Overview of Bibliometrics & Citation Analyses

coronis group
Jeno Mavzer

Overview of Citation-Based Analyses

coronis group

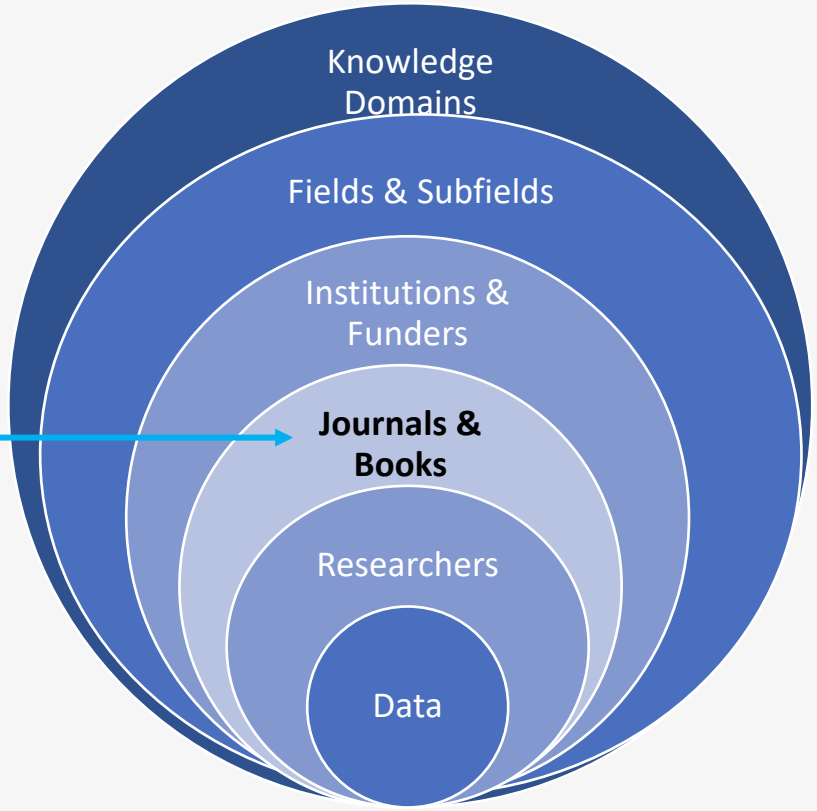
What We Do

We are the only consulting group using bibliometric and citation analysis tools to improve publishing and editorial products.

Analyses are accompanied with subject-specific recommendations — we recognize that publications are complex and multidimensional entities, requiring custom approaches.

Above all, our aim is to make publications better.

We work with the producers of research media directly. That is, within the ecosystem of research interaction, we focus on one element.



Client Profile

- Base includes small to mid-size scholarly associations, editorial teams, and other producers of research media
- Focus on contextualizing and translating citation data into actionable editorial strategy
- We help organizations understand their publication's identity, efficacy, and development opportunities in the research ecosystem
- The tools we use can be adopted by clients for continuous process monitoring and improvement

Reporting Aims

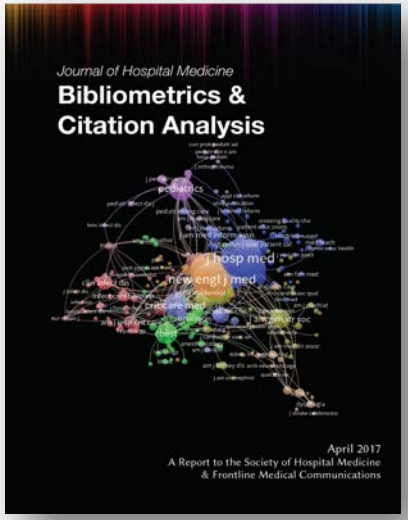
- Provide data at the field level and identify subfield-specific structure and dynamics
- Evaluate traditional and alternative metrics in light of target journal's citation dynamics and goals
- Compile article-level data informing metrics best fitting journal composition
- Identify composite citation peaks for target journal content performance over time
- Chart citable to non-citable article ratios to document historic content heterogeneity
- Examine citation patterns to define co-citation communities
- Identify top-cited historical and recent content
- Visualize and interpret selected citation dynamics

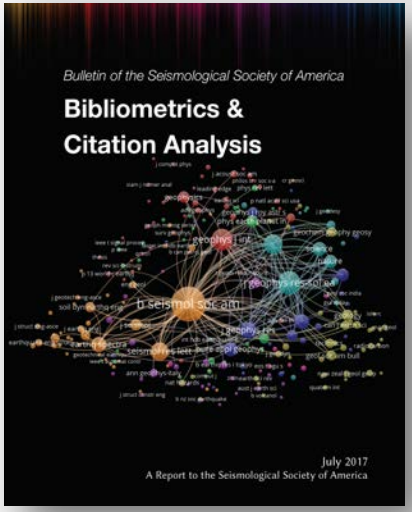
Techniques

Editorial interest	Analytic Strategy
Classic papers	• Citation count • Citation peak map
Competitor title identification	• Cited by/citing data • Network co-citation mapping
Competitor analysis	• Article type breakdowns & citation counts • Lists of highly cited articles • Never-cited rates and h-indices
Editorial board candidates, author discovery	• H-index corrected for career maturity • Altmetrics and community engagement • Centrality in journal network space
Digital efficacy	• Altmetrics, backlinks, and other usage statistics • Indices of citation speed
Highly cited topics	• Citation velocity/density • Topical citation maps

Data Sources

- All discussions begin with description of data sources (Clarivate, Scopus, MA, Google), their associated indicators, and limitations.
- We emphasize that **no one indicator fits all journals** — encourage return to source data in evaluating performance or making content decisions.
- Have been vocal advocates of participation in Open Citations Corpus to:
  - wean publishers/editors from blind reliance on proprietary metrics that may not be appropriate to content
  - ground editorial decisions in matrix of citation-based data points





Elements of Typical Reports

coronis group

Structure

Description of Data Sources

- Web of Science (WoS)/Scopus
- Google Scholar
- Microsoft Academic
- Open Citations Corpus (OCC)

Types of Indicators

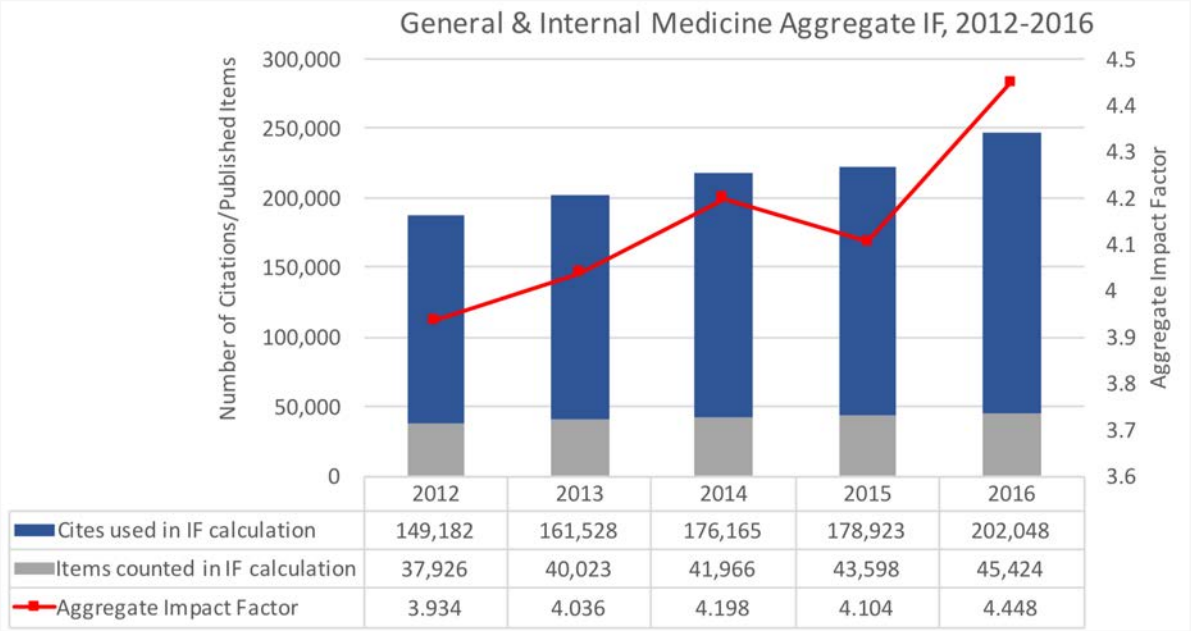
- Field-level (FCR, manual cohort clustering)
- Journal-level (Eigenfactor, Impact Factor, CiteScore, SJR, SNIP/IPP, h5-index)
- Article-level (citation counts, RCR, Altmetrics)
- Author-level (h-index, i10-index, m-parameter)

Citation Analysis

- Citation Peaks
- Networks of Influence
- Topic Mapping

Field-Level

Clients are generally shown JCR or Scopus field aggregate data to broadly contextualize journal performance.



	2012	2013	2014	2015	2016
Cites used in IF calculation	149,182	161,528	176,165	178,923	202,048
Items counted in IF calculation	37,926	40,023	41,966	43,598	45,424
Aggregate Impact Factor	3.934	4.036	4.198	4.104	4.448

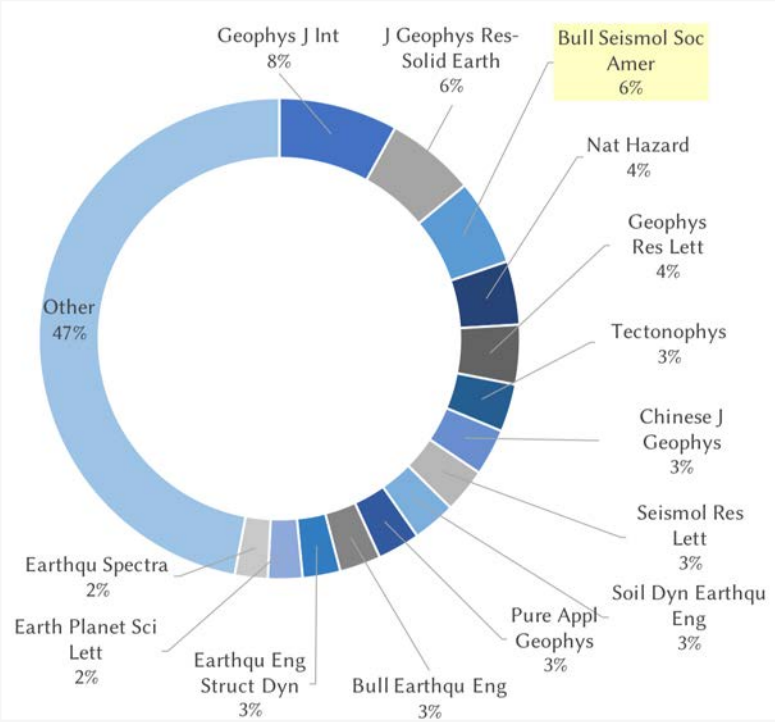
Presenting field citation ratios (captured from a defined set of related titles in OCC) would likely be better practice…

Field-Level

Most platforms define subjects too broadly for niche journals to draw fair comparison with other titles.

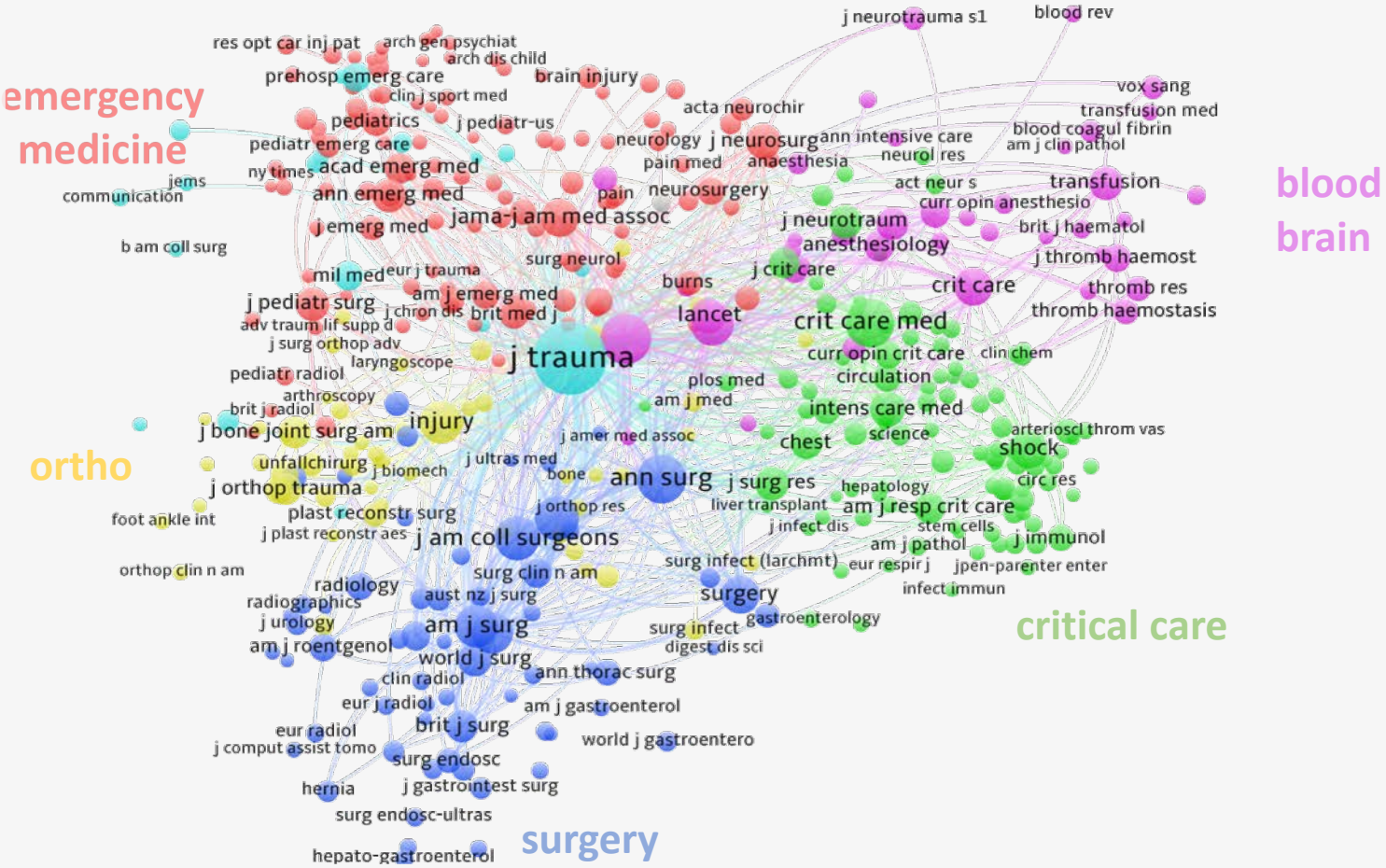
To decouple the data from platform-specific indicators, we define narrower subject cohorts.

Smaller sets illustrate where relevant content is landing (and, subsequently, how well a publication is attracting desired content).



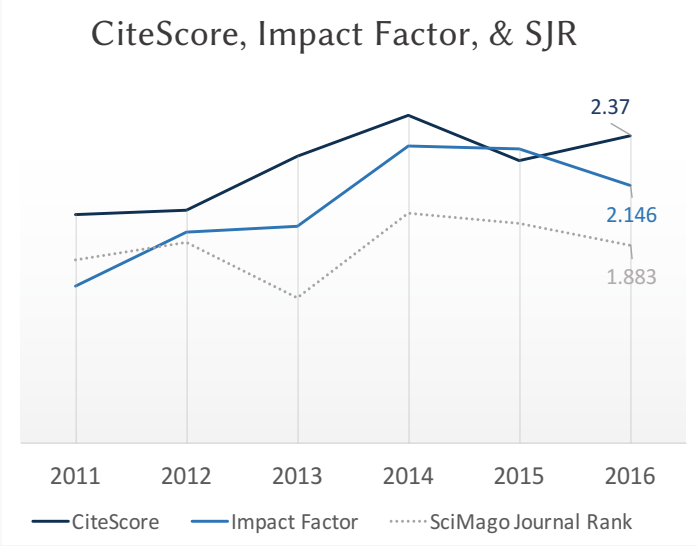
Field-Level

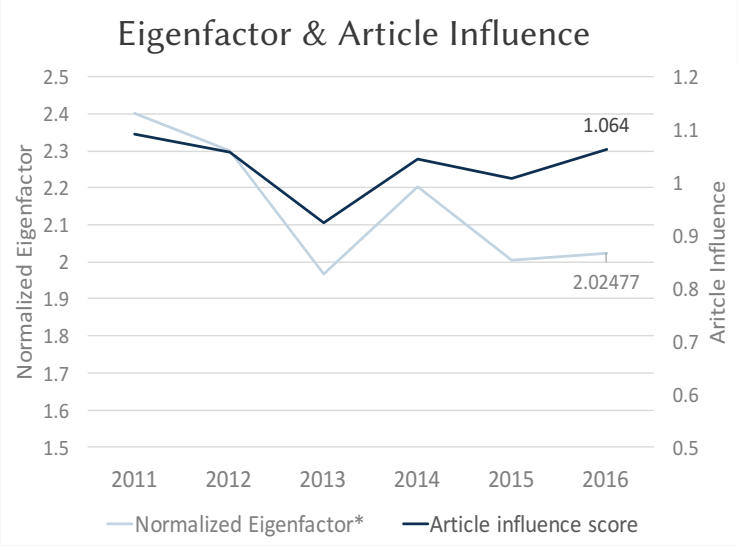
For interdisciplinary journals, we construct an artificial set of all relevant articles (identified by MeSH term) and use fractional citation weighting to examine subject reach, citation share, and dynamics between titles.

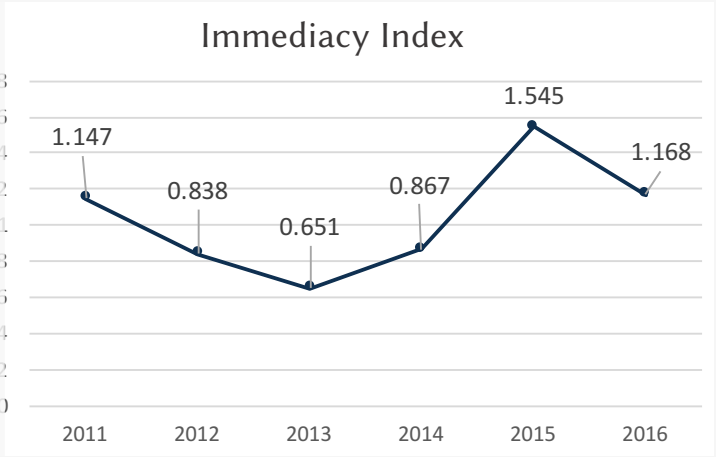


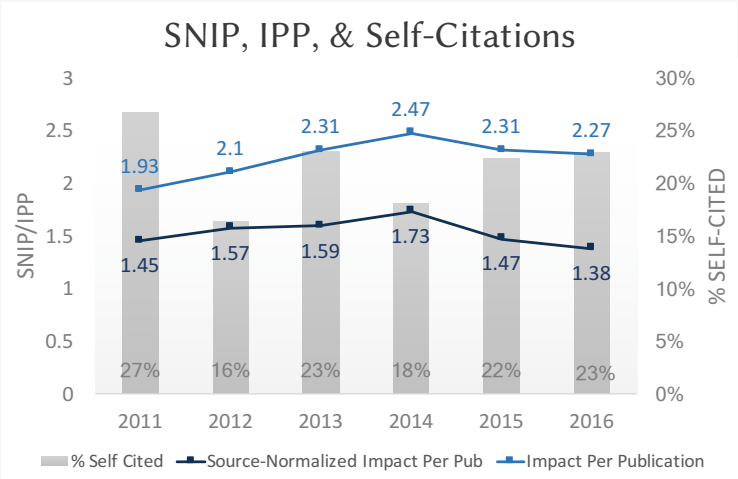
Journal-Level

Most analyses at the journal-level begin with a rundown of historical performance (5- or 10-year) across extant indicators.







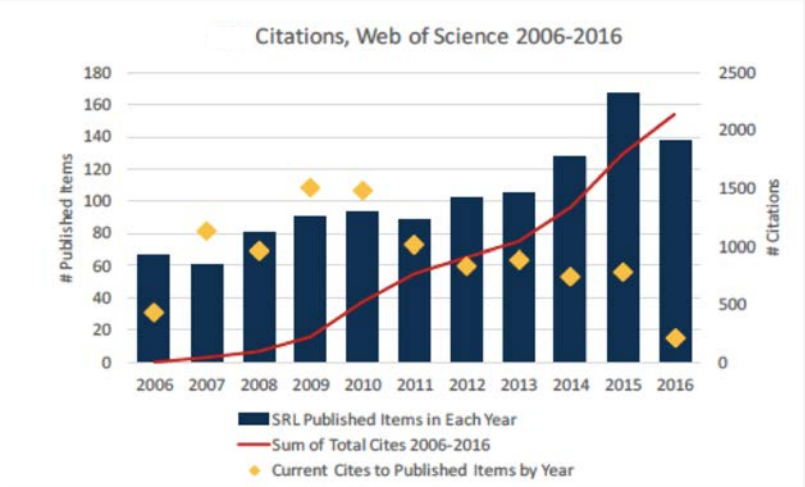


Journal-Level

The tour of indicators opens the ground for a deep dive into content decisions.

Starting from a wide view of citation history, we drill down into the origin of articles and content types.

We supplement discussion of reasons for rises/falls of particular indicators via recourse back to field trends, content collection performance (or lack thereof), and editorial policy effects.

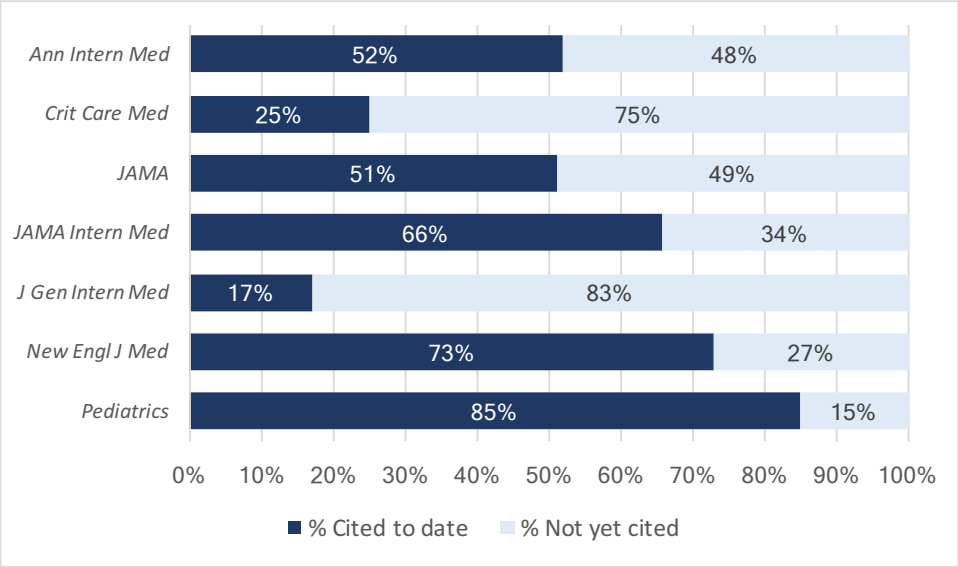


Article Origin	% Published (# articles/all pubs)	% Citations Received to Date
Affiliate Society A	12.5%	16.4%
Affiliate Society B	8.7%	5.9%
Affiliate Society C	0.8%	0.2%
Affiliate Society D	0.7%	0.8%
Affiliate Society E	10.4%	6.8%
Supplements	6.9%	7.1%
Independent	60%	62.7%

Key: Citation Benefit | Citation Neutral | Citation Deficit

Journal-Level

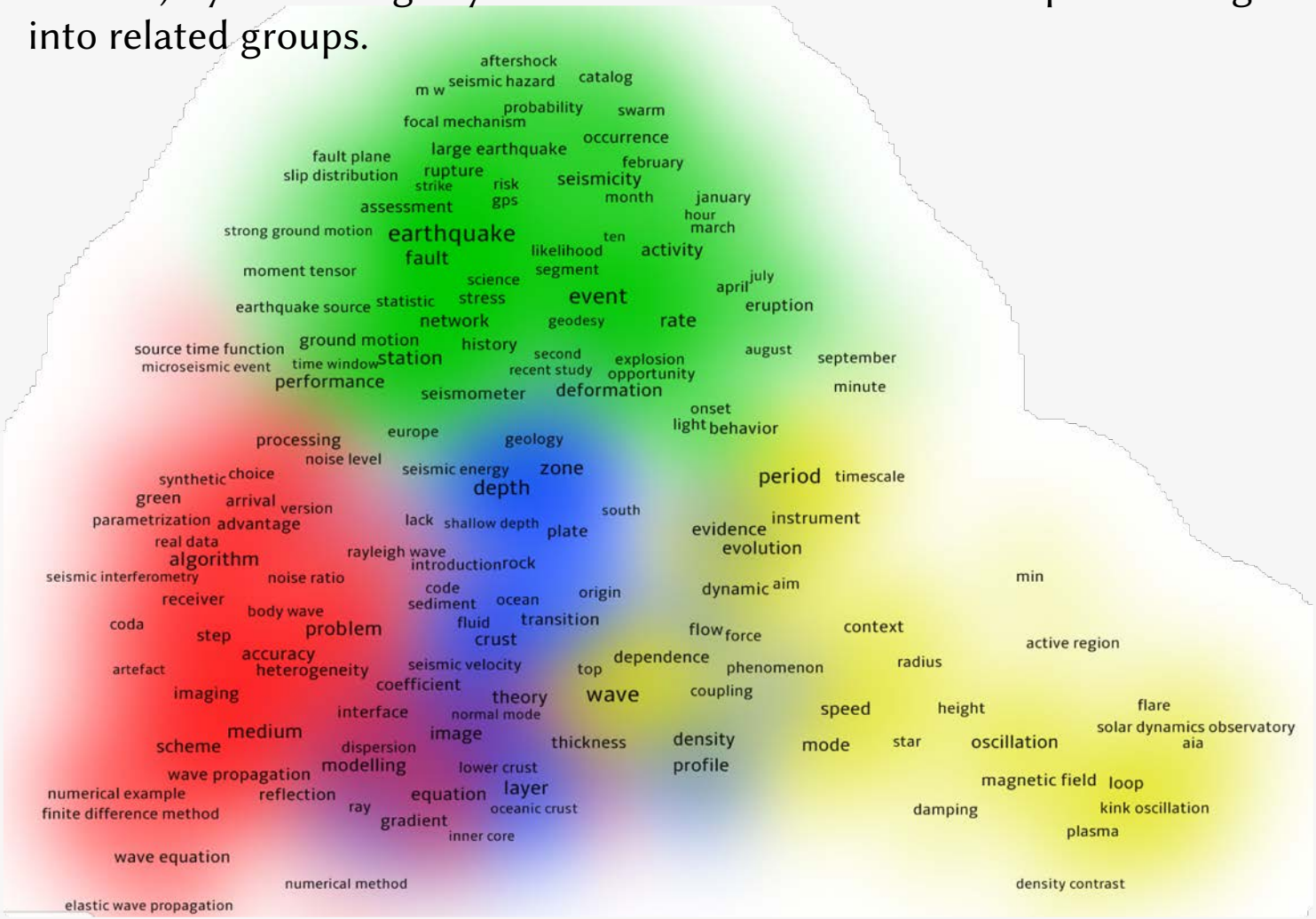
We also address competitor analysis queries raised by editorial teams (e.g. “how does our never-cited rate compare with other titles?”)



Journal	Total Pubs 2014-2015	Total "Citable" Pubs 2014-2015	Total Cites to Date	# Pubs Not Cited to Date	h-Index
Ann Intern Med	1,333	326	12,933	645	53
Crit Care Med	3,949	622	10,398	2,966	39
JAMA	3,236	425	41,310	1,574	91
JAMA Intern Med	1,300	316	11,720	448	49
J Gen Intern Med	3,453	438	3,534	2,863	22
New Engl J Med	3,140	695	106,938	847	165
Pediatrics	1,623	1,329	15,802	246	42

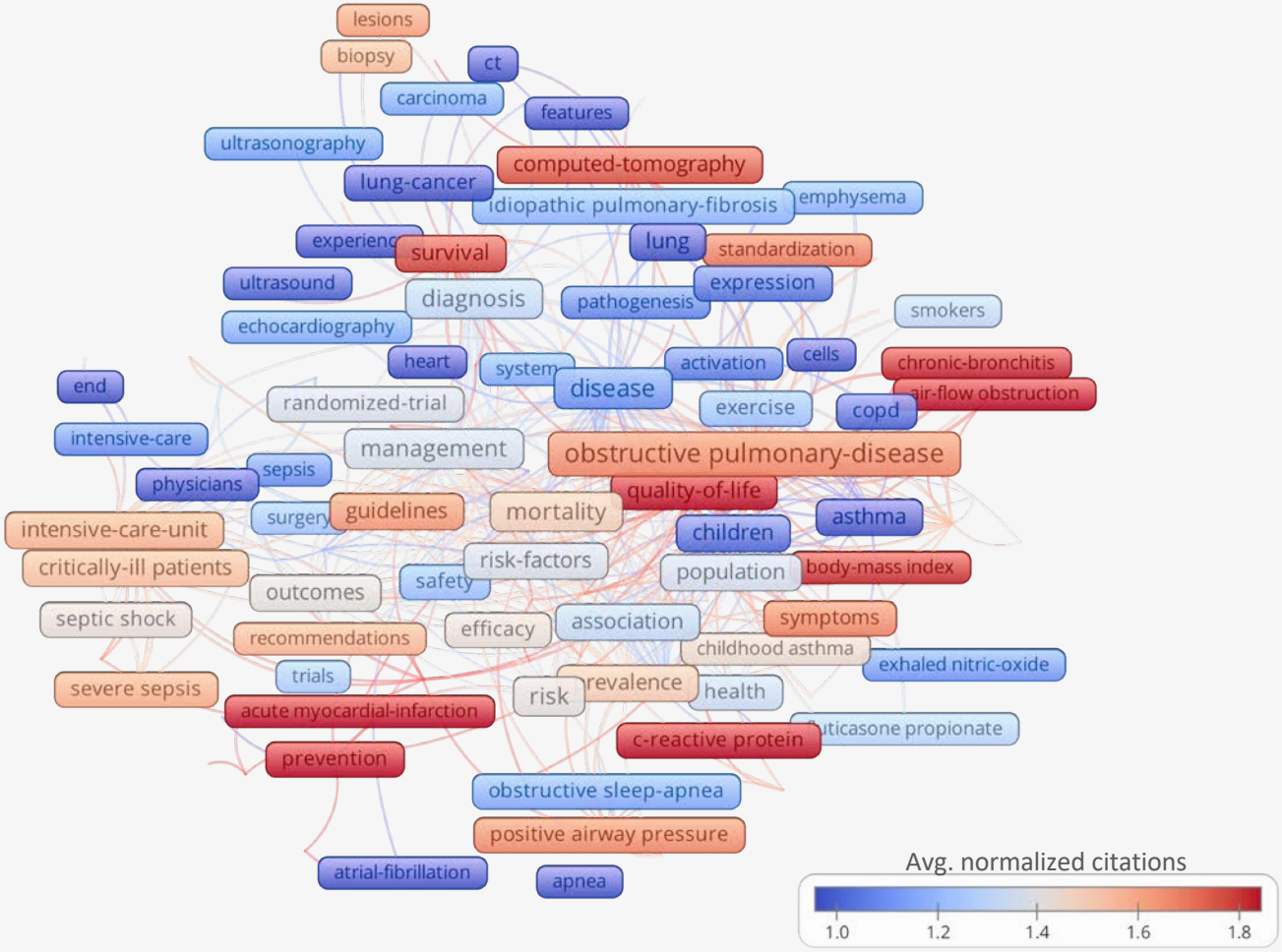
Journal-Level

We visualize topic density and structure in a journal (or across a group of titles) by extracting keywords from titles/abstracts and partitioning into related groups.



Journal-Level

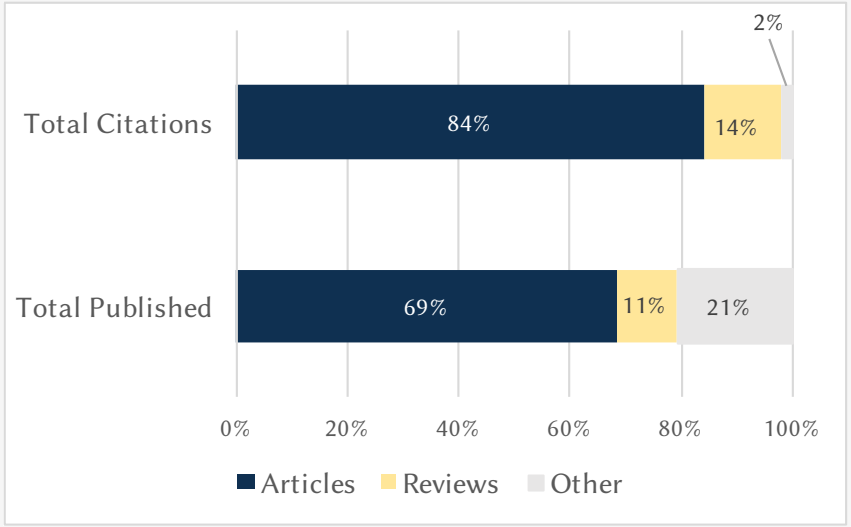
Overlaying keyword frequency maps with citation data (gross or through time) permits identification of hot topics and informs discussion of content direction and new product development.



Article-Level

Editors are concerned with expending resources on impactful content.

To that end, we pull historical submission information from peer-review platforms and correlate each item with its subsequent citation performance.

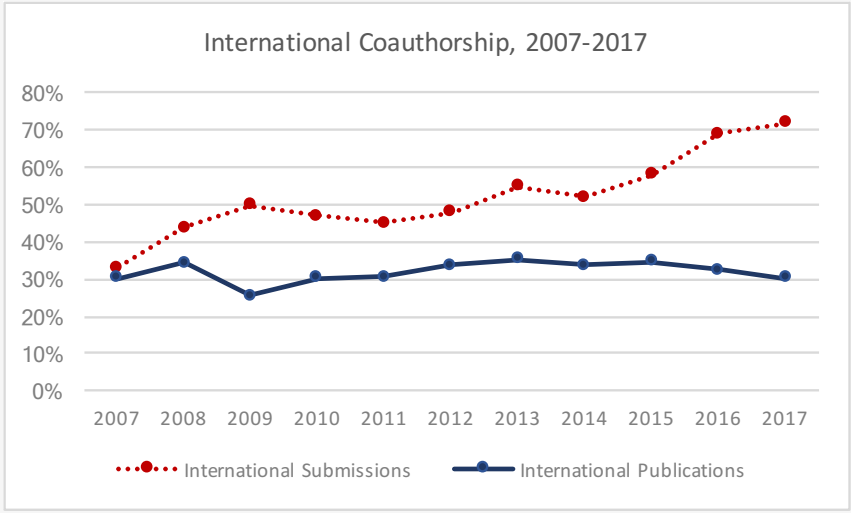


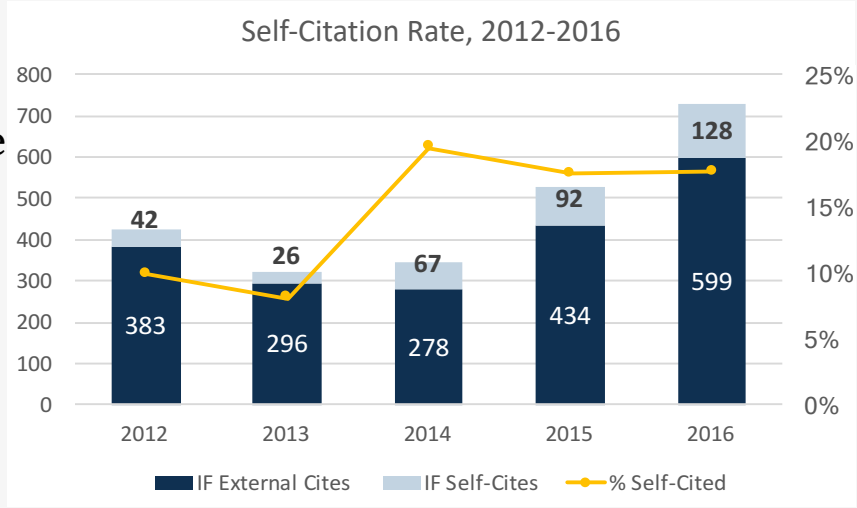
	Article type	Total Submitted	Total Published	% All Published	Total Citations	% All Citations	h-Index	Cited to Date
Citable	Articles	1233	548	68.6%	3165	84.2%	22	82%
	Reviews	114	87	10.9%	521	13.9%	12	91%
Non-citable	Editorial	94	86	10.8%	48	1.3%	3	54%
	Letters	89	76	9.5%	25	0.7%	3	21%
	Corrections	2	2	0.3%	0	0	0	0

Article-Level

Other article-level data points that we interpret include:

- Altmetrics trends for individual articles and content groups
- International co-authorship rates and country demographics (submission through publication)
- Self-citation and never-cited rates
- Lists of highly cited articles by year or according to type (i.e. clinical vs. basic science, supplements)
- For clinical outlets, highly ranked collections by RCR

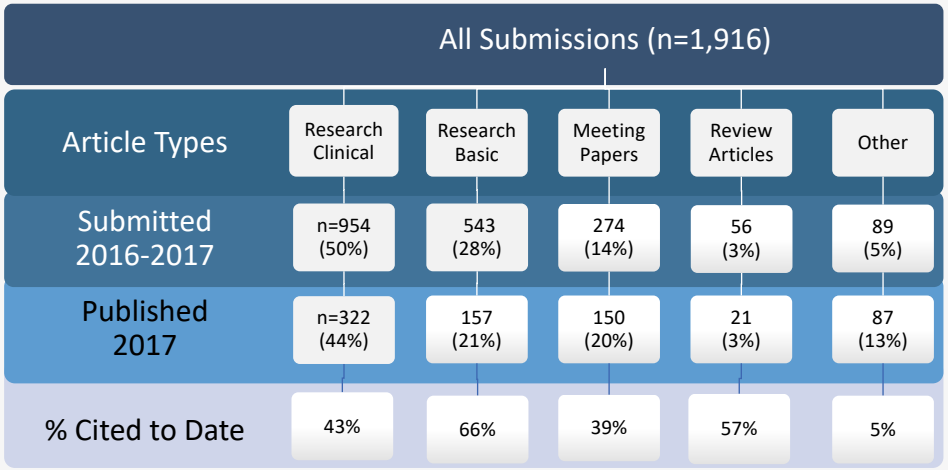




Journal-Level

For underperforming titles, we examine the full publication lifecycle, from peer review through post-publication.

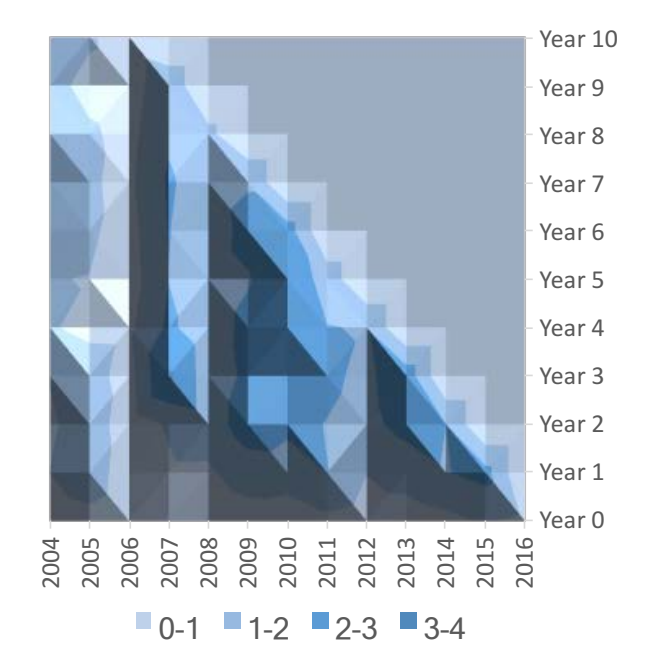
We correlate submission information from peer-review platforms with subsequent publication and citation data.

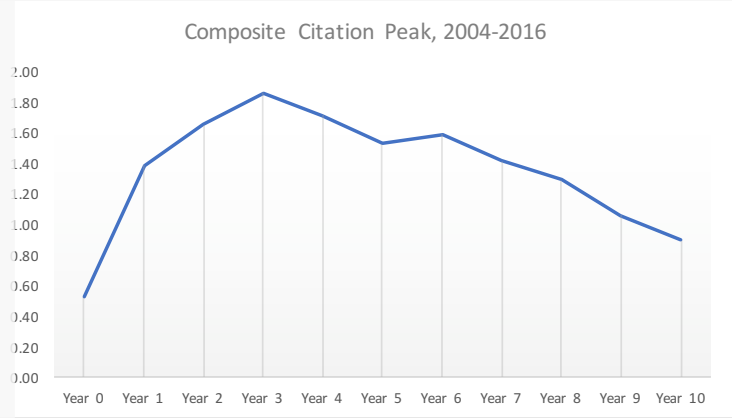


This yields insight into processes and policies that can be amended long before publication to improve the quality of overall research output. Common needs include:

- Adoption of clear reporting standards
- Increasing information density of original research reports
- Implementation of statistical review checklists
- Reproducibility or data transparency guidelines
- Commissioning to bridge content gaps

Article-Level



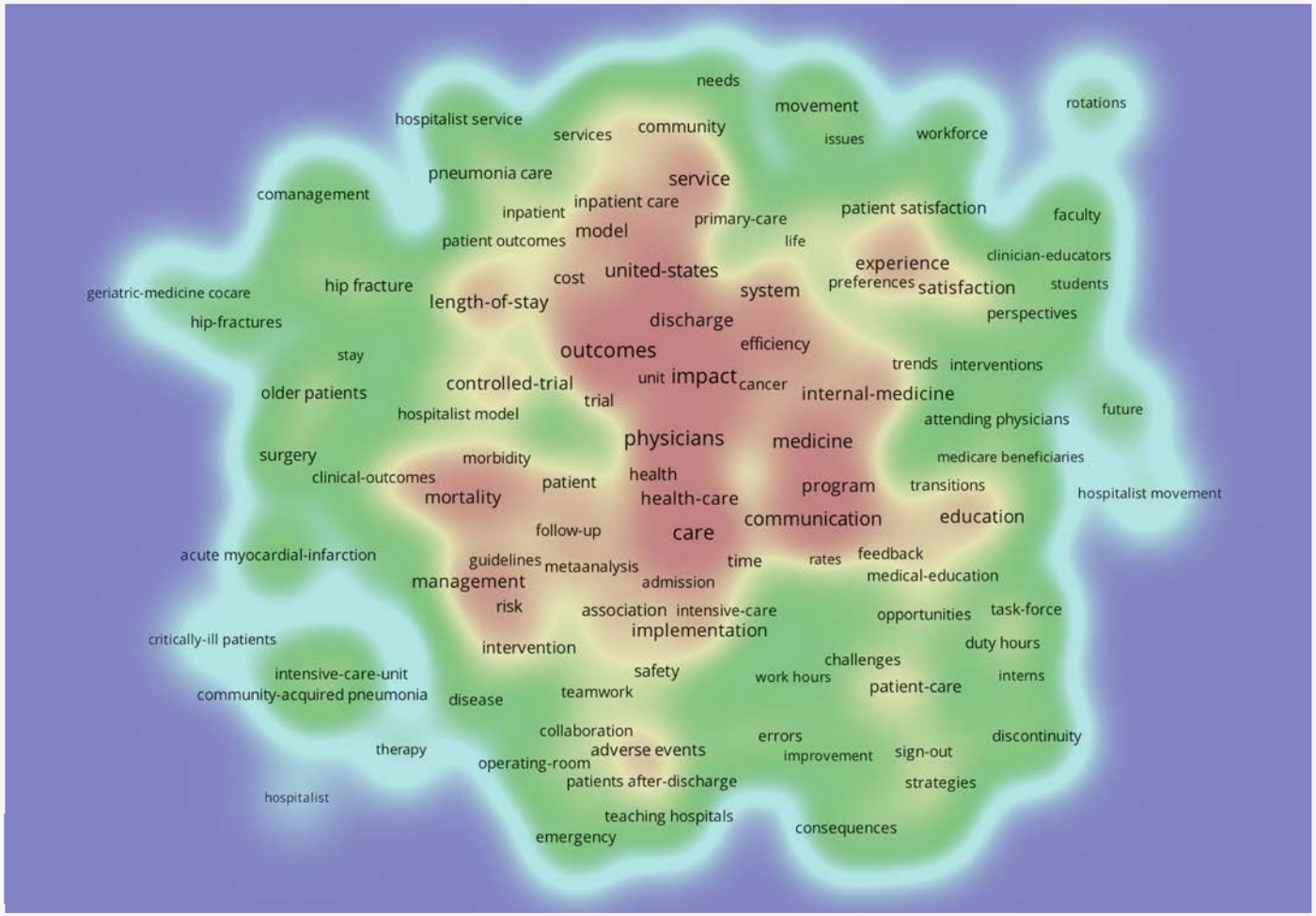


We calculate citation peaks for publications. Peaks are calculated for all articles published over 10-20 years to:

- map citation accrual differences for major article types
- identify classics and “sleeping beauty” papers
- provide data-driven advise on opening up content archives
- guide clients on platform selection and adoption of metrics appropriate to citing norms

Article-Level

Overlaying topics identified via keyword frequency with citation data permits identification of topics gaining momentum.



Author-Level

At the author level, end use of data determines collection strategy.

For general intelligence or author discovery:

- Most frequent contributors and their domains of expertise
- Highly cited authors, optionally with publication key words
- Author groups and affiliate organizations

For editorial board candidates:

