Journal of Multidisciplinary Healthcare

Dovepress
Taylor & Francis Group

Open Access Full Text Article

ORIGINAL RESEARCH

Integrating Bibliometrics and Bioinformatics to Map Knowledge Structure, Trends, and Genetic Insights in Polycystic Ovary Syndrome and Tumors (2015–2024)

Meng Liu[1], Ying Xiong[2], Shao-Hua Zhang[3], Jun Yuan[4], Zhi-Qiang Cheng[1]

[1]National Center for Integrative Medicine, Oncology Department of Integrated Traditional Chinese and Western Medicine, China-Japan Friendship Hospital, Beijing, People's Republic of China; [2]Department of Radiation Oncology, China-Japan Friendship Hospital, Beijing, People's Republic of China; [3]Day Ward, China-Japan Friendship Hospital, Beijing, People's Republic of China; [4]Graduate School, Beijing University of Chinese Medicine, Beijing, People's Republic of China

Correspondence: Zhi-Qiang Cheng, National Center for Integrative Medicine, Oncology Department of Integrated Traditional Chinese and Western Medicine, China-Japan Friendship Hospital, Beijing, People's Republic of China, Email zhiqiangcheng@163.com

Objective: This study aims to construct a knowledge map of polycystic ovary syndrome (PCOS)-cancer research through bibliometric analysis to elucidate its developmental trajectory and global research landscape, and further employ bioinformatics approaches to investigate the underlying molecular mechanisms linking PCOS and related cancer.

Methods: Utilizing the Web of Science Core Collection as the data source, English-language publications from 2015 to 2024 were retrieved. CiteSpace and VOSviewer were employed for co-occurrence analysis, co-citation network construction, cluster identification, and keyword burst detection. PCOS and endometrial cancer-related genes were extracted from the Genecards database, followed by screening of overlapping genes for protein-protein interaction (PPI) network analysis to identify key targets. Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis was performed to pinpoint critical signaling pathways.

Results: Publications on PCOS and cancer exhibited a significant and steady growth over the past decade, with the United States and China demonstrating prominent contributions in both output volume and collaborative networks. *Frontiers in Endocrinology* and *Gynecological Endocrinology* jointly ranked first in publication count, while *The Journal of Clinical Endocrinology & Metabolism* received the highest citations. Keyword co-occurrence cluster analysis revealed major research hotspots including endometrial cancer, gene expression, and cardiovascular disease. Bioinformatics analysis identified 250 overlapping genes between PCOS and endometrial cancer. PPI network analysis highlighted TP53 as the most critical hub gene, and KEGG enrichment analysis underscored the pivotal role of the PI3K/AKT signaling pathway.

Conclusion: By integrating bibliometric analysis with bioinformatics, this study systematically maps the knowledge structure, emerging trends, and molecular mechanisms linking PCOS and cancer. Our findings specifically highlight the association between PCOS and endometrial cancer, may driven by dysregulation of the TP53 and PI3K/AKT signaling pathways. This work provides valuable insights for researchers to understand the foundational knowledge framework, identify emerging trends, potential collaborators, and mechanistic targets for future studies.

Keywords: polycystic ovary syndrome, cancer, PI3K/AKT signaling pathway, bibliometric analysis, CiteSpace, VOSviewer

Introduction

Polycystic Ovary Syndrome (PCOS) is a prevalent endocrine and metabolic disorder characterized by ovulatory dysfunction, hyperandrogenism, and polycystic ovarian morphology, predominantly affecting women of reproductive age.[1] The Global Burden of Disease (GBD) study reports a global age-standardized prevalence of PCOS at 1.68%, with marked geographical disparities—surpassing 7.9% in some countries.[2] Although PCOS is traditionally perceived as a disorder predominantly affecting the reproductive system, it is in fact a systemic condition encompassing metabolic,

Received: 23 April 2025
Accepted: 24 July 2025
Published: 5 August 2025

reproductive, cardiovascular, and psychological disturbances, accompanied by a spectrum of comorbidities.[3] The association between PCOS and neoplasms has garnered increasing attention, with compelling evidence from epidemiological investigations demonstrating a significantly elevated incidence of endometrial cancer, breast cancer, and ovarian cancer among affected individuals.[4]

Bibliometrics, a discipline employing statistical methods to quantitatively analyze academic literature, investigates patterns in publication volume, authorship, citation networks, and keyword distribution.[5,6] It serves to uncover disciplinary trends, identify core research communities, assess scholarly impact, and inform scientific decision-making. Despite rapid growth in research on PCOS and cancer, the global knowledge structure, research trajectories, and frontier topics linking PCOS to cancer remain systematically underexplored. To address this gap, this study leverages the Web of Science Core Collection (WoSCC) database and employs CiteSpace, VOSviewer and Scimago Graphica to analyze and visualize scholarly outputs from 2015 to 2024. Complementarily, we overlap genes linked to PCOS and cancer are extracted from the Genecards database, followed by protein-protein interaction (PPI) network analysis to identify key targets and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment to pinpoint critical signaling pathways. By synthesizing these dual perspectives, this work aims to aims to provide a macroscopic perspective and strategic guidance for future investigations into the PCOS-cancer association.

Materials and Methods

Literature Search Strategy

The data collection for this study was conducted on April 3, 2025, through the WoSCC (https://www.webofscience.com/wos/woscc/), with the search parameters configured as follows: databases were set to Science Citation Index Expanded (SCI-EXPANDED) and Social Sciences Citation Index (SSCI); the timeframe spanned from January 1, 2015, to December 31, 2024; document types were limited to research articles and reviews; and the language was restricted to English. The specific keyword combinations and detailed search strategy are provided in Table 1. The exported data were saved in plain text format and tab-delimited format to meet the input requirements of CiteSpace (6.4.R1) and VOSviewer, respectively.

Analytical Tools

CiteSpace Parameter Settings

CiteSpace 6.4.R1 was used to construct literature co-citation networks and analyze keyword clustering and evolution.[7] The data retrieved from WoSCC were imported into CiteSpace, with parameters set as follows: time slicing spanned from 2015 to 2024 (annual slices), threshold selection used the g-index (K=10), network optimization applied the Pathfinder algorithm to prune slice networks followed by secondary pruning after merging the networks, and visualization metrics linked node diameter to frequency of occurrence and line width to co-occurrence strength.

Table 1 Literature Search Strategy

Set	Results	Search Query
#1	16692	TS = ("Polycystic Ovary Syndrome" OR "PCOS" OR "Polycystic Ovarian Syndrome" OR "Ovary Syndrome, Polycystic" OR "Syndrome, Polycystic Ovary" OR "Ovarian Syndrome, Polycystic" OR "Polycystic Ovary Syndrome 1" OR "Sclerocystic Ovarian Degeneration" OR "Ovarian Degeneration, Sclerocystic" OR "Sclerocystic Ovary Syndrome" OR "Stein-Leventhal Syndrome" OR "Stein Leventhal Syndrome" OR "Syndrome, Stein-Leventhal" OR "Sclerocystic Ovaries" OR "Ovary, Sclerocystic" OR "Sclerocystic Ovary" OR "Ovarian Dysfunction" OR "Androgen Excess" OR "Hyperandrogenism" OR "Chronic Anovulation" OR "Polycystic Ovarian Morphology" OR "Polycystic Ovary Disease" OR "Insulin Resistance Syndrome")
#2	2220761	TS = ("Cancer" OR "Neoplasm" OR "Malignancy" OR "Tumor" OR "Carcinoma*" OR "Oncogenesis" OR "Oncology" OR "Metastasis" OR "Endometrial Cancer" OR "Endometrial Neoplasm*" OR "Uterine Cancer" OR "Uterine Neoplasm*" OR "Ovarian Cancer" OR "Ovarian Neoplasm*" OR "Breast Cancer" OR "Breast Neoplasm*" OR "Gynecologic Cancer" OR "Gynecologic Neoplasm*" OR "Hormone-dependent Cancer" OR "Hormone-related Malignancy" OR "Reproductive Cancer" OR "Reproductive Neoplasm*")
#3	1869	#1 AND #2

VOSviewer Analysis Workflow

VOSviewer 1.6.19 was utilized to process collaboration networks and co-occurrence analysis.[8] WoSCC export files were converted to UTF-8 encoding, and the network layout was optimized using the Linlog modularity algorithm. In the visualized networks, node size was proportional to either document count (publications) or citation count.

Application of Scimago Graphica

Scimago Graphica 1.0.25 was primarily employed to visualize country/region collaboration networks.[9] The GML-formatted country collaboration table obtained from VOSviewer was imported into Scimago Graphica, with parameters configured as follows: the label menu corresponded to "Country" and the cluster menu corresponded to "String". The generated country collaboration network mapped node diameter to national publication volume and line thickness to inter-country collaboration frequency.

WPS Excel

WPS Excel 2023 was used for basic statistical analysis, creating visualizations such as donut charts, bar charts, and column charts based on the extracted data.

