licensing. They had no role in the study design, data analysis, interpretation, or writing of the report. This study was not supported by funding from pharmaceutical companies or other institutional sources.

Conflicts of Interest

The authors declare no conflicts of interest.

Supplementary Material

Supplementary material associated with this article can be found, in the online version, at https://doi.org/10.31083/BJHM54896.

References

[1] Vahanian A, Beyersdorf F, Praz F, Milojevic M, Baldus S, Bauersachs J, et al. 2021 ESC/EACTS Guidelines for the management of valvular heart disease. European Journal of Cardio-thoracic Surgery. 2021; 60: 727–800. https://doi.org/10.1093/ejcts/ezab389

[2] Otto CM, Nishimura RA, Bonow RO, Carabello BA, Erwin JP, 3rd, Gentile F, et al. 2020 ACC/AHA Guideline for the Management of Patients With Valvular Heart Disease: A Report of the American College of Cardiology/American Heart Association Joint Committee on Clinical Practice Guidelines. Circulation. 2021; 143: e72–e227. https://doi.org/10.1161/CIR.0000000000000923

[3] d'Arcy JL, Coffey S, Loudon MA, Kennedy A, Pearson-Stuttard J, Birks J, et al. Large-scale community echocardiographic screening reveals a major burden of undiagnosed valvular heart disease in older people: the OxVALVE Population Cohort Study. European Heart Journal. 2016; 37: 3515–3522. https://doi.org/10.1093/eurheartj/ehw229

[4] Mack MJ, Leon MB, Thourani VH, Makkar R, Kodali SK, Russo M, et al. Transcatheter Aortic-Valve Replacement with a Balloon-Expandable Valve in Low-Risk Patients. The New England Journal of Medicine. 2019; 380: 1695–1705. https://doi.org/10.1056/NEJMoa1814052

[5] Krittanawong C, Johnson KW, Rosenson RS, Wang Z, Aydar M, Baber U, et al. Deep learning for cardiovascular medicine: a practical primer. European Heart Journal. 2019; 40: 2058–2073. https://doi.org/10.1093/eurheartj/ehz056

[6] Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. Nature Medicine. 2019; 25: 44–56. https://doi.org/10.1038/s41591-018-0300-7

[7] Zhang J, Gajjala S, Agrawal P, Tison GH, Hallock LA, Beussink-Nelson L, et al. Fully Automated Echocardiogram Interpretation in Clinical Practice: Feasibility and Diagnostic Accuracy. Circulation. 2018; 138: 1623–1635. https://doi.org/10.1161/CIRCULATIONAHA.118.034338

[8] Ouyang D, He B, Ghorbani A, Yuan N, Ebinger J, Langlotz CP, et al. Video-based AI for beat-to-beat assessment of cardiac function. Nature. 2020; 580: 252–256. https://doi.org/10.1038/s41586-020-2145-8

[9] Johnson KW, Torres Soto J, Glicksberg BS, Shameer K, Miotto R, Ali M, et al. Artificial Intelligence in Cardiology. Journal of the American College of Cardiology. 2018; 71: 2668–2679. https://doi.org/10.1016/j.jacc.2018.03.521

[10] Donthu N, Kumar S, Mukherjee D, Pandey N, Lim WM. How to conduct a bibliometric analysis: An overview and guidelines. Journal of Business Research. 2021; 133: 285–296. https://doi.org/10.1016/j.jbusres.2021.04.070

[11] Zupic I, Čater T. Bibliometric Methods in Management and Organization. Organizational Research Methods. 2015; 18: 429–472. https://doi.org/10.1177/1094428114562629

[12] Aria M, Cuccurullo C. *bibliometrix*: An R-tool for comprehensive science mapping analysis. Journal of Informetrics. 2017; 11: 959–975. https://doi.org/10.1016/j.joi.2017.08.007

[13] Echchakoui S. Why and how to merge Scopus and Web of Science during bibliometric analysis: the case of sales force literature from 1912 to 2019. Journal of Marketing Analytics. 2020; 8: 165–184. https://doi.org/10.1057/s41270-020-00081-9

[14] Chen C. Searching for intellectual turning points: progressive knowledge domain visualization. Proceedings of the National Academy of Sciences of the United States of America. 2004; 101: 5303–5310. https://doi.org/10.1073/pnas.0307513100

[15] Chen C. CiteSpace II: Detecting and visualizing emerging trends and transient patterns in scientific literature. Journal of the American Society for Information Science and Technology. 2006; 57: 359–377. https://doi.org/10.1002/asi.20317

[16] Newman MEJ. Modularity and community structure in networks. Proceedings of the National Academy of Sciences of the United States of America. 2006; 103: 8577–8582. https://doi.org/10.1073/pnas.0601602103

[17] Chen C, Ibekwe-SanJuan F, Hou J. The structure and dynamics of cocitation clusters: A multiple-perspective cocitation analysis. Journal of the American Society for Information Science and Technology. 2010; 61: 1386–1409. https://doi.org/10.1002/asi.21309

[18] Vahanian A, Beyersdorf F, Praz F, Milojevic M, Baldus S, Bauersachs J, et al. 2021 ESC/EACTS Guidelines for the management of valvular heart disease. European Heart Journal. 2022; 43: 561–632. https://doi.org/10.1093/eurheartj/ehab395

