Journal of Digestive Diseases and Hepatology
Chang CS, et al. J Dig Dis Hepatol 10: 237.
www.doi.org/10.29011/2574-3511.100237
www.gavinpublishers.com

OPEN ACCESS

GAVIN PUBLISHERS

Research Article

Scopus Cluster: A Web-Based Platform for Bibliometric Analysis of Scopus-Indexed Publications, with a Structural-hole analysis to the 2025 Top 2% Scientists in Gastroenterology and Hepatology

Chi-Sheng Chang[1,2], Lifan Chen[3], Tsair-Wei Chien[4], Kun-Feng Tsai[5*]

[1]Quality Management Center, Linyin Chi-Mei Hospital, Tainan, Taiwan

[2]Department of Nursing, Min-Hwei College of Health Care Management, Tainan, Taiwan

[3]Department of Nursing, An Nan Hospital, China Medical University, Tainan, Taiwan

[4]Department of Medical Research, Chi Mei Medical Center, Taina, Taiwan

[5]Gastroenterology and Hepatology Section, Department of Internal Medicine, An Nan Hospital, China Medical University, Tainan, Taiwan.

***Corresponding Author:** Kun-Feng Tsai, Gastroenterology and Hepatology Section, Department of Internal Medicine, An Nan Hospital, China Medical University, Tainan, No. 66, Sec. 2, Changhe Rd., Zhounan Village, Annan District, Tainan City, Taiwan

Citation: Chang CS, Chen L, Chien TW, Tsai KF (2026) Scopus Cluster: A Web-Based Platform for Bibliometric Analysis of Scopus-Indexed Publications, with a Structural-hole analysis to the 2025 Top 2% Scientists in Gastroenterology and Hepatology. J Dig Dis Hepatol 11: 237. DOI: 10.29011/2574-3511.100237

Received Date: 05 September 2026; **Accepted Date:** 05 October 2026; **Published Date:** 08 October 2026.

Abstract

Background: Scopus-based bibliometric analysis often requires separate procedures for metadata preparation, network analysis, and visualization. Scopus Cluster is a browser-based platform that integrates these tasks and provides reporting of cluster structure, leadership concentration, and structural-hole positions. **Methods:** Scopus Cluster was applied to Scopus-indexed publications associated with 2,152 scientists in the 2025 Top 2% Scientists dataset for Gastroenterology and Hepatology, representing 54 countries and 1,220 institutions. Uploaded Scopus CSV meta data were transformed into automated summaries and more than 80 visualizations across authors, countries, institutions, departments, keywords, journals, publication years, document types, and cited references. Following-leading clustering algorithm (FLCA) networks were assessed using silhouette scores (SS) for local cluster separation, normalized modularity (Q) for global community structure, AAC for leadership concentration, and structural-hole measures of degree, effective size, and Burt constraint. **Results:** Scopus Cluster generated reports, forest plots, choropleth maps, Kano plots, slope plots, SS plots, Sankey diagrams, chord plots, and network visualizations within a unified workflow. Dr. Younossi showed substantial coauthor leadership concentration (AAC=0.93), whereas institutional collaboration was moderately concentrated (AAC=0.68). In the displayed first/corresponding-author network, Younossi also had the strongest structural-hole position, with the highest degree and effective size and the lowest Burt constraint. A composite structural-hole coefficient combined normalized degree, effective size, and inverse constraint; when all three components equaled 1, the composite AAC was 0.91. The outputs revealed geographic, institutional, thematic, temporal, and brokerage patterns not apparent from publication counts alone. **Conclusions:** Scopus Cluster provides an accessible and reproducible workflow for Scopus-based bibliometric analysis. By integrating FLCA, SS, Q, AAC, and structural-hole reporting, it complements conventional bibliometric software with transparent assessment of cluster quality, collaboration concentration, and network brokerage. Future studies should evaluate metadata quality, longitudinal change, and cross-database compatibility.

Citation: Chang CS, Chen L, Chien TW, Tsai KF (2026) Scopus Cluster: A Web-Based Platform for Bibliometric Analysis of Scopus-Indexed Publications, with a Structural-hole analysis to the 2025 Top 2% Scientists in Gastroenterology and Hepatology. J Dig Dis Hepatol 11: 237. DOI: 10.29011/2574-3511.100237

Keywords: Absolute Advantage Coefficient; Bibliometric Analysis; Following-leading clustering Algorithm; Scopus; Structural Hole

List of abbreviations

AAC : Absolute Advantage Coefficient

CSV : Comma-Separated Values

FLCA: Following-Leading Clustering Algorithm

PubMed: Public/Publisher MEDLINE

Q* : Normalized Modularity

Scopus : SciVerse Scopus database

SS : Silhouette Score

WoSCC: Web of Science Core Collection

Highlights

1. ScopusCluster converts Scopus CSV metadata into automated bibliometric reports and more than 80 visualizations.

2. FLCA, silhouette scores, normalized modularity, and AAC jointly assess network structure and leadership concentration.

3. The 2025 Top 2% gastroenterology and hepatology application demonstrates reproducible geographic, institutional, thematic, and scholar-level mapping.

Introduction

Bibliometric analysis is widely used to evaluate research productivity, identify emerging topics, map knowledge structures, and characterize scholarly impact across scientific disciplines [1-3]. The annually updated *Standardized Citation Indicators Database for All Scientists*, developed by Ioannidis and colleagues from Scopus data, provides a particularly valuable resource for this purpose [4-6]. It classifies scientists across 22 scientific fields and 176 subfields and reports standardized citation indicators, including total citations, the *h*-index, and a composite citation indicator (*c*-score), enabling cross-disciplinary comparison of career-long and single-year research impact [4-8].

Nevertheless, ranked lists alone provide only a limited view of scholarly profiles [9-10]. Citation-based indicators are essential for identifying influential scientists[10], but may favor longer careers and do not directly show authorship roles, collaboration structures, institutional connections, geographic participation, research topics, or journal patterns [11-13].

These contextual features are especially relevant when examining scientists in a defined specialty, such as gastroenterology and hepatology, where research output often arises from multidisciplinary and international collaborations [9]. Accordingly, there is a need to complement rankings with transparent visual and quantitative descriptions of the bibliographic metadata underlying researchers' Scopus-indexed publications.