Identification of Key Gene Targets and Enrichment Analysis

Target Gene Identification

High-relevance genes were retrieved from the GeneCards database (https://www.genecards.org/) using the keywords "Polycystic Ovary Syndrome" and "Endometrial Cancer"[10] Genes with a relevance score >25 were selected. Overlapping high-score genes shared between both diseases were identified as candidate genes. The intersecting genes was obtained through the Bioinformatics.com.cn platform (http://www.bioinformatics.com.cn/).[11,12]

PPI Network Construction

Overlapping genes were imported into the STRING database (https://cn.string-db.org/)[13] with the following parameters: Organism: Homo sapiens; Minimum interaction score: High confidence (0.700); Disconnected nodes hidden. The PPI network was exported in TSV format and visualized using Cytoscape software (version 3.7.2, https://cytoscape.org/).[14,15]

KEGG Pathway Enrichment Analysis

The overlapping genes were subjected to KEGG pathway analysis using the clusterProfiler package[16,17] (v4.6.2) on R 4.4.2 (https://www.r-project.org/). Significant pathways were filtered using the Benjamini-Hochberg method (FDR<0.05). Disease-related pathways were excluded. The top 10 enriched pathways were selected to construct a "gene-pathway" regulatory network, which was visualized in Cytoscape.

Results

Number of Articles Published in the Past Decade

A total of 1869 publications related to PCOS and cancer were identified through a comprehensive search of WoSCC database spanning from January 2015 to December 2024. Following rigorous screening procedures, 1746 articles were ultimately included in this study. The annual publication count exhibited two distinct phases. During the first phase (2015–2020), the number of articles showed a gradual upward trend. In the second phase (2021–2024), a rapid surge in publications was observed in 2021 compared to 2020 (68 articles, a 39.3% increase), followed by a plateau in subsequent years. In terms of citations, the cumulative number of citations for these articles demonstrated an exponential growth trend (Figure 1).

Country/Region Collaboration Analysis

The collaboration networks among countries/regions were clustered and visualized using Scimago Graphica (Figure 2A). The six collaboration clusters included the following countries/regions: Cluster 1: USA, Canada, United Kingdom, Spain, Italy, Poland, Finland, Denmark, China, and Australia; Cluster 2: Brazil and South Africa; Cluster 3:Turkey, India and Egypt; Cluster 4: Iran and malaysia; Cluster 5: Romania; Cluster 6: China Taiwan. The collaboration frequency and

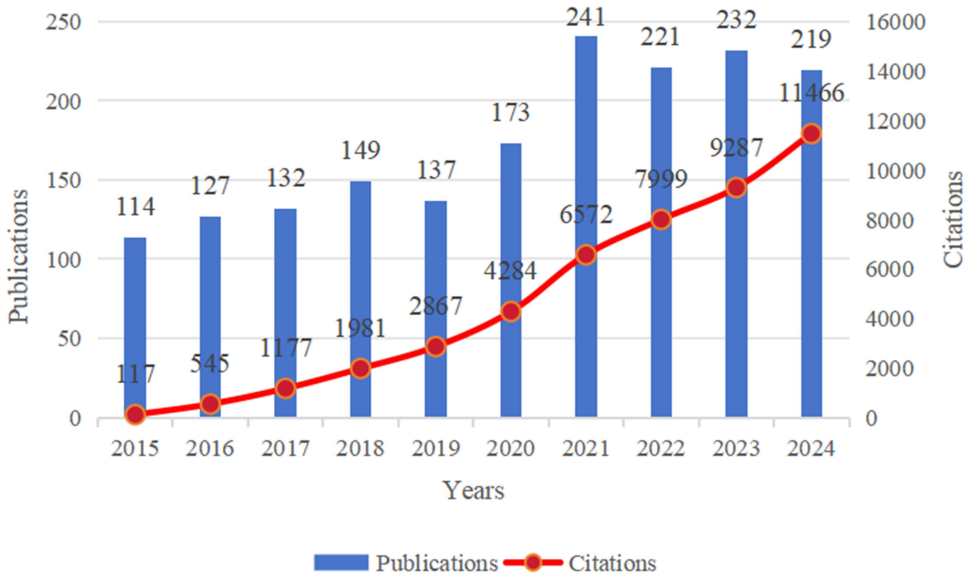

Figure 1 Trends in Annual Publications and Total Citations.

linkages between countries are shown in Figure 2B. China and the USA exhibited the highest international collaboration frequency. Additionally, China established close collaborations with Iran, the United Kingdom, and Canada, while the USA maintained frequent interactions with Turkey, Iran, and other countries. The overall network displayed a multi-hub structure, with key nodes such as China, the USA, Iran, and the United Kingdom, reflecting their dominant roles and connective capabilities in international collaboration. Annual publication counts by country are presented in Figure 2C. China's output surged rapidly after 2015, increasing from 14 articles in 2015 to 95 in 2021 and remaining high in 2024 (84 articles). In contrast, the USA maintained stable output with minimal fluctuations. National contributions to the field are detailed in Figure 2D. China's share rose from approximately 14% in 2015 to 39% in 2024, becoming the largest contributor, while the USA declined from 38% to 17% during the same period. To further quantify research performance, key metrics (publications, total citations, and average citations) for major countries are listed in Table 2. China ranked first in total publications (562 articles) with 11216 total citations but had a lower average citation rate (20.0 citations/article). The USA ranked second in publications (347 articles) but achieved the highest total citations (15929) and average citations (45.9 citations/article), reflecting its leading research quality and influence. Other countries, such as Australia (62.2 citations/article), France (57.4), and Italy (49.6), also demonstrated strong citation performance.

Academic Institution Influence and Collaboration Analysis

Collaboration networks among institutional affiliations were visualized using VOSviewer. Chinese universities and research institutions dominated in quantity, forming dense collaborative clusters centered around institutions such as Fudan University and Shanghai Jiao Tong University, which not only established close collaborations with domestic partners but also developed stable international linkages with universities in Iran, Sweden, the USA, and other countries (Figure 3A). The top 10 institutions by publication volume included Fudan University (34 articles), Tehran University of Medical Sciences (31 articles), Zhejiang University (28 articles), Shanghai Jiao Tong University (27 articles), Sichuan University (25 articles), Shandong University (23 articles), Zhengzhou University (22 articles), Karolinska Institute (21 articles), China Medical University (20 articles), and Harvard Medical School (20 articles) (Figure 3B). These institutions demonstrated strong research activity in the field, with their total citation counts shown in Figure 3C and Table 3. Despite Fudan University ranking first in output, Shandong University led in academic impact with 1293 total citations. Similarly, Karolinska Institute and Harvard Medical School, despite moderate publication volumes, achieved 794 and 545 citations, respectively, highlighting their significant scholarly influence in this research domain.

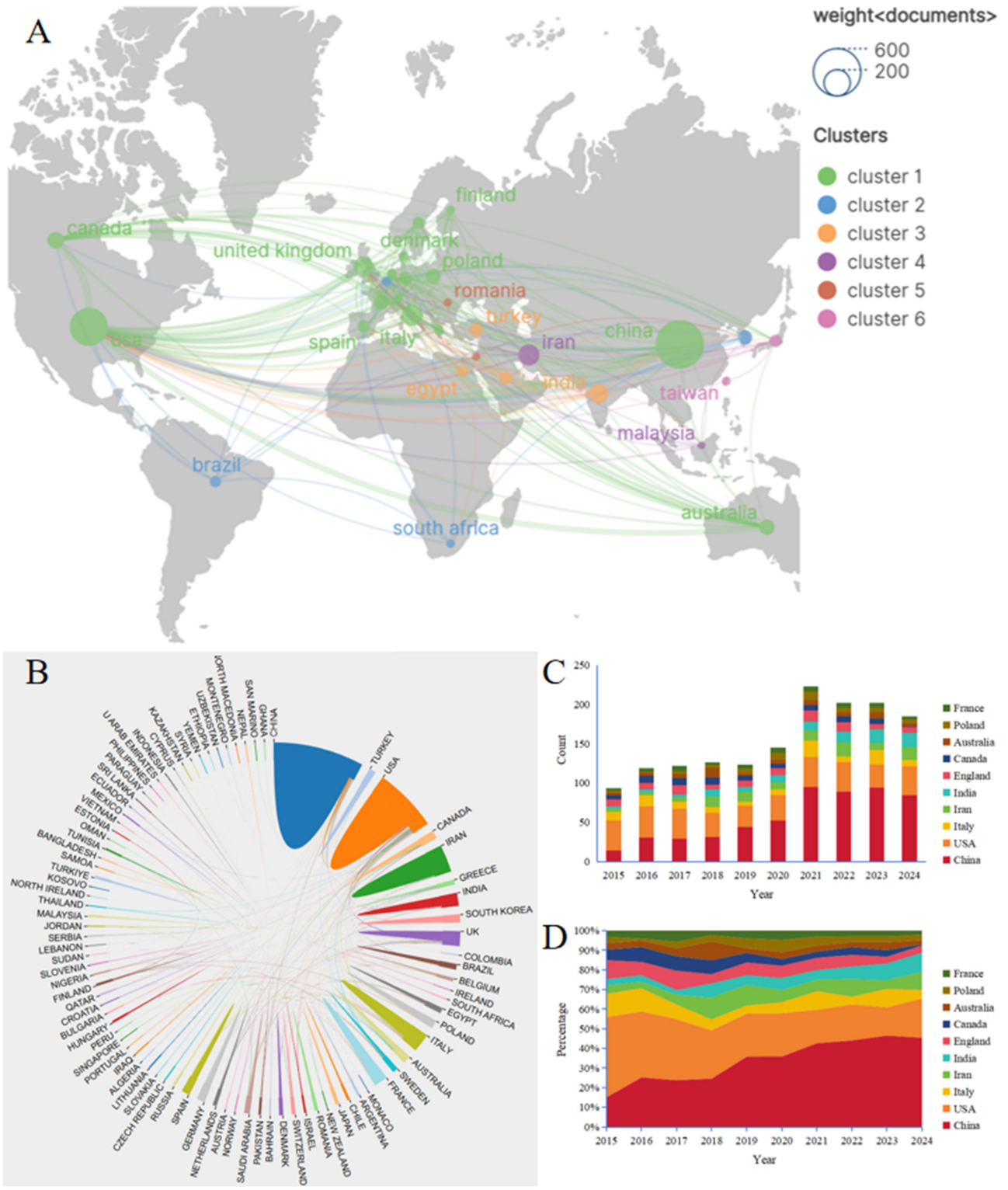

Figure 2 Global contribution and collaboration in PCOS and cancer research. (**A**) Country-level collaboration network. (**B**) Chord diagram of inter-country cooperation. (**C**) Annual publication trends of top 10 countries. (**D**) Proportional contributions of top countries over time.

https://doi.org/10.2147/JMDH.S536122

Table 2 Top 10 Countries by Publication Volume in PCOS and Cancer Research

Rank	Country/Region	Documents	Citations	Average Citations
1	China	562	11,216	20.0
2	USA	347	15,929	45.9
3	Italy	110	5454	49.6
4	Iran	107	2360	22.1
5	India	93	2188	23.5
6	United Kingdom	91	3917	43.0
7	Canada	66	1948	29.5
8	Australia	56	3483	62.2
9	Poland	54	1639	30.4
10	France	53	3042	57.4

Journal Distribution and Reference Analysis

The collaboration networks and citation networks of publication journals were visualized using VOSviewer. In terms of journal publication volume, the top 10 journals included *Frontiers in Endocrinology* (51 articles), *Gynecological Endocrinology* (51 articles), *Journal of Ovarian Research* (38 articles), *International Journal of Molecular Sciences* (36 articles), *Reproductive Sciences* (32 articles), *Nutrients* (26 articles), *Human Reproduction* (24 articles), *Journal of Clinical Endocrinology & Metabolism* (22 articles), *Molecular and Cellular Endocrinology* (19 articles), and *PLOS ONE* (19 articles). These journals demonstrated significant academic influence in the research field, with detailed metrics including Total Citations, Total Link Strength, Impact Factor, and Average Citations per Article provided in Table 4. Notably, *Journal of Clinical Endocrinology & Metabolism*, despite its moderate publication count (22 articles), high total citations (971 citations) and highest average citations per article (73 citations/article), reflecting its strong scholarly authority. The inter-journal collaboration network is shown in Figure 4A, and the bubble map illustrating journal publication volume versus citation counts is presented in Figure 4B.

Regarding cited references, the top 10 journals were *The Journal of Clinical Endocrinology & Metabolism, Fertility and Sterility, Human Reproduction*, PLOS ONE, Endocrinology, Human Reproduction Update, The New England Journal of Medicine, Proceedings of the National Academy of Sciences of the USA, The Journal of Biological Chemistry, International Journal of Molecular Sciences. highlighting their pivotal role in shaping interdisciplinary research on PCOS and cancer (Figure 4C). Detailed metrics are summarized in Table 5.

Author Contributions and Collaboration Networks

The collaboration networks among authors were analyzed using VOSviewer, revealing multiple stable collaborative clusters with distinct modular structures. For example, Chavarro Jorge E, Chen Zi-jiang, and Zhang Jing occupied central positions in separate collaborative cores (Figure 5A). In terms of publication count (Figure 5B and Table 6), Yang Jing (12 articles) and Asemi Zatollah (11 articles) ranked highest, followed by Chen Zi-jiang, Grynberg Michael, Li Li, and Li Yan (8 articles each). Arlt Wiebke, Azziz Ricardo, Chavarro Jorge E, and Du Toit Therina each published 7 articles. Among the top 10 authors, four were Chinese scholars: Yang Jing, Chen Zi-jiang, Li Li, and Li Yan. Further analysis of the author collaboration network (Figure 5C) identified Azziz Ricardo, Fauser BCJM, and Escobar-Morreale HF as central nodes, indicating their strong academic influence. Citation metrics for these authors are shown in Figure 5D. Azziz Ricardo led with 435 co-citations, followed by Legro *RS* (262 co-citations), Fauser BCJM (252 co-citations), Escobar-Morreale HF (240 co-citations), and Diamanti-Kandarakis E (239 co-citations). Despite moderate publication volumes, these scholars exhibited significant dissemination power and structural influence within the academic network.

Keyword Analysis

Through CiteSpace visualization and cluster analysis of the keyword co-occurrence network, 14 keyword clusters were identified (Figure 6A). Specific information and timelines of each cluster are as follows: Cluster #0 (fertility preservation)

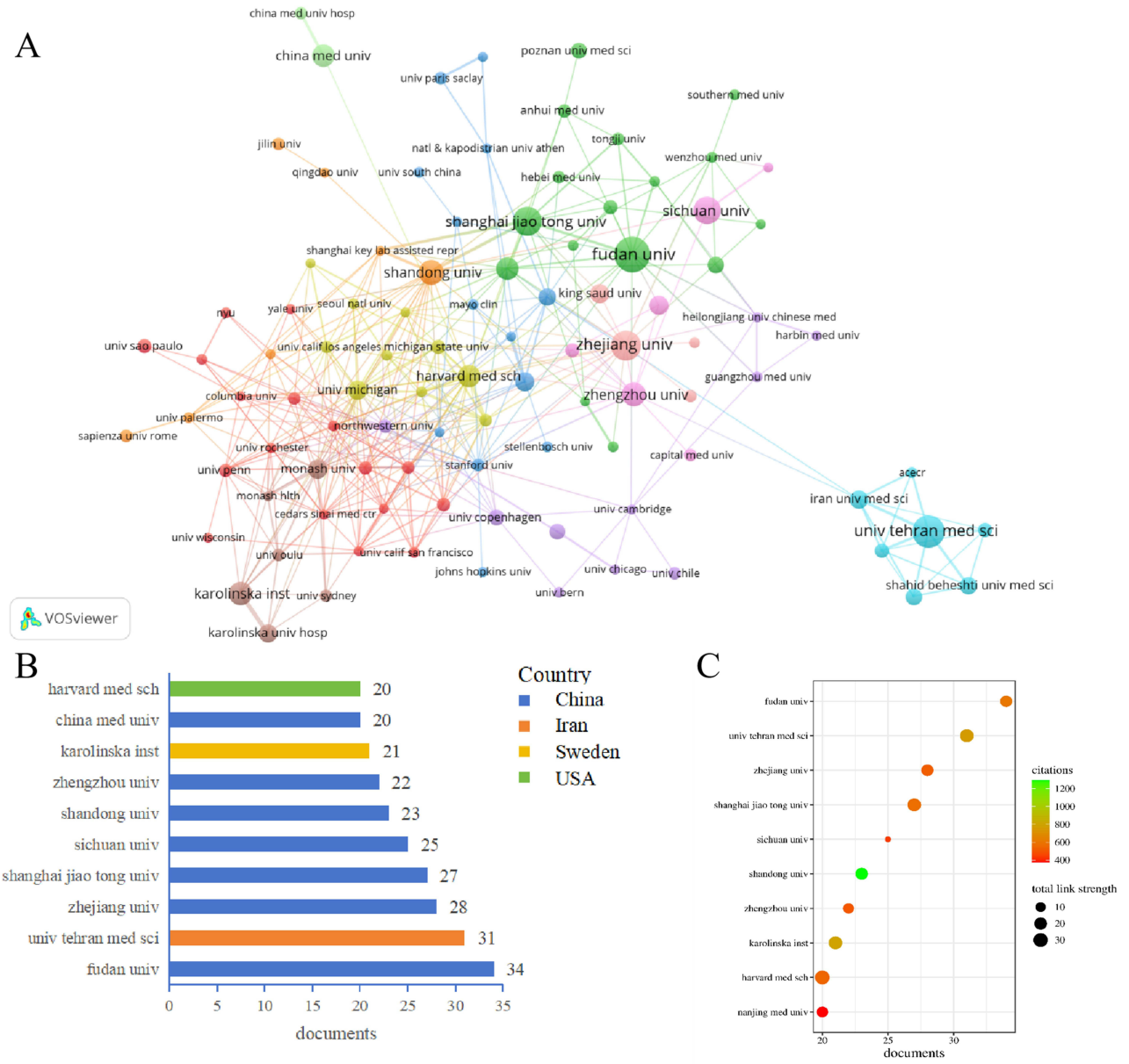

Figure 3 Institutional performance and collaboration in PCOS and cancer research. (**A**) Co-authorship network of research institutions. (**B**) Top 10 institutions ranked by publication volume. (**C**) Bubble chart of publications and citations by institution.