[19] Sengupta PP, Shrestha S, Berthon B, Messas E, Donal E, Tison GH, et al. Proposed Requirements for Cardiovascular Imaging-Related Machine Learning Evaluation (PRIME): A Checklist: Reviewed by the American College of Cardiology Healthcare Innovation Council. JACC. Cardiovascular Imaging. 2020; 13: 2017–2035. https://doi.org/10.1016/j.jcmg.2020.07.015

[20] Samad MD, Ulloa A, Wehner GJ, Jing L, Hartzel D, Good CW, et al. Predicting Survival From Large Echocardiography and Electronic Health Record Datasets: Optimization With Machine Learning. JACC. Cardiovascular Imaging. 2019; 12: 681–689. https://doi.org/10.1016/j.jcmg.2018.04.026

[21] Deo RC. Machine Learning in Medicine. Circulation. 2015; 132: 1920–1930. https://doi.org/10.1161/CIRCULATIONAHA.115.001593

[22] Tran BX, Vu GT, Ha GH, Vuong QH, Ho MT, Vuong TT, et al. Global Evolution of Research in Artificial Intelligence in Health and Medicine: A Bibliometric Study. Journal of Clinical Medicine. 2019; 8: 360. https://doi.org/10.3390/jcm8030360

[23] Esteva A, Robicquet A, Ramsundar B, Kuleshov V, DePristo M, Chou K, et al. A guide to deep learning in healthcare. Nature Medicine. 2019; 25: 24–29. https://doi.org/10.1038/s41591-018-0316-z

[24] Guo Y, Hao Z, Zhao S, Gong J, Yang F. Artificial Intelligence in Health Care: Bibliometric Analysis. Journal of Medical Internet Research. 2020; 22: e18228. https://doi.org/10.2196/18228

[25] Coudray N, Ocampo PS, Sakellaropoulos T, Narula N, Snuderl M, Fenyö D, et al. Classification and mutation prediction from non-small cell lung cancer histopathology images using deep learning. Nature Medicine. 2018; 24: 1559–1567. https://doi.org/10.1038/s41591-018-0177-5

[26] Savadjiev P, Chong J, Dohan A, Vakalopoulou M, Reinhold C, Paragios N, et al. Demystification of AI-driven medical image interpretation: past, present and future. European Radiology. 2019; 29: 1616–1624. https://doi.org/10.1007/s00330-018-5674-x

[27] Noble JA, Boukerroui D. Ultrasound image segmentation: a survey. IEEE Transactions on Medical Imaging. 2006; 25: 987–1010. https://doi.org/10.1109/tmi.2006.877092

[28] Krittanawong C, Zhang H, Wang Z, Aydar M, Kitai T. Artificial Intelligence in Precision Cardiovascular Medicine. Journal of the American College of Cardiology. 2017; 69: 2657–2664. https://doi.org/10.1016/j.jacc.2017.03.571

[29] Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. Communications of the ACM. 2017; 60: 84–90. https://doi.org/10.1145/3065386

[30] Madani A, Arnaout R, Mofrad M, Arnaout R. Fast and accurate view classification of echocardiograms using deep learn-

ing. NPJ Digital Medicine. 2018; 1: 6. https://doi.org/10.1038/s41746-017-0013-1

[31] Attia ZI, Kapa S, Lopez-Jimenez F, McKie PM, Ladewig DJ, Satam G, et al. Screening for cardiac contractile dysfunction using an artificial intelligence–enabled electrocardiogram. Nature Medicine. 2019; 25: 70–74. https://doi.org/10.1038/s41591-018-0240-2

[32] Yang XY, Li YM, Wang JY, Jia YH, Yi Z, Chen M. Utilizing multimodal artificial intelligence to advance cardiovascular diseases. Precision Clinical Medicine. 2025; 8: pbaf016. https://doi.org/10.1093/pcmedi/pbaf016

[33] Ioannidis JPA. Why most published research findings are false. PLoS Medicine. 2005; 2: e124. https://doi.org/10.1371/journal.pmed.0020124 Erratum in: PLoS Medicine. 2022; 19: e1004085. https://doi.org/10.1371/journal.pmed.1004085

[34] Bornmann L, Daniel H. What do citation counts measure? A review of studies on citing behavior. Journal of Documentation. 2008; 64: 45–80. https://doi.org/10.1108/00220410810844150

[35] Tahamtan I, Safipour Afshar A, Ahamdzadeh K. Factors affecting number of citations: a comprehensive review of the literature. Scientometrics. 2016; 107: 1195–1225. https://doi.org/10.1007/s11192-016-1889-2

[36] Díez-Vidal A, Arribas JR. How recent is recent? Retrospective analysis of suspiciously timeless citations. British Medical Journal. 2025; 391: e086941. https://doi.org/10.1136/bmj-2025-086941

[37] Falagas ME, Pitsouni EI, Malietzis GA, Pappas G. Comparison of PubMed, Scopus, Web of Science, and Google Scholar: strengths and weaknesses. FASEB Journal. 2008; 22: 338–342. https://doi.org/10.1096/fj.07-9492LSF

[38] Waltman L. A review of the literature on citation impact indicators. Journal of Informetrics. 2016; 10: 365–391. https://doi.org/10.1016/j.joi.2016.02.007