Existing tools and remaining workflow challenges

CiteSpace, VOSviewer, App4WoSCCinR, and bibliometrix are well-established tools for performance analysis, science mapping, network construction, and bibliometric visualization [14–17]. Their widespread use demonstrates the maturity and value of the current bibliometric software ecosystem. However, practical bibliometric analysis of Scopus-indexed publications often remains fragmented. Users may need to export Scopus records, clean and standardize fields, restructure multi-value metadata, select analytic settings, construct networks, and generate publication-ready Figures across several software environments [16, 18-20].

This workflow can be challenging when investigators wish to analyze common Scopus CSV exports directly(e.g., the 2025 Top 2% Dermatology Scientists [9]). Author names, affiliations, countries, departments, keywords, references, source titles, and document types frequently require normalization before meaningful co-occurrence or profile analysis can be performed. Such preprocessing can be time-consuming, susceptible to inconsistency, and difficult to reproduce, particularly for users without programming expertise [16]. Although existing tools offer powerful network displays, users may still need additional procedures to obtain standardized summary reports and diverse visual outputs from the same uploaded dataset [21].

Need for interpretable network and leadership reporting

Conventional bibliometric reports commonly emphasize counts, rankings, tables, and network diagrams [1-3,18-20]. However, a network visualization alone does not necessarily indicate whether its detected clusters are well separated, whether the overall community structure is strong, or whether contributions are concentrated around a small number of highly prominent entities [9, 22-25]. These dimensions should be quantified to improve the interpretation of bibliometric maps.

Three complementary indicators are useful in this context. First, silhouette scores (SS) assess local cluster separation by comparing within-cluster cohesion with between-cluster distinction [22-24]. Second, normalized modularity ($Q = Q\text{w}/Q\text{max}$) can describe the strength of global community structure within a network [25,26]. Third, the Absolute Advantage Coefficient (AAC) quantifies dominance by comparing the leading entities with their nearest competitors [9, 27]. With them, AAC with a cutting point at 0.70 is particularly useful for distinguishing a sharply dominant leading author, institution, country, or keyword from a more evenly

distributed pattern. These indicators do not replace established bibliometric measures; rather, they add interpretable information concerning clustering quality, community organization, and leadership concentration.

Scopus Cluster and illustrative application

To address these workflow and interpretation needs, we developed Scopus Cluster [28], a browser-based platform for bibliometric analysis of Scopus-indexed publications. The platform accepts Scopus CSV metadata, automatically maps key bibliographic entities, constructs co-occurrence networks, and produces structured summary reports and exportable visualizations in a unified environment. Its functions cover first- and corresponding-author entities, countries, regions, institutes, departments, author keywords, index keywords, journals, publication years, document types, and cited references.

Scopus Cluster incorporates FLCA-based network analysis [25,29] and reporting tools, including network plots, reduced networks, Sankey diagrams, chord plots, Kano plots, SS plots, slope plots, choropleth maps, and AAC-enhanced summary tables [16,21,23-25]. The platform is intended as a practical complement to established bibliometric software by reducing the need for repeated manual preprocessing and by integrating analytic outputs within a reproducible upload-and-run workflow.

Study goals

Using Scopus-indexed publications associated with the 2025 Top 2% Scientists in Gastroenterology and Hepatology as an illustrative application, this study had three goals:

1. To demonstrate the workflow integration and usability of Scopus Cluster [28] for transforming a Scopus CSV export into automated metadata summaries and bibliometric visualizations.

2. To characterize local and global bibliometric network structure using FLCA-based co-occurrence analyses [25,29], SS for local cluster separation, and normalized modularity (Q) for global community organization [22-24,26].

3. To assess leadership concentration across bibliographic entities including authors, countries, institutions, departments, keywords, and journals using AAC-enhanced reporting, thereby complementing simple rankings and publication counts with interpretable dominance patterns [9, 26].

Methods

Data Source

The 2025 Stanford/Ioannidis Top 2% Scientists dataset [13] in Gastroenterology and Hepatology was downloaded, and scholars affiliated with their institutions were identified. A total of 2,152 scholars from 54 countries and 1,220 institutions (i.e., University: 711; Hospital: 460; and Center/Centre: 551) were analyzed. Study data and instruction for drawing visualizations are deposited in Appendices 1, 2, and 3.

This study did not involve human participants or medical records and was exempt from institutional ethics review.

Network-quality and leadership metrics

To address the second and third study goals, we applied Absolute Advantage Coefficient (AAC) to FLCA-based co-occurrence networks. AAC quantifies leadership concentration by comparing the leading entity with the next two entities [9,27,30].

Together, these measures complement the automated Scopus Cluster visualizations by reporting dominance patterns (AAC) across authors, countries, institutions, departments, keywords, and journals. An AAC $\geq$0.70 was interpreted as substantial leadership concentration [9,30].

Analysis and visualization

Scopus Cluster [28] was used to generate the study visualizations. Forest plots were produced with Demo2, whereas the summary report, choropleth maps, Kano plot, and slope plots were produced with Demo1 (Figure 1) of Forest plot.

The platform automatically generates more than 80 visualizations from uploaded Scopus-indexed publication metadata with Demo1 (Figure 1) of Bibliometrics.

AAC values for multiple authors were calculated only after all matching PubMed, records had been retrieved using Demo3 (Figure 1).

Detailed procedures and plotting specifications are provided in Appendix 3**,** accompanying code repository in Github [31].

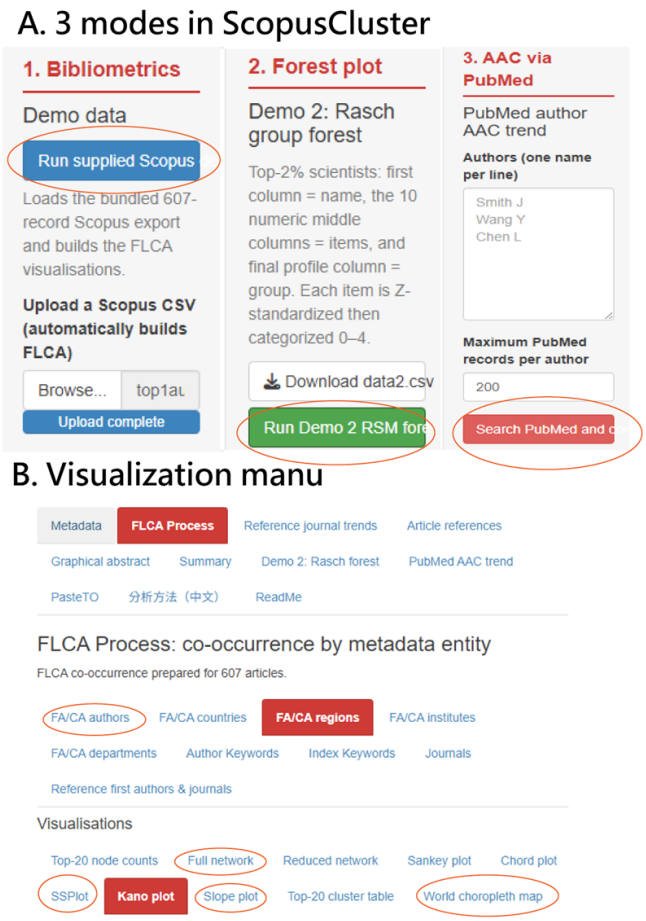


Figure 1: Homepage of Scopus Cluster [29] with 3 modes to create visualizations.

Structural-hole analysis

Composite AAC for structural-role node in computation

Structural-hole analysis was conducted using the unreduced direct co-word network. Each node represented an author (or another selected bibliometric entity), and each weighted undirected edge represented a direct co-occurrence relationship. Edge weight was defined as the number of records in which two nodes co-occurred.

Three node-level structural-hole indicators were calculated: degree (D), Burt's constraint (C), and effective size (ES). Degree represents the number of distinct direct connections. Burt's constraint represents the redundancy among a node's relationships; lower values indicate that the node's contacts are less mutually redundant and that the node has greater bridging potential. Effective size represents the number of nonredundant contacts in the ego network; higher values indicate broader access to distinct network segments.

For comparability in visualization, the three indicators were normalized to the 0–1 range across connected nodes:

$$D_i^* = (D_i - \min(D)) / (\max(D) - \min(D)) \quad (1)$$

$$C_i^* = (C_i - \min(C)) / (\max(C) - \min(C)) \quad (2)$$

$$ES_i^* = (ES_i - \min(ES)) / (\max(ES) - \min(ES)) \quad (3)$$

Because lower Burt's constraint indicates stronger structural-hole potential, its direction was reversed for interpretation:

$$RC_i^* = 1 - C_i^* \quad (4)$$

AAC was then calculated separately for normalized degree, reversed normalized constraint, and normalized effective size. For each indicator, the three largest values were denoted by v_1, v_2, and v_3, where $v_1 \geq v_2 \geq v_3$. The dominance ratio was calculated as:

$$r = (v_1 / v_2) / (v_2 / v_3) \quad (5)$$

and the AAC was calculated as:

$$AAC = r / (1 + r) \quad (6)$$

Accordingly, three AAC values were obtained: AAC^{D}, AAC^{Re}, and AAC^{Es}. The overall structural-hole dominance index named composite AAC, denoted as AACa, was defined as their arithmetic mean:

$$AACa = (AAC^{D} + AAC^{Re} + AAC^{Es}) / 3 \quad (7)$$

Interpretation of AACa and the 3D plot

A higher AACa indicates that the leading node shows greater concentration or dominance relative to the second- and third-ranked nodes across degree, low-constraint bridging potential, and effective size. Thus, AACa describes the prominence of the leading structural-hole position in the network, rather than a node-specific Burt coefficient.

A three-dimensional plot was generated with normalized degree (D*) on the horizontal axis, normalized Burt's constraint (C*) on the depth axis, and normalized effective size (ES*) on the vertical axis. Point size represented total co-word strength. Nodes located toward the right and upper regions, while positioned toward the low-constraint side of the depth axis, demonstrate a more favorable structural-hole profile: extensive direct reach, greater effective size, and lower relational redundancy. In contrast, a node with high degree but high Burt's constraint may be well connected but embedded in a redundant neighborhood and therefore may not function as an effective bridge.

Importantly, AACa is a network-level dominance measure derived from the top three nodes in each structural-hole dimension. It should therefore be reported as a summary index for the network

and should not be used as the color of individual nodes in the 3D plot.

Statistics and tools used in this study

PubMed publications for the top 100 scientists were retrieved and analyzed using Scopus Cluster [28]. That is to select the mode of Series author AAC mode: treat each pasted name as an [Author] query.

When checked, paste one author per line in the search box. If the AMA/PubMed textarea option above is checked, the app uses that textarea instead and can extract author names from full references. AAC is computed only when PubMed hits are fully fetched (pubmed_hits <= fetched_pmids).

Co-occurrence networks were evaluated using AAC to quantify the centralization of mentoring junior researchers [9,30]. With Scopus Cluster [29] Demo 2, the 11 performance indicators were standardized and linearly transformed into five ordinal categories (0–4). A Rasch Rating Scale Model [32] was then fitted to estimate individual scholar and group measures; forest plots display each group estimate with its uncertainty interval and compare it with a selected individual scholar.

Results

Forest plot comparison of countries and individual scholars

(Figure 2) presents forest plots comparing country-level academic-achievement estimates with those of individual scholars. Point estimates and their confidence intervals allow rapid assessment of an individual scholar's position relative to the country groups.

Dr. Younossi ranked Top 1 in Gastroenterology and Hepatology is highlighted at the bottom of Figure 1.

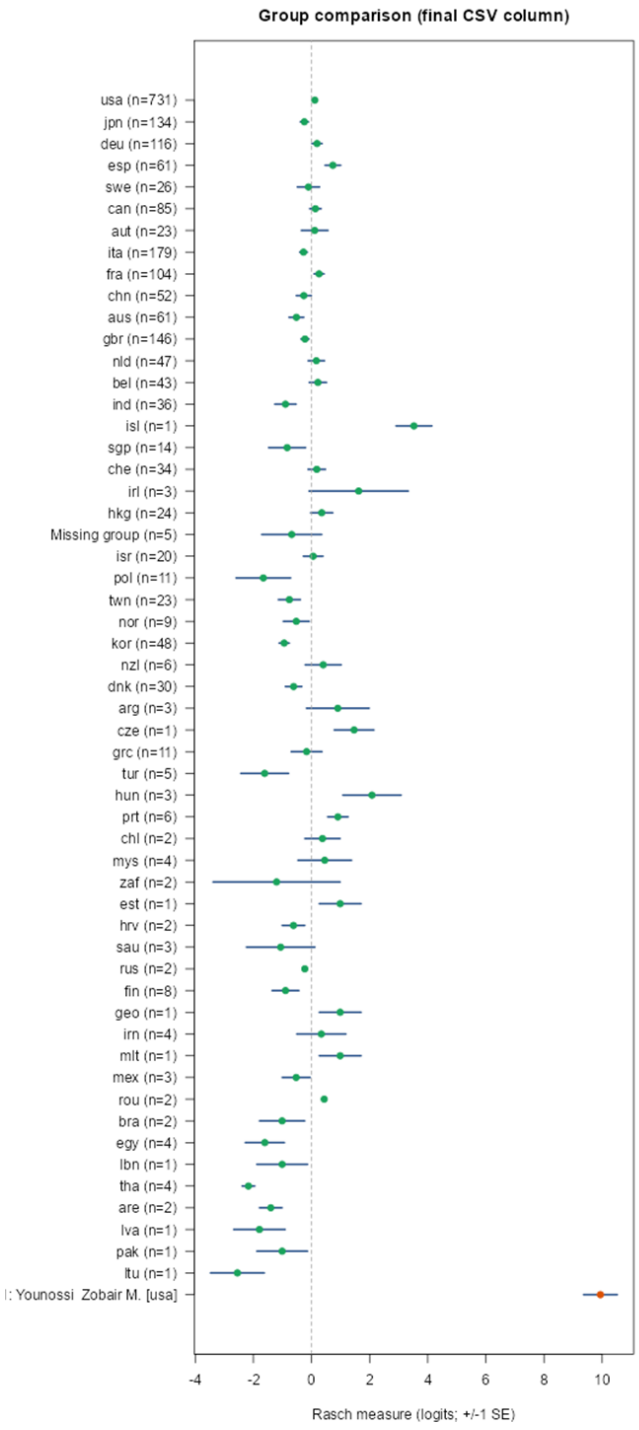


Figure 2 : Forest plots used to compare academic achievements between country groups and individual scientists using Demo 2 in Figure 1.

Bibliometric summary report

(Figure 3) provides a concise profile of Dr. Younossi's Scopus-indexed publications (n=607 and h-index=104). The report summarizes the leading first/corresponding authors, countries, institutions, departments, regions, journals, keywords, publication years, and document types, together with AAC values indicating the concentration of the leading entities.

Summary Report (n=607, h=104)

FA/CA authors	h	n
Younossi, Zabair M.	102	518
Stepanova, Maria	23	33
Golabi, Pegah	22	28
Paik, James M.	15	24
Baranova, Ancha	12	17
	AAC = 0.93	

FA/CA countries	h	n
United States	103	563
China	7	12
Saudi Arabia	7	10
United Kingdom	4	7
Japan	2	4
	AAC = 0.98	

FA/CA institutes	h	n
Inova Fairfax Hospital	84	331
George Mason University	27	47
College of Science	10	14
Translational Research Institut	10	12
School of Systems Biology	8	12
	AAC = 0.68	

FA/CA departments	h	n
Medicine	81	336
Gastroenterology	6	11
Internal Medicine	3	4
Medical Sciences	2	3
Medicine and Betty and Guy Beat	2	2
	AAC = 0.92	

FA/CA regions	h	n
United States: Washington	31	114
United States: Virginia	10	12
United States: District of Colu	4	7
United States: Florida	4	4
China: Shanghai	2	3
	AAC = 0.85	

Journal	h	n
Alimentary Pharmacology and The	29	50
Hepatology	29	42
Clinical Gastroenterology and H	27	40
Liver International	21	35
Journal of Hepatology	19	31
	AAC = 0.53	

Author keywords	h	n
NAFLD	36	66
NASH	34	51
HCV	21	29
MASLD	15	24
Obesity	19	23
	AAC = 0.42	

Index keywords	h	n
nonalcoholic fatty liver	90	301
liver cirrhosis	65	217
Non-alcoholic Fatty Liver Disea	63	159
obesity	56	158
hepatitis C	49	148
	AAC = 0.50	

Year	h	n
2019	26	49
2016	28	47
2018	23	46
2023	17	38
2025	14	35
	AAC = 0.51	

Document type	h	n
Article	90	409
Review	46	101
Letter	6	26
Editorial	7	24
Note	5	11
	AAC = 0.51	

Figure 3: Summary reports of Dr. Younossi created by tab of Summary via Demo 2 in Figure 1.

Geographic distribution of author collaborations

(Figure 4) displays choropleth maps of author collaborations for Dr. Younossi's Scopus-indexed publications (n=607) at global, United States state, and China provincial levels. Darker shading represents a greater number of Scopus-indexed articles, enabling visualization of the geographic concentration of collaboration.

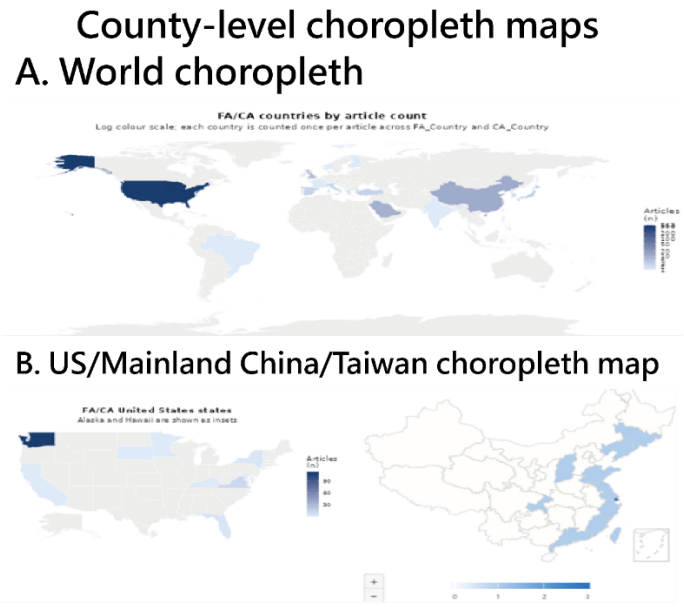


Figure 4 : Choropleth maps of author collaborations for Dr. Younossi.

Coauthor collaboration and leadership concentration

(Figure 5) shows the author co-occurrence network using a Kano plot. Dr. Younossi was the dominant author node, with an AAC of 0.93, indicating substantial concentration of collaboration around the leading investigator.

Figure 5 Author collaborations for Dr. Younossi using Kano plot(AAC=0.93).

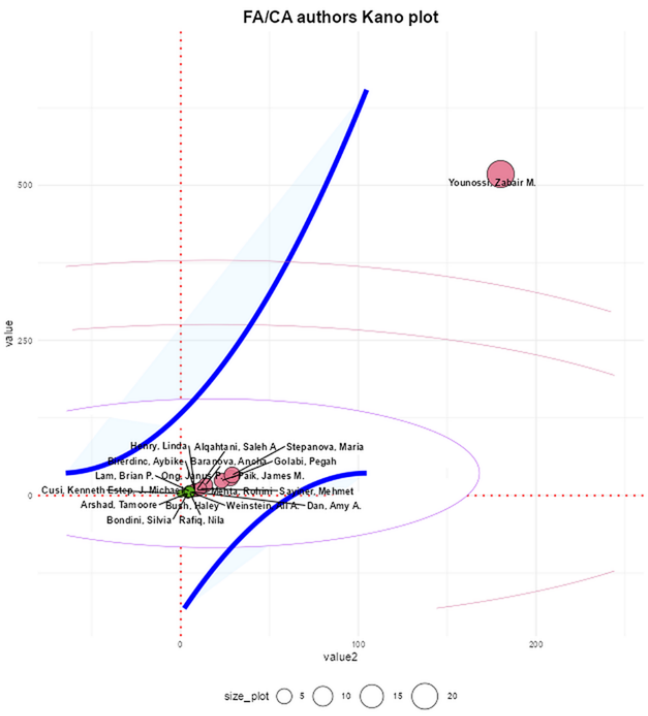


Figure 5 : Author collaborations for Dr. Younossi using Kano plot(AAC=0.93).

AAC for Top 100 scientists in Gastroenterology and Hepatology

(Figure 6) presents AAC distribution for the top 100 gastroenterology and hepatology scientists based on PubMed-indexed publications. The orange line shows each scientist's AAC, calculated using App4PubMedinR. The blue dotted line denotes AAC = 0.70, indicating substantial leadership concentration. Higher AAC values indicate that publication contributions are more strongly concentrated around the focal scientist. We found that not all of the top 100 scientists had author-collaboration patterns with substantial leadership concentration (AAC >0.70).

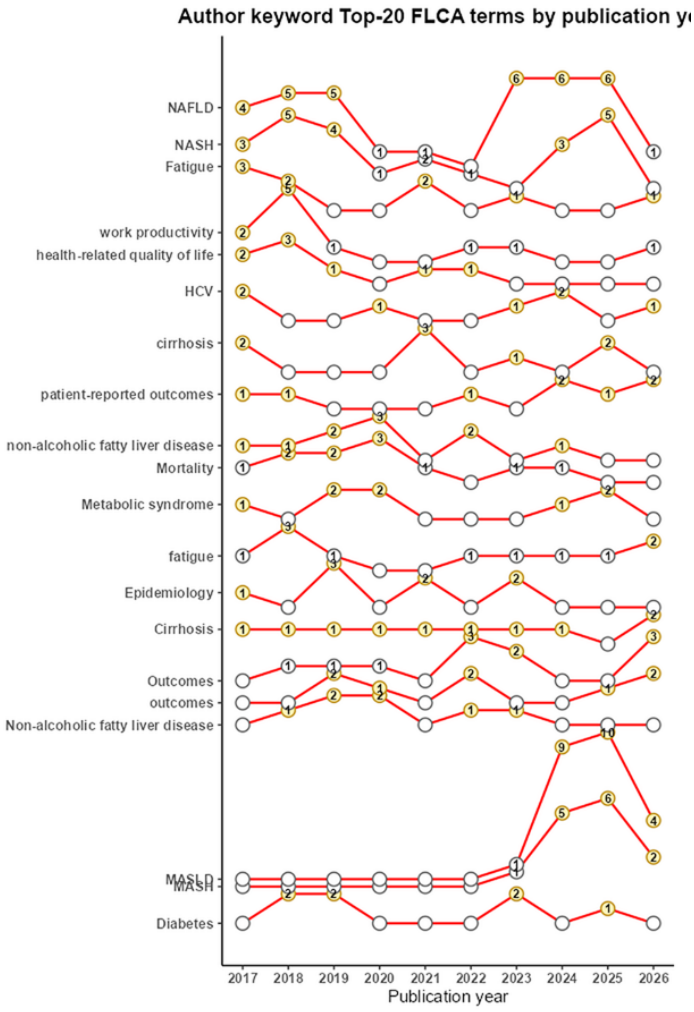


Figure 6: Trend analysis of author-defined keywords using slope plot (yellow circles denotes values >=its mean).

Structural-hole analysis

For the displayed FA/CA-author network, Younossi had the strongest structural-hole position, with the composite AAC was 0.91, because this node showed the highest Degree and effective size and the lowest Burt constraint [33,34] (Figure 7).

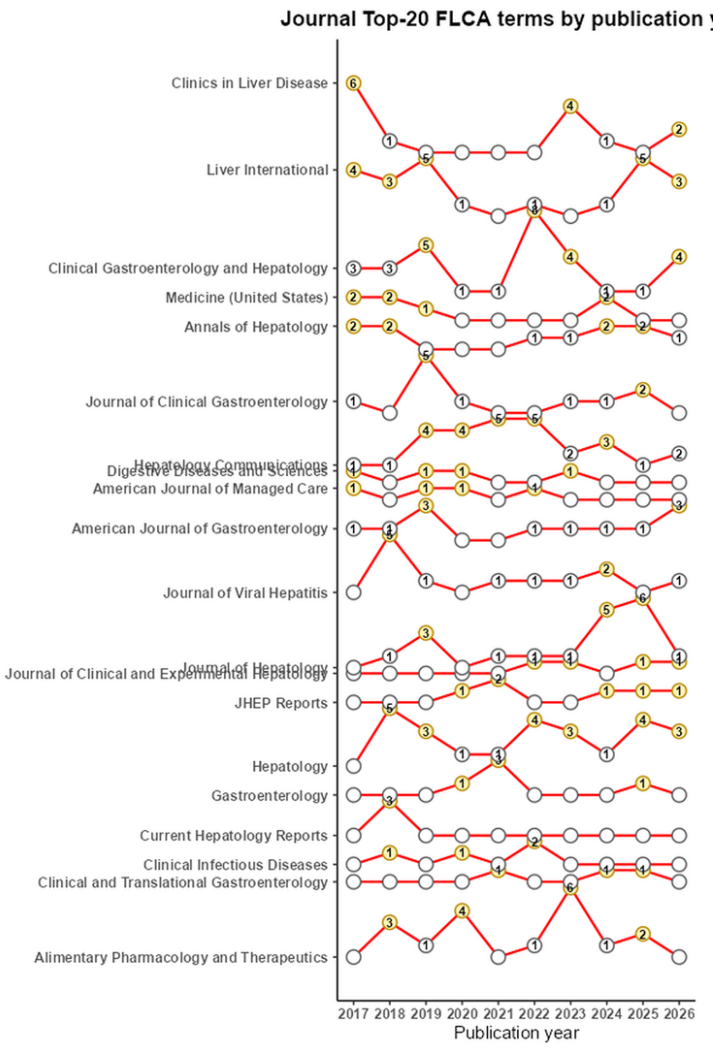


Figure 7: Trend analysis of index keywords by Scopus using slope plot (yellow circles denotes values >=its mean).

Discussions

Principal Findings

All three study goals were achieved. Scopus Cluster [29] transformed Scopus metadata for 2,152 2025 Top 2% gastroenterology and hepatology scientists from 54 countries and 1,220 institutions into automated summaries and more than 80 visualizations. FLCA-based analyses, including network, choropleth, slope, and SS plots, characterized collaboration patterns, local cluster separation, and global community structure. AAC-enhanced reporting identified leadership concentration across bibliographic entities: Dr. Younossi's coauthor network showed strong dominance (AAC=0.93), whereas institutional collaboration was moderately concentrated (AAC=0.68). A composite structural-hole coefficient combined normalized degree, effective size, and inverse constraint; when all three components equaled 1, the composite AAC was 0.91.

Interpretation of the network patterns created by Scopus Cluster

Scopus Cluster [28] extends Scopus-based bibliometric analysis from raw CSV metadata to automated summary reports and more than 80 visualizations. By integrating metadata mapping, FLCA-based co-occurrence analysis [25,29], SS plots [22–24], normalized modularity (Q) [26], and AAC-enhanced reporting [9,27,30] within a browser-based interface, the application reduces the need for repeated preprocessing, programming, and software switching.

Compared with established tools such as CiteSpace, VOSviewer, App4WoSCCinR, and bibliometrix [14–17], Scopus Cluster is intended as a practical complement rather than a replacement. Its contribution is an upload-and-run workflow that supports simultaneous examination of authors, countries, institutions, departments, keywords, journals, and publication trends. The illustrative Figures demonstrated how collaboration patterns can be explored geographically, institutionally, and thematically, rather than being represented only by publication counts or ranked lists.

Metadata quality remains an important limitation in all bibliometric analyses. Scopus exports may contain incomplete, inconsistent, or variably formatted author, affiliation, keyword, and reference fields, requiring transparent checking and normalization before interpretation [11–14]. Scopus Cluster does not eliminate such source-level limitations; however, its structured workflow makes the transformation from uploaded metadata to analytic output more transparent and reproducible. Future development should expand metadata-quality checks, user-defined normalization rules, longitudinal analyses, and interoperability with other bibliographic databases (Figure 8).

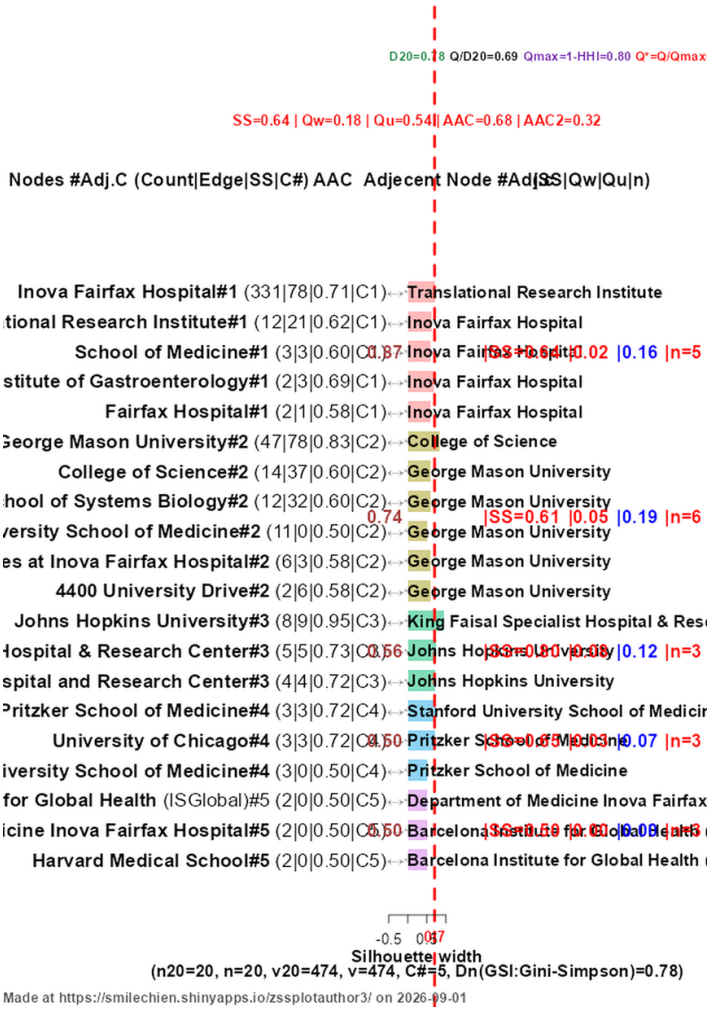


Figure 8: SSPlot for top 20 Institutes by count observed in this research (AAC=0.68).

Illustrative scholar-level findings

The scholar-level application showed the value of AAC-enhanced reporting beyond conventional citation rankings. Dr. Younossi's coauthor network had an AAC of 0.93, indicating substantial concentration around the leading investigator, whereas the institutional network had an AAC of 0.68, suggesting a more distributed structure. AAC therefore distinguishes leadership concentration from citation impact and publication volume [9,27,30].

This interpretation was further supported by Scopus Cluster [28] analysis of PubMed-indexed publications for the top 100 gastroenterology and hepatology scientists (Figure 9). Although many scientists had AAC values above 0.70, not all showed substantial author-collaboration concentration. Thus, a top-ranked scientist does not necessarily have a predominantly single-core collaboration pattern. AAC describes observed contribution and collaboration structures; it does not establish supervision, mentorship, or the quality of individual scholarly contributions.

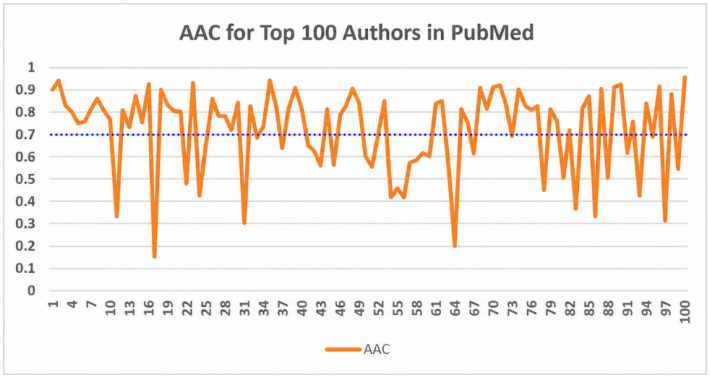


Figure 9: AAC distribution for the top 100 gastroenterology and hepatology scientists based on PubMed-indexed publications.

Methodological contribution and future potential

Scopus Cluster complements established bibliometric tools, including CiteSpace, VOSviewer, App4WoSCCinR and bibliometrix [14–17], by providing a browser-based upload-and-run workflow for Scopus CSV metadata. Its practical contribution is workflow integration: metadata mapping, FLCA-based analysis [25,29], SS and *Q* reporting [22–24,26], AAC calculation [9,27,30], and exportable visual outputs are available within one platform. The application is not intended to replace existing software; instead, it offers an accessible route from Scopus export files to interpretable reports. Future development could extend standardized data cleaning, cross-database input support, longitudinal comparisons, and customizable reporting templates (Figure 10).

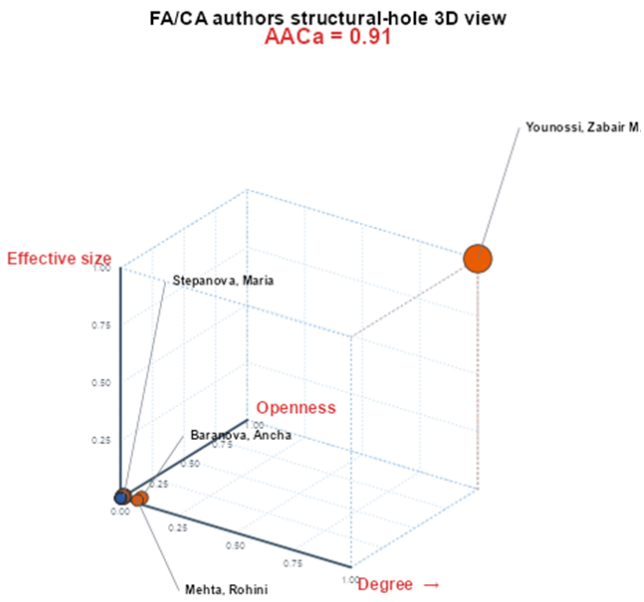


Figure 10: Structural-hole analysis for author-level network.

Limitations

This study used the 2025 Top 2% Scientists dataset [13] and Scopus-indexed publications; findings may differ with other databases, search strategies, metadata completeness, or author-name disambiguation procedures.

Network structure and SS, *Q*, and AAC values also depend on entity definitions, threshold settings, and clustering rules [22–26]. The illustrative analysis was cross-sectional and focused on a selected specialty; it does not establish temporal change, research quality, or mentorship. Future studies should apply consistent protocols across disciplines and annual datasets and evaluate whether observed network patterns remain stable over time.

Conclusion

Scopus Cluster provides an integrated, browser-based workflow for transforming Scopus CSV metadata into automated bibliometric summaries and visualizations. In the 2025 Top 2% gastroenterology and hepatology application, FLCA-based networks, SS, normalized modularity, and AAC jointly described collaboration structure, cluster quality, and leadership concentration. As a complement to existing bibliometric software, Scopus Cluster may facilitate reproducible Scopus-based research mapping and comparative monitoring across specialties and time periods.

Declarations

Ethics approval and consent to participate

Not applicable.

All data are publicly available in the WoSCC.

Consent to publish

Not applicable.

Availability of data and materials

All data used in this study are available in Appendices.

Competing interests

The authors declare that they have no competing interests.

Funding

There are no sources of funding to be declared.

Authors' Contributions

CSC and LF developed the study concept and design. CSC analyzed and interpreted the data. KFT monitored the process of this study and helped in responding to the reviewers' advice and comments. TWC drafted the manuscript, and all authors provided critical revisions for important intellectual content. The study was supervised by KFT. All the authors have read and approved the final manuscript.

Acknowledgments

We thank Enago (www.enago.tw) for the English language review of this manuscript.

Appendices

Appendix 1: XLSM (an Excel module used in this study)

Appendix 2: XLSM (Authors' publications)

Appendix 3: PDF (How to conduct this study)

Appendix 4: PDF (Silhouette and Modularity Metrics)

References

1. Chen B, Zhang X (2026) Global research trends of digital applications in medical education: A bibliometric analysis based on VOSviewer. Medicine (Baltimore)105: e50357.
2. Zhang M, Lin X, Gao W, Ren Y, Wang D, et al. (2026) Perioperative sleep disturbance: A bibliometric analysis. Medicine (Baltimore) 105: e50386.
3. Tang Y, Tan Y, Chen G, Ren C (2026) Research hotspots and trends of urology education in China and worldwide over the past decade: A Bibliometric Comparative Analysis. Medicine (Baltimore) 105: e50085.
4. Ioannidis JPA, Baas J, Klavans R, Boyack KW (2019) A standardized citation metrics author database annotated for scientific field. PLoS Biol 17: e3000384.
5. Yeung AWK, Tsoi JKH (2025) A Critical Correlate of QS Subject Ranking and Stanford's Top 2% Scientists in Top Dental Schools. Int Dent J 75: 100954.

6. Abdalla BA, Mustafa AM, Fattah FH, KaKamad FH, Omar SS, et al. (2025) Self-citation pattern among world's top 2 % of the scientists. Heliyon 11: e42471.

7. Ioannidis JPA, Boyack KW, Baas J (2020) Updated science-wide author databases of standardized citation indicators. PLoS Biol 18: e3000918.

8. Wu AL, Chou W (2024) Identifying China's distinctive academic fields among the top 2% of influential scientists: A bibliometric analysis using Rasch KIDMAP. Medicine (Baltimore) 103 : e36706.

9. Wang YA, Lai PC, Lai FJ (2026) Beyond Rankings: Mapping Local Cluster Separation, Global Community Structure, and Leadership Concentration Among the 2025 Top 2% Dermatology Scientists. Medicine (Baltimore). 2026; in print

10. Hirsch JE (2005) An index to quantify an individual's scientific research output. Proc Natl Acad Sci U S A. 102: 16569-16572.

11. Van Noorden R, Singh Chawla D (2019) Hundreds of extreme self-citing scientists revealed in new database. Nature 572: 578-579.

12. Ioannidis JP, Klavans R, Boyack KW (2016) Multiple Citation Indicators and Their Composite across Scientific Disciplines PLoS Biol 14: e1002501.

13. Elsevier Data Repository. August 2024 data-update for "Updated science-wide author databases of standardized.

14. Chen C (2006) CiteSpace II: Detecting and visualizing emerging trends and transient patterns in scientific literature. J Am Soc Inf Sci Technol 57: 359-377.

15. van Eck NJ, Waltman L (2010) Software Survey: VOSviewer, a computer program for bibliometric mapping. Scientometrics 84: 523-538.

16. Chien TW, Chen L, Chou W (2026) App4WoSCCinR: A web-based bibliometric analysis platform for the Web of Science Core Collection. SoftwareX 35:102770.

17. Aria M, Cuccurullo C (2017) bibliometrix: An R-tool for comprehensive science mapping analysis. J Informetrics 11: 959-975.

18. Stirrat T, Hu B, Flatley M, Apte S, Thiru S, et al. (2026) Gender authorship in gastroenterology and hepatology: trends and persistent seniority gaps. J Gastroenterol Hepatol 41: 2605-2612.

19. Zhao Q, Liang L, Zhai F, Ling G, Xiang R, et al. (2023) A bibliometric and visualized analysis of liver fibrosis from 2002 to 2022. J Gastroenterol Hepatol 38: 359-369.

20. Tang X, Gong W, Yuan F, Li R, Han X, et al. (2016) Top-cited articles in digestive system disease from 1950 to 2013. J Gastroenterol Hepatol 31: 107-111.

21. Chien TW, Chou W (2025) 10StepsinExcelforWoS: Streamlining bibliometric analysis in web of science using Microsoft excel. SoftwareX 32:102403.

22. Rousseeuw PJ (1987) Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. J Comput Appl Math 20:53-65.

23. Chuang HY, Chou W (2025) SilhouetteScoreinR: Beyond traditional network layouts by leveraging local cohesion and nearest neighbor separation. MethodsX 15: 103622.

24. Chien TW, Chen L, Chou W (2026) SilhouettePlot: An R-based app integrating silhouette scores into network plots for enhanced clarity and meaning. SoftwareX 34: 102652.

25. Chien TW, Chen L, Chou W (2026) Cluster Choice: A web-based tool for interpretable clustering and algorithm comparison in bibliometric networks. SoftwareX 35: 102794.

26. Newman MEJ (2006) Modularity and community structure in networks. Proc Natl Acad Sci U S A. 103: 8577-8582.

27. Yang DH, Chien TW, Yeh YT, Yang TY, Chou W, et al. (2021) Using the absolute advantage coefficient (AAC) to measure the strength of damage hit by COVID-19 in India on a growth-share matrix. European Journal of Medical Research 26: 61.

28. Chien TW (2026) Scopus Cluster for bibliometrics using Scopus database.

29. Lai PC, Chou W, Chien TW, Lai FJ (2023) A modern approach with follower-leading clustering algorithm for visualizing author collaborations and article themes in skin cancer research: A bibliometric analysis. Medicine (Baltimore)102: e34801.

30. Li MJ, Lai FJ (2025) Classifying research leadership in infantile hemangioma studies: A bibliometric analysis using Kano diagrams and absolute advantage coefficients. Medicine (Baltimore) 104: e44905.

31. Github (2026) Scopus Cluster with R scripts.

32. Andrich D (1978) A rating formulation for ordered response categories. Psychometrika 43: 561-573.

33. Burt RS (1992) Structural Holes: The Social Structure of Competition. Cambridge, MA: Harvard University.

34. Schoch D, Valente TW, Brandes U (2020) Unpacking Burt's constraint measure. Social Networks 62: 50-57.