includes representative keywords such as fertility preservation and carcinoma, primarily active in recent years; #1 (endometrial cancer) contains endometrial cancer and metabolic syndrome, concentrated around 2016; #2 (gene expression) focuses on granulosa cells, estrogen receptor, and follicular fluid; #3 (cardiovascular disease) involves terms like oxidative stress, androgen, and insulin resistance, with peak activity around 2017; #4 (body composition) includes type 2 diabetes mellitus and mechanisms; #5 (polycystic ovary syndrome), as the core cluster, covers prostate cancer, obesity, and hyperandrogenism; #6 (ovarian granulosa cells) contains keywords such as management, therapy, and PCOS; #7 (ovarian reserve) addresses themes like inflammation, syndrome PCOS, and adipose tissue; #8 (breast cancer risk) incorporates postmenopausal women, receptor, endometriosis, activation, expression, and proliferation; #9 (granulosa cell) focuses on apoptosis, breast cancer, disease, and associated polycystic ovary syndrome; #10 (breast cancer) is an independent cluster; #11 (polycystic ovarian syndrome) includes polycystic ovary syndrome, associated, and disease; #12

https://doi.org/10.2147/JMDH.S536122

Table 3 Top 10 Research Institutions by Publication Volume and Citation Count

Rank	Organization	Documents	Citations	Country
1	Fudan University	34	603	China
2	University of Tehran Medical Sciences	31	757	Iran
3	Zhejiang University	28	518	China
4	Shanghai Jiao Tong University	27	574	China
5	Sichuan University	25	435	China
6	Shandong University	23	1293	China
7	Zhengzhou University	22	500	China
8	Karolinska Institute	21	794	Sweden
9	China Medical University	20	243	China
10	Harvard Medical School	20	545	USA

Table 4 Top 10 Journals by Publication Volume in PCOS and Cancer Research

Rank	Journal Name	Publications	Citations	Total Link Strength	IF (2023)	Average Citations
1	Frontiers in Endocrinology	51	1110	53	5.2	21.8
2	Gynecological Endocrinology	51	556	70	2.0	10.9
3	Journal of Ovarian Research	38	634	80	3.8	16.7
4	International Journal of Molecular Sciences	36	2467	51	4.9	68.5
5	Reproductive Sciences	32	636	80	2.6	19.9
6	Nutrients	26	1187	24	4.8	45.7
7	Human Reproduction	24	679	46	6.0	28.3
8	Journal of Clinical Endocrinology & Metabolism	22	971	73	5.0	44.1
9	Molecular and Cellular Endocrinology	19	644	56	3.8	33.9
10	PLOS ONE	19	388	29	2.9	20.4

(anti-müllerian hormone) exclusively contains anti-müllerian hormone and health; #13 (ovarian function) features cancer, gene, and identification (Figure 6B).

In terms of keyword bursts (Figure 6C), keywords with bursts initiating in 2015 include tumor necrosis factor (8.92), impaired glucose tolerance (6.07), quality of life (5.81), cardiovascular risk (5.65), in vitro fertilization (5), diabetes mellitus (4.72), breast cancer (3.96), serum (3.92), metabolic syndrome (3.92). Bursts starting in 2017 include overian reserve (4.81) and bone mineral density (4.81). Necrosis factor alpha (6.02), nfκb (4.35) and association (4.03) emerged in 2019; pathogenesis (4.1) and hormone (3.96) appeared in 2020. 2021 include mechanisms (4.66) and 2022 include apoptosis (5.38), gut microbiota (4.59) and infertility (3.94). Polycystic ovary syndrome is the most frequent keyword (863 occurrences), occupying a central position, while in vitro fertilization has the highest centrality score (1.49), indicating its role as a bridging topic across research directions. Additionally, ovarian reserve, breast, and chemotherapy exhibit high network centrality (Table 7).

Key Targets of PCOS and Endometrial Cancer

Target genes associated with PCOS and endometrial carcinoma were retrieved from the GeneCards database, yielding 876 and 623 candidate genes, respectively. A total of 250 overlapping genes were identified between PCOS and endometrial carcinoma. A PPI network was constructed using these overlapping genes, comprising 208 nodes and 1148 edges (Figure 7A). The top 20 hub targets ranked by connectivity included: TP53 (56), BRCA1 (52), SRC (42), CTNNB1 (38), ATM (36), PIK3CA (36), AKT1 (34), EGFR (33), KRAS (33), PIK3R1 (33), STAT3 (33), BRCA2 (31), RAD51 (31), HRAS (29), NRAS (28), PIK3CD (28), PIK3R2 (28), EP300 (26), MSH2 (26), and PTPN11 (26) (Figure 7B).

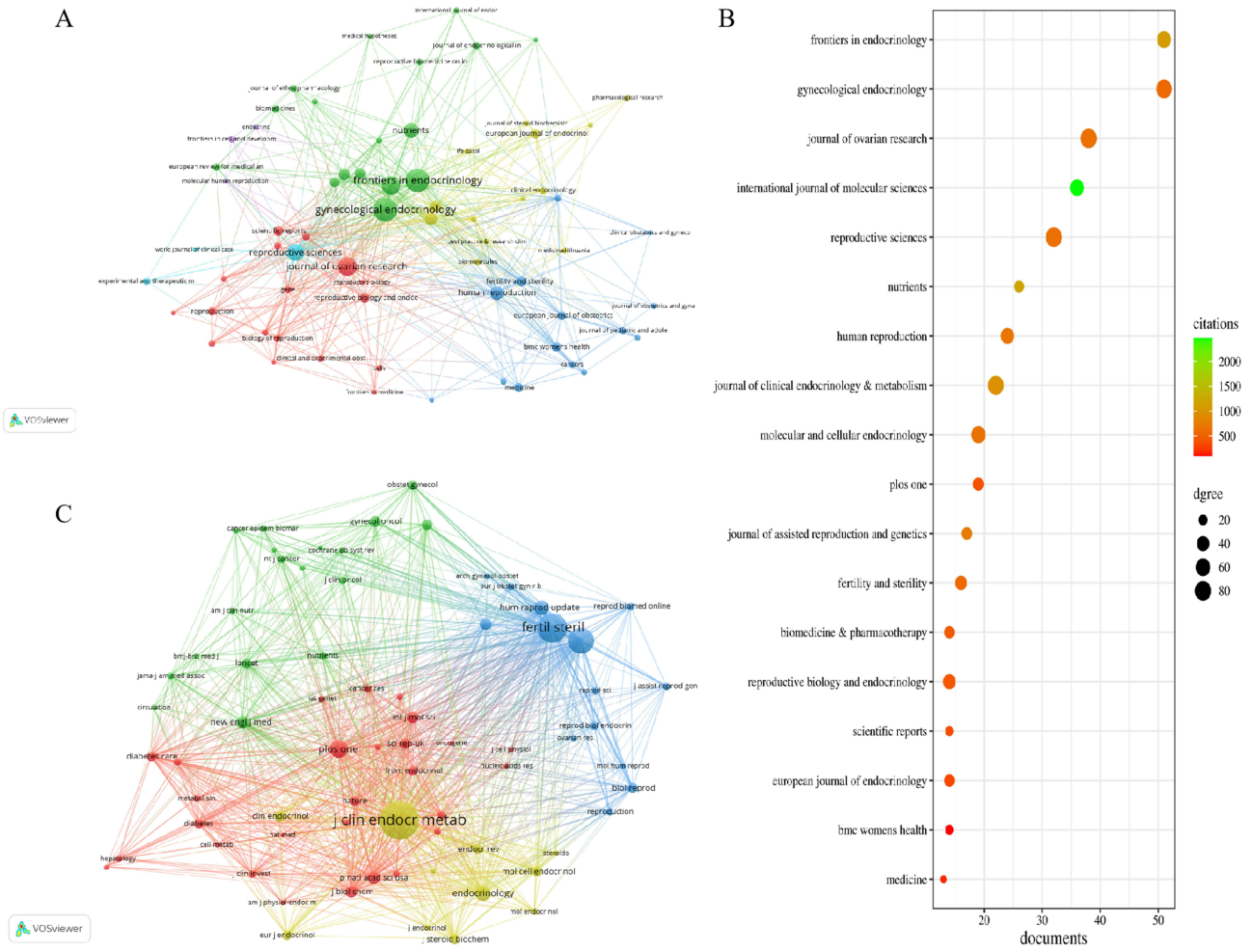

Figure 4 Journal visualization in PCOS and cancer research. (**A**) Co-citation network of source journals (**B**) Bubble chart of publication volume and citation count (**C**) Co-citation network of highly referenced journals.

Key Signaling Pathways in PCOS and Endometrial Carcinoma

KEGG enrichment analysis of overlapping genes between PCOS and endometrial carcinoma identified 160 signaling pathways. After excluding human disease-related pathways, 86 pathways remained. The top 10 pathways ranked by ascending P-value were: PI3K-Akt signaling pathway, MAPK signaling pathway, Cellular senescence, FoxO signaling pathway, Ras signaling pathway, Rap1 signaling pathway, Signaling pathways regulating pluripotency of stem cells,

Table 5 Top Cited Source Journals in PCOS and Cancer Research

Rank	Journal Name	Citations	Total Link Strength	IF (2023)	JCR Quartile
1	The Journal of Clinical Endocrinology & Metabolism	5381	283,447	5	Q1
2	Fertility and Sterility	3698	186,418	6.6	Q1
3	Human Reproduction	3080	156,986	6	Q1
4	PLOS ONE	1862	95,385	2.9	Q2
5	Endocrinology	1517	95,949	3.8	Q2
6	Human Reproduction Update	1381	74,868	14.8	Q1
7	The New England Journal of Medicine	1135	68,447	96.2	Q1
8	Proceedings of the National Academy of Sciences of the USA	1076	71,029	9.4	Q1
9	The Journal of Biological Chemistry	1065	68,759	4	Q2
10	International Journal of Molecular Sciences	1061	50,782	4.9	Q1

https://doi.org/10.2147/JMDH.S536122

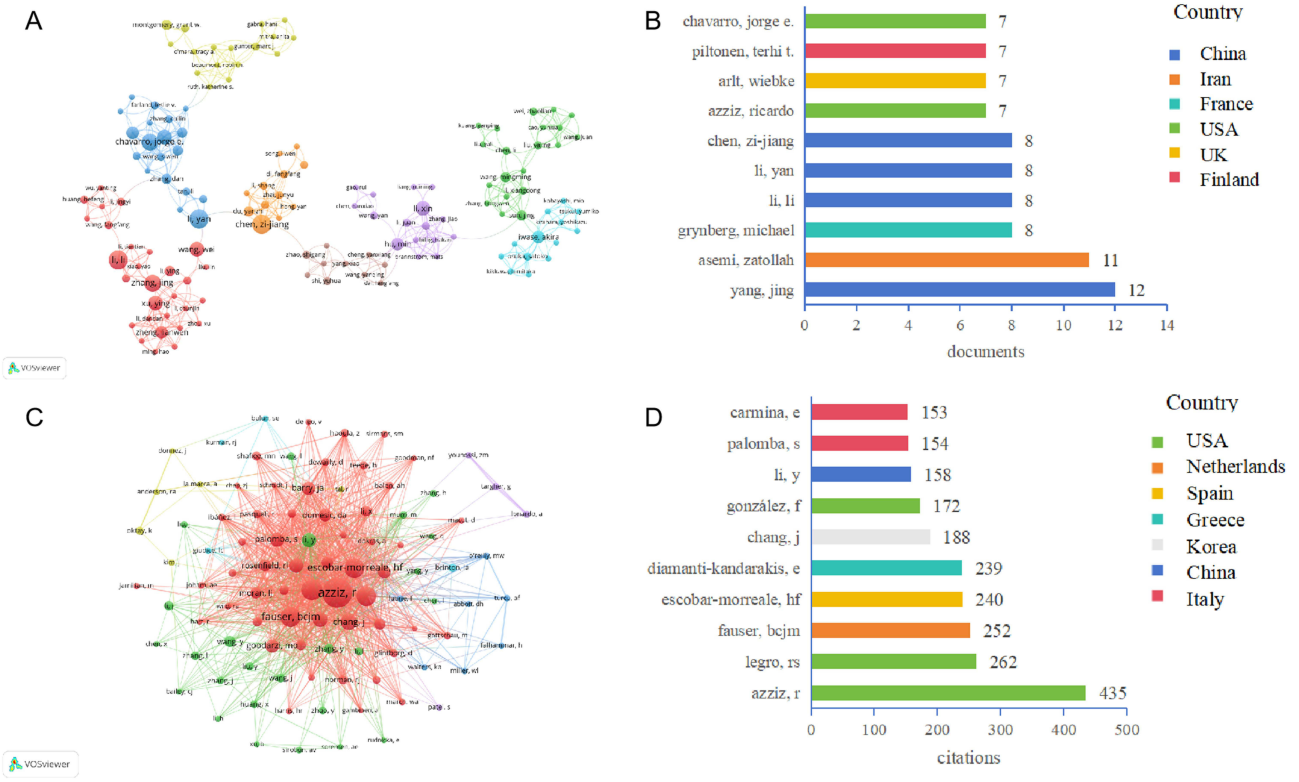

Figure 5 Author productivity and collaboration in PCOS and cancer research. (**A**) Co-authorship network of active authors. (**B**) Top 10 authors ranked by number of publications. (**C**) Collaboration clusters based on co-authorship links. (**D**) Top 10 authors ranked by total citation counts.

ErbB signaling pathway, Thyroid hormone signaling pathway, and Focal adhesion (Figure 8A). These top 10 pathways were associated with 95 genes. Based on their connectivity to the pathways, the top 20 prioritized genes included: MAPK1, AKT3, AKT2, AKT1, HRAS, MAP2K1, RAF1, MAP2K2, PIK3R2, KRAS, PIK3CA, PIK3R1, PIK3CD, NRAS, EGF, EGFR, IGF1, IGF1R, SOS1, and MYC (Figure 8B).

Discussion

In the past, research on the comorbid conditions of PCOS primarily focused on infertility,[18] type 2 diabetes,[19] obesity,[20] non-alcoholic fatty liver disease,[21] and depression.[22] The association between PCOS and cancer has received increasing attention in recent years. The long-term anovulation and progesterone deficiency in PCOS patients may be key factors contributing to the development of endometrial cancer, breast cancer, and ovarian cancer.[23] A nationwide Danish cohort

Table 6 Top 10 Most Prolific Authors in PCOS and Cancer Research

Rank	Author	Documents	Citations	Average Citations
1	Yang, Jing	12	278	23.2
2	Asemi, Zatollah	11	315	28.6
3	Chen, Zi-jiang	8	115	14.4
4	Grynberg, Michael	8	286	35.8
5	Li, Li	8	247	30.9
6	Li, Yan	8	144	18.0
7	Arlt, Wiebke	7	544	77.7
8	Azziz, Ricardo	7	1351	193.0
9	Chavarro, Jorge E.	7	266	38.0
10	Du Toit, Therina	7	219	31.3

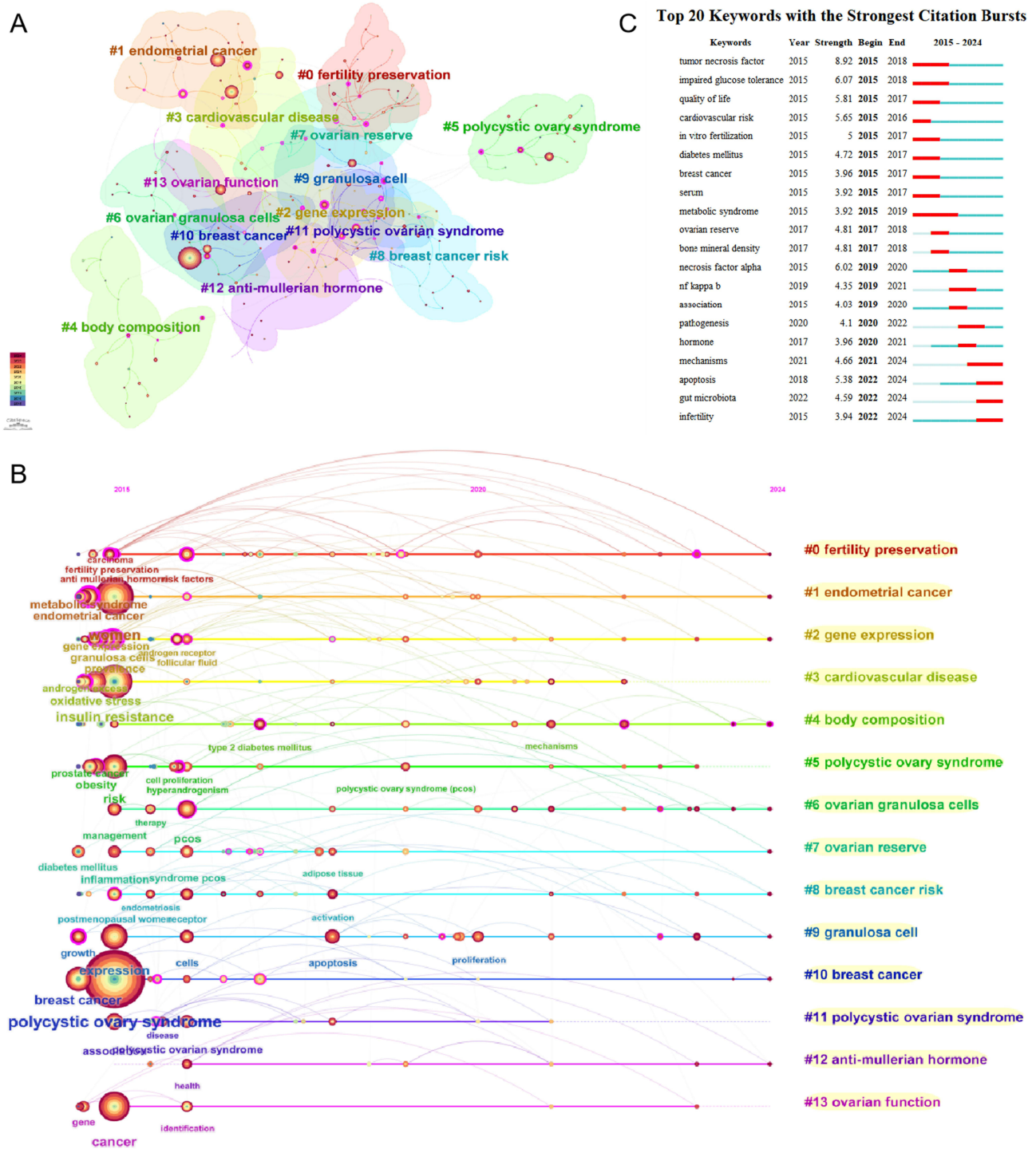

Figure 6 Keyword visualization in PCOS and cancer research. (**A**) Co-occurrence clustering of keywords. (**B**) Timeline evolution of high-frequency keywords. (**C**) Top 20 keywords with strongest citation bursts.

study, utilizing a national database and including 1,719,452 women, found that women with PCOS had a threefold higher risk of endometrial cancer compared to women without PCOS. Among these, women of reproductive age exhibited an even greater risk, reaching nearly sixfold higher than the control group.[24] Women with PCOS have an increased overall risk of breast cancer, and subgroup analyses revealed that this risk elevation was confined to postmenopausal women.[25] A study utilizing the National Inpatient Sample database revealed that women with PCOS exhibited a nearly four-fold

https://doi.org/10.2147/JMDH.S536122

Table 7 Top Keywords by Frequency and Centrality

Rank	Keyword	Count	Keyword	Centrality
1	Polycystic ovary syndrome	863	In vitro fertilization	1.49
2	Women	407	Ovarian reserve	0.95
3	Insulin resistance	301	Breast	0.83
4	Cancer	246	Chemotherapy	0.82
5	Expression	213	Fertility drugs	0.82
6	Breast cancer	195	Cohort	0.76
7	Risk	182	Cardiovascular disease	0.58
8	Oxidative stress	146	Follow up	0.58
9	Endometrial cancer	126	Estrogen	0.55
10	Metabolic syndrome	119	Assisted reproductive technology	0.54

higher risk of endometrial cancer compared to those without the condition, while no significant increase in ovarian or cervical cancer risk was observed.[26] A meta-analysis demonstrated that, after excluding postmenopausal women, individuals with PCOS had a more than five-fold increased risk of endometrial cancer compared to those without the syndrome.[27] The aforementioned studies confirm a significant association between PCOS and malignancies, particularly highlighting its strong linkage with endometrial cancer risk.

Bibliometric analysis reveals that publications investigating the association between PCOS and neoplasms exhibited a steady growth trend from 2015 to 2021, with a notable surge in 2021 followed by stabilization, indicating sustained academic attention to this field in recent years. China and the United States dominate the global publication output, with China demonstrating a rapid increase in annual publications, whereas the US maintains a stable production volume. The comparatively lower influence of Chinese academic publications relative to other nations may be associated with the nation's research evaluation system, which historically emphasized quantitative output over qualitative impact. At the institutional level, Chinese institutions account for the majority of high-yield organizations. Fudan University leads in publication volume, yet Shandong University demonstrates disproportionately higher influence within this domain. Author-level analysis further indicates that the most influential researchers in this field typically maintain moderate publication volumes. These findings collectively suggest a non-linear relationship between quantity and quality in PCOS-

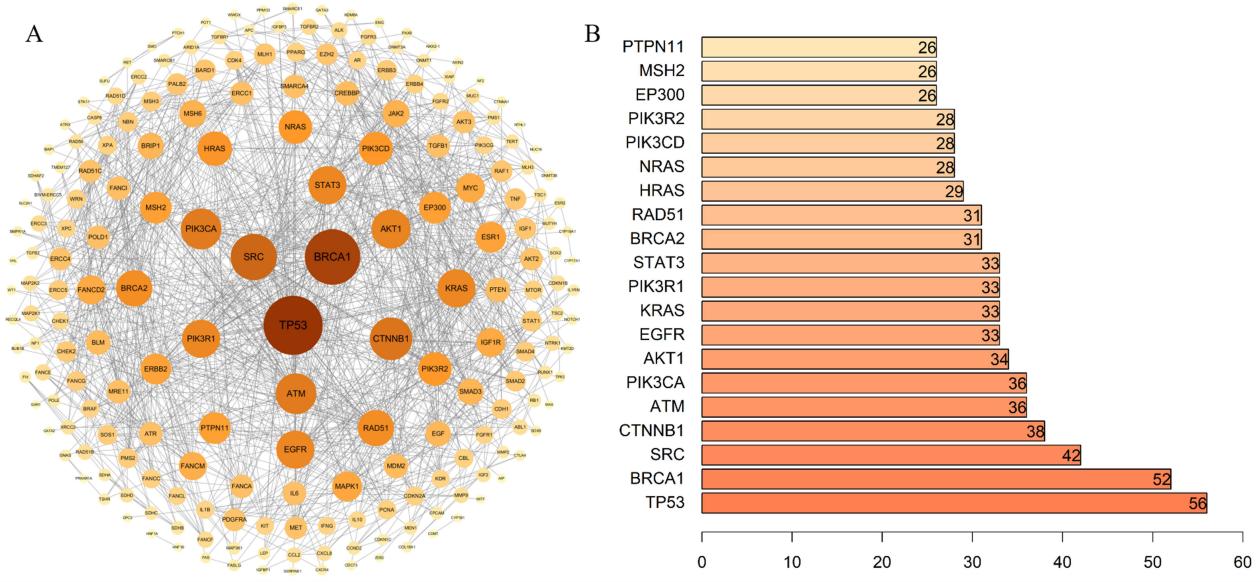

Figure 7 PPI Network Analysis of Shared Genes Between PCOS and Endometrial Carcinoma. (**A**) Visualization of the PPI network. (**B**) Top 20 hub targets ranked within the PPI network.

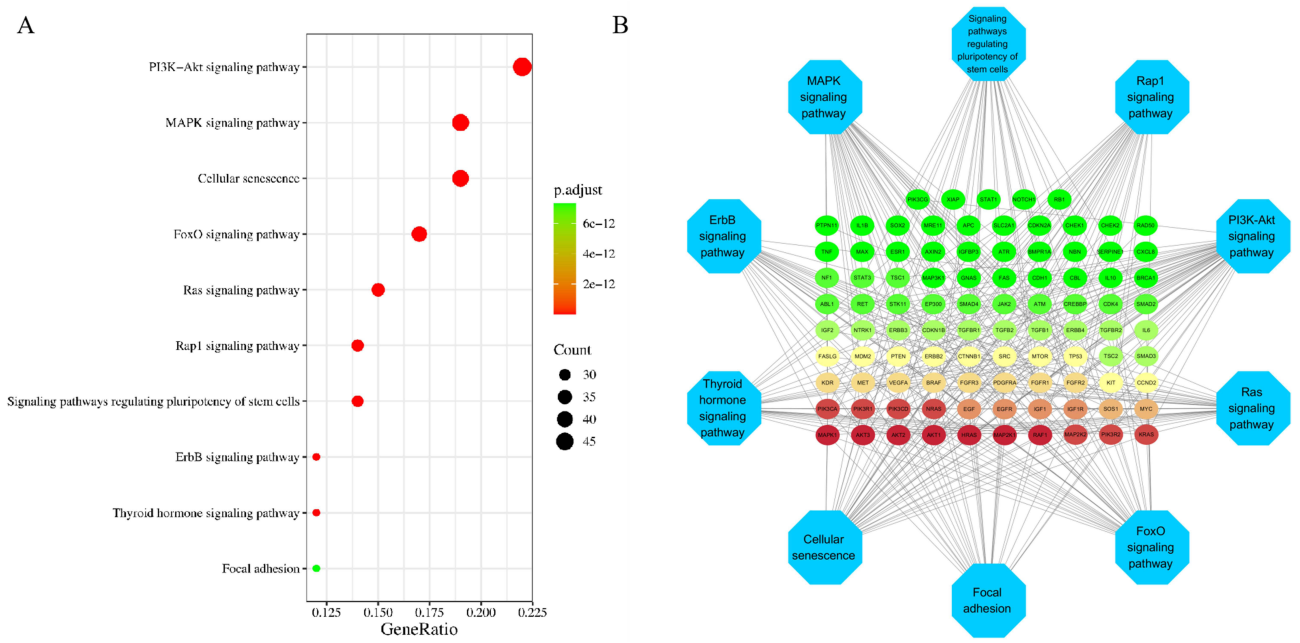

Figure 8 KEGG Enrichment Analysis of Shared Genes Between PCOS and Endometrial Carcinoma. (**A**) Top 10 enriched pathways (bubble plot). (**B**) Association between the top 10 pathways and target genes.

cancer research. High-impact institutions and authors likely prioritize research depth and innovation, strategically allocating resources to cultivate seminal studies with enduring academic value rather than pursuing indiscriminate productivity. Among journals, *Frontiers in Endocrinology* and *Gynecological Endocrinology* rank first in publication volume related to PCOS and cancer. However, their citation impact markedly lags behind *International Journal of Molecular Sciences*. Reference analysis identifies The *Journal of Clinical Endocrinology & Metabolism* as the most frequently cited source in foundational literature, underscoring its pivotal role in shaping the intellectual framework of PCOS-cancer research.

Keyword-based analysis revealed that endometrial cancer is a major research focus in the field of PCOS and cancer-related studies, garnering increasing attention. Based on this finding, we retrieved PCOS and endometrial cancer-associated gene targets from the Genecards database and identified overlapping genes as potential shared molecular targets for subsequent analysis. We then performed PPI network analysis to pinpoint key hub genes, followed by KEGG pathway enrichment analysis to elucidate critical signaling pathways. The PPI analysis identified the top 20 hub genes, including: TP53, BRCA1, SRC, CTNNB1, ATM, PIK3CA, AKT1, EGFR, KRAS, PIK3R1, STAT3, BRCA2, RAD51, HRAS, NRAS, PIK3CD, PIK3R2, EP300, MSH2, and PTPN11. Among these gene targets, TP53, BRCA1, BRCA2, ATM, MSH2, EP300 are key tumor suppressor genes, and mutations in these genes lead to abnormal cell cycle regulation, impaired DNA repair, and other cellular functions.[28–30] SRC encodes a non-receptor tyrosine kinase, which is activated by phosphorylating tyrosine residues of the corresponding target proteins to activate MAPK, STAT, PI3K/AKT, EGFR, and other tumor-associated signaling pathways.[31,32] CTNNB1 affects tumor progression by regulating the Wnt/β-catenin signaling pathway.[33] PIK3CA, AKT1, PIK3CD, PIK3R2 encode key proteins in the PI3K/AKT signaling pathway, which is critically involved in tumorigenesis.[34] The epidermal growth factor receptor (EGFR) orchestrates tissue homeostasis through regulating cellular proliferation, differentiation, migration, and apoptosis, while functioning as a central oncogenic driver that promotes tumorigenesis, disease progression, and therapeutic resistance across multiple cancer types.[35]

In the KEGG enrichment analysis, the PI3K/AKT signaling pathway was significantly enriched, indicating its central involvement in both PCOS and endometrial cancer. The PI3K/AKT signaling pathway plays a critical role in insulin resistance, hyperandrogenism, and abnormal follicular development in PCOS, participating in the entire disease progression.[36,37] The PI3K/AKT signaling pathway serves as a central regulatory hub in tumorigenesis and cancer progression by modulating critical cellular processes including proliferation, survival, metabolic reprogramming, inflammatory responses, and angiogenesis, with PI3K

https://doi.org/10.2147/JMDH.S536122

inhibitors and AKT inhibitors having advanced to clinical trial phases for anticancer therapeutic development.[38–42] Furthermore, the PI3K/AKT signaling pathway regulates key targets and pathways identified through PPI network analysis and KEGG enrichment, thereby driving tumorigenesis. Regarding the impact on key targets, the PI3K/AKT pathway phosphorylates and activates MDM2 (an E3 ubiquitin ligase), thereby promoting ubiquitination-mediated degradation of p53 and reducing its protein stability.[43] AKT phosphorylates BRCA1 at Serine 694, which inhibits its binding to BARD1, thereby impairing DNA damage repair capacity and concurrently promoting tumor cell migration, invasion, and epithelial-mesenchymal transition.[44] AKT phosphorylates β-catenin at Serine 552, preventing its degradation by GSK3β and promoting nuclear translocation and Wnt signaling activation, thereby enhancing tumor stemness and metastasis.[45,46] In terms of signaling pathway interactions, the PI3K/AKT and MAPK pathways form a complex regulatory network through upstream crosstalk, direct phosphorylation, negative feedback loops, and transcriptional synergy, collectively driving tumor proliferation, survival, metastasis, and therapy resistance.[47,48] Compensatory activation between these pathways represents a major cause of failure in targeted cancer therapies. Concurrently, phosphorylation-mediated activation of AKT induces the inactivation of FoxO proteins via phosphorylation, thereby abolishing the tumor-suppressive effects of the FoxO signaling pathway.[49–51] In summary, the PI3K/AKT signaling pathway plays a pivotal role in both PCOS and endometrial carcinogenesis.

Conclusion

This study comprehensively, scientifically, and systematically analyzed the research landscape and frontiers of PCOS and cancer based on bibliometrics, visualization analysis, and bioinformatics, while identifying potential targets and pathways linking PCOS to endometrial cancer. Our research observed that global interest in PCOS and cancer-related studies has significantly increased over the past decade, with research hotspots predominantly focused on endometrial cancer. Further bioinformatics investigations revealed that the PI3K/AKT signaling pathway may play the most critical role in the association between PCOS and endometrial cancer.

Limitations and Future Directions

This study inevitably has several limitations. First, the methodology exclusively utilized WoSCC as the sole data source, which may compromise the comprehensiveness of findings due to the exclusion of materials from major biomedical repositories such as PubMed, Scopus, and Embase. Second, the selection criteria were confined to articles and review papers, potentially overlooking valuable insights contained in alternative publication formats like conference proceedings and technical reports. Future research should prioritize international collaboration across three critical dimensions: investigating the dynamic interplay between PCOS-associated metabolic dysregulation and tumor microenvironment evolution, validating the clinical efficacy of PI3K/AKT pathway-targeted therapies, and developing multi-omics risk prediction models to refine personalized cancer surveillance protocols for PCOS populations. Such an integrative approach bridging mechanistic exploration and clinical implementation could establish novel paradigms for precision medicine in PCOS-related oncopathology.

Funding

This study was supported by the National High Level Hospital Clinical Research Funding (No. 2022-NHLHCRF-LX-02-0109).

Disclosure

The authors report no conflicts of interest in this work.

References

1. Sadeghi HM, Adeli I, Calina D, et al. Polycystic ovary syndrome: a comprehensive review of pathogenesis, management, and drug repurposing. *Int J Mol Sci*. 2022;23(2):583. doi:10.3390/ijms23020583
2. Safiri S, Noori M, Nejadghaderi SA, et al. Prevalence, incidence and years lived with disability due to polycystic ovary syndrome in 204 countries and territories, 1990–2019. *Hum Reprod*. 2022;37(8):1919–1931. doi:10.1093/humrep/deac091
3. Sanchez-Garrido MA, Tena-Sempere M. Metabolic dysfunction in polycystic ovary syndrome: pathogenic role of androgen excess and potential therapeutic strategies. *Mol Metab*. 2020;35:100937. doi:10.1016/j.molmet.2020.01.001

4. Meczekalski B, Pérez-Roncero GR, López-Baena MT, Chedraui P, Pérez-López FR. The polycystic ovary syndrome and gynecological cancer risk. *Gynecol Endocrinol*. 2020;36(4):289–293. doi:10.1080/09513590.2020.1730794
5. Zhao Y, Chen GY, Fang M. Research trends of rheumatoid arthritis and depression from 2019 to 2023: a bibliometric analysis. *J Multidiscip Healthc*. 2024;17:4465–4474. doi:10.2147/JMDH.S478748
6. Lai Y, Jiang J, Zhang H, Gong K. Bibliometric analysis of curcumin based on citespace: landscapes, hotspots, and frontiers. *Drug Des Devel Ther*. 2024;18:5743–5758. doi:10.2147/DDDT.S494758
7. Chen C. CiteSpace II: detecting and visualizing emerging trends and transient patterns in scientific literature. *J Am Soc Inf Sci*. 2006;57(3):359–377. doi:10.1002/asi.20317
8. van Eck NJ, Waltman L. Software survey: vOSviewer, a computer program for bibliometric mapping. *Scientometrics*. 2010;84(2):523–538. doi:10.1007/s11192-009-0146-3
9. Hassan-Montero Y, De-Moya-Anegón F, Guerrero-Bote V. SCImago graphica: a new tool for exploring and visually communicating data. *El Profesional de la Informacion*. 2022;35. doi:10.3145/epi.2022.sep.02
10. Stelzer G, Rosen N, Plaschkes I, et al. The genecards suite: from gene data mining to disease genome sequence analyses. *Curr Protoc Bioinformatics*. 2016;54:1.30.1–1.30.33. doi:10.1002/cpbi.5
11. Chen GY, Ji XY, Li Y, Zheng SS, Jin Q, Tao QW. Mechanisms of total glucosides of paeony in alleviating methotrexate-induced liver injury. *Drug Des Devel Ther*. 2025;19:3407–3423. doi:10.2147/DDDT.S521740
12. Tang D, Chen M, Huang X, et al. SRplot: a free online platform for data visualization and graphing. *PLoS One*. 2023;18(11):e0294236. doi:10.1371/journal.pone.0294236
13. Szklarczyk D, Gable AL, Lyon D, et al. STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res*. 2019;47(D1):D607–D613. doi:10.1093/nar/gky1131
14. Shannon P, Markiel A, Ozier O, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res*. 2003;13(11):2498–2504. doi:10.1101/gr.1239303
15. Chen GY, Liu XY, Luo J, Yu XB, Liu Y, Tao QW. Integrating network pharmacology and experimental validation to explore the key mechanism of gubitong recipe in the treatment of osteoarthritis. *Comput Math Methods Med*. 2022;2022:7858925. doi:10.1155/2022/7858925
16. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS*. 2012;16(5):284–287. doi:10.1089/omi.2011.0118
17. Chen GY, Liu XY, Yan XE, et al. Total flavonoids of rhizoma drynariae treat osteoarthritis by inhibiting arachidonic acid metabolites through AMPK/NFκB pathway. *J Inflamm Res*. 2023;16:4123–4140. doi:10.2147/JIR.S418345
18. Collée J, Mawet M, Tebache L, Nisolle M, Brichant G. Polycystic ovarian syndrome and infertility: overview and insights of the putative treatments. *Gynecol Endocrinol*. 2021;37(10):869–874. doi:10.1080/09513590.2021.1958310
19. Xu Y, Qiao J. Association of insulin resistance and elevated androgen levels with Polycystic Ovarian Syndrome (PCOS): a review of literature. *J Healthc Eng*. 2022;2022:9240569. doi:10.1155/2022/9240569
20. Hu L, Ma L, Xia X, et al. Efficacy of bariatric surgery in the treatment of women with obesity and polycystic ovary syndrome. *J Clin Endocrinol Metab*. 2022;107(8):e3217–e3229. doi:10.1210/clinem/dgac294
21. Chen Y, Ma L, Ge Z, Pan Y, Xie L. Key genes associated with non-alcoholic fatty liver disease and polycystic ovary syndrome. *Front Mol Biosci*. 2022;9:888194. doi:10.3389/fmolb.2022.888194
22. Brutocao C, Zaiem F, Alsawas M, Morrow AS, Murad MH, Javed A. Psychiatric disorders in women with polycystic ovary syndrome: a systematic review and meta-analysis. *Endocrine*. 2018;62(2):318–325. doi:10.1007/s12020-018-1692-3
23. Throwba HPK, Unnikrishnan L, Pangath M, et al. The epigenetic correlation among ovarian cancer, endometriosis and PCOS: a review. *Crit Rev Oncol Hematol*. 2022;180:103852. doi:10.1016/j.critrevonc.2022.103852
24. Frandsen CLB, Gottschau M, Nøhr B, et al. Polycystic ovary syndrome and endometrial cancer risk: results from a nationwide cohort study. *Am J Epidemiol*. 2024;193(10):1399–1406. doi:10.1093/aje/kwae061
25. Frandsen CLB, Nøhr B, Gottschau M, et al. Polycystic ovary syndrome and risk of breast cancer in premenopausal and postmenopausal women: a nationwide population-based cohort study. *Breast Cancer Res Treat*. 2024;208(3):535–542. doi:10.1007/s10549-024-07467-8
26. Abu-Zaid A, Baradwan S, Alyafi M, et al. Association between polycystic ovary syndrome and the risk of malignant gynecologic cancers (ovarian, endometrial, and cervical): a population-based study from the U.S.A. national inpatient sample 2016–2019. *Eur J Obstet Gynecol Reprod Biol*. 2024;299:283–288. doi:10.1016/j.ejogrb.2024.06.031
27. Johnson JE, Daley D, Tarta C, Stanciu PI. Risk of endometrial cancer in patients with polycystic ovarian syndrome: a meta-analysis. *Oncol Lett*. 2023;25(4):168. doi:10.3892/ol.2023.13754
28. Mansur MB, Greaves M. Convergent TP53 loss and evolvability in cancer. *BMC Ecol Evol*. 2023;23(1):54. doi:10.1186/s12862-023-02146-6
29. Oubaddou Y, Ben Ali F, Oubaqui FE, Qmichou Z, Bakri Y, Ameziane El Hassani R. The tumor suppressor BRCA1/2, cancer susceptibility and genome instability in gynecological and mammary cancers. *Asian Pac J Cancer Prev*. 2023;24(9):3139–3153. doi:10.31557/APJCP.2023.24.9.3139
30. Wang Y, Liu Y, Wang C, Kang R, Tang D, Liu J. EP300 promotes ferroptosis via HSPA5 acetylation in pancreatic cancer. *Sci Rep*. 2023;13(1):15004. doi:10.1038/s41598-023-42136-8
31. Pelaz SG, Tabernero A. Src: coordinating metabolism in cancer. *Oncogene*. 2022;41(45):4917–4928. doi:10.1038/s41388-022-02487-4
32. Xu R, Song J, Ruze R, et al. SQLE promotes pancreatic cancer growth by attenuating ER stress and activating lipid rafts-regulated Src/PI3K/Akt signaling pathway. *Cell Death Dis*. 2023;14(8):497. doi:10.1038/s41419-023-05987-7
33. Moroney MR, Woodruff E, Qamar L, et al. Inhibiting Wnt/beta-catenin in CTNNB1-mutated endometrial cancer. *Mol, Carcinog*. 2021;60(8):511–523. doi:10.1002/mc.23308
34. Glaviano A, Foo ASC, Lam HY, et al. PI3K/AKT/mTOR signaling transduction pathway and targeted therapies in cancer. *Mol Cancer*. 2023;22(1):138. doi:10.1186/s12943-023-01827-6
35. Levantini E, Maroni G, Del Re M, Tenen DG. EGFR signaling pathway as therapeutic target in human cancers. *Semin Cancer Biol*. 2022;85:253–275. doi:10.1016/j.semcancer.2022.04.002
36. Chang KJ, Chen JH, Chen KH. The pathophysiological mechanism and clinical treatment of polycystic ovary syndrome: a molecular and cellular review of the literature. *Int J Mol Sci*. 2024;25(16):9037. doi:10.3390/ijms25169037

37. Tong C, Wu Y, Zhang L, Yu Y. Insulin resistance, autophagy and apoptosis in patients with polycystic ovary syndrome: association with PI3K signaling pathway. *Front Endocrinol*. 2022;13:1091147. doi:10.3389/fendo.2022.1091147
38. Yu L, Wei J, Liu P. Attacking the PI3K/Akt/mTOR signaling pathway for targeted therapeutic treatment in human cancer. *Semin Cancer Biol*. 2022;85:69–94. doi:10.1016/j.semcancer.2021.06.019
39. He Y, Sun MM, Zhang GG, et al. Targeting PI3K/Akt signal transduction for cancer therapy. *Signal Transduct Target Ther*. 2021;6(1):425. doi:10.1038/s41392-021-00828-5
40. Ediriweera MK, Tennekoon KH, Samarakoon SR. Role of the PI3K/AKT/mTOR signaling pathway in ovarian cancer: biological and therapeutic significance. *Semin Cancer Biol*. 2019;59:147–160. doi:10.1016/j.semcancer.2019.05.012
41. Chen GY, Liu XY, Chen JQ, et al. Prediction of rhizoma drynariae targets in the treatment of osteoarthritis based on network pharmacology and experimental verification. *Evid Based Complement Alternat Med*. 2021;2021:5233462. doi:10.1155/2021/5233462
42. Yap TA, Kristeleit R, Michalarea V, et al. Phase I trial of the PARP inhibitor olaparib and AKT inhibitor capivasertib in patients with BRCA1/2- and Non-BRCA1/2-mutant cancers. *Cancer Discov*. 2020;10(10):1528–1543. doi:10.1158/2159-8290.CD-20-0163
43. Zhang P, Guo H, Zhao F, Jia K, Yang F, Liu X. UBE2J1 knockdown promotes cell apoptosis in endometrial cancer via regulating PI3K/AKT and MDM2/p53 signaling. *Open Med*. 2023;18(1):20220567. doi:10.1515/med-2022-0567
44. Nelson AC, Lyons TR, Young CD, Hansen KC, Anderson SM, Holt JT. AKT regulates BRCA1 stability in response to hormone signaling. *Mol Cell Endocrinol*. 2010;319(1–2):129–142. doi:10.1016/j.mce.2010.01.019
45. Jere SW, Abrahamse H, Houreld NN. Interaction of the AKT and β-catenin signalling pathways and the influence of photobiomodulation on cellular signalling proteins in diabetic wound healing. *J Biomed Sci*. 2023;30(1):81. doi:10.1186/s12929-023-00974-8
46. Fang D, Hawke D, Zheng Y, et al. Phosphorylation of beta-catenin by AKT promotes beta-catenin transcriptional activity. *J Biol Chem*. 2007;282(15):11221–11229. doi:10.1074/jbc.M611871200
47. Morgos DT, Stefani C, Miricescu D, et al. Targeting PI3K/AKT/mTOR and MAPK Signaling Pathways in Gastric Cancer. *Int J Mol Sci*. 2024;25(3):1848. doi:10.3390/ijms25031848
48. Shorning BY, Dass MS, Smalley MJ, Pearson HB. The PI3K-AKT-mTOR pathway and prostate cancer: at the crossroads of AR, MAPK, and WNT signaling. *Int J Mol Sci*. 2020;21(12):4507. doi:10.3390/ijms21124507
49. Farhan M, Wang H, Gaur U, Little PJ, Xu J, Zheng W. FOXO signaling pathways as therapeutic targets in cancer. *Int J Biol Sci*. 2017;13(7):815–827. doi:10.7150/ijbs.20052
50. Tzivion G, Dobson M, Ramakrishnan G. FoxO transcription factors; Regulation by AKT and 14-3-3 proteins. *Biochim Biophys Acta*. 2011;1813(11):1938–1945. doi:10.1016/j.bbamcr.2011.06.002
51. Jiramongkol Y, Lam EW. FOXO transcription factor family in cancer and metastasis. *Cancer Metastasis Rev*. 2020;39(3):681–709. doi:10.1007/s10555-020-09883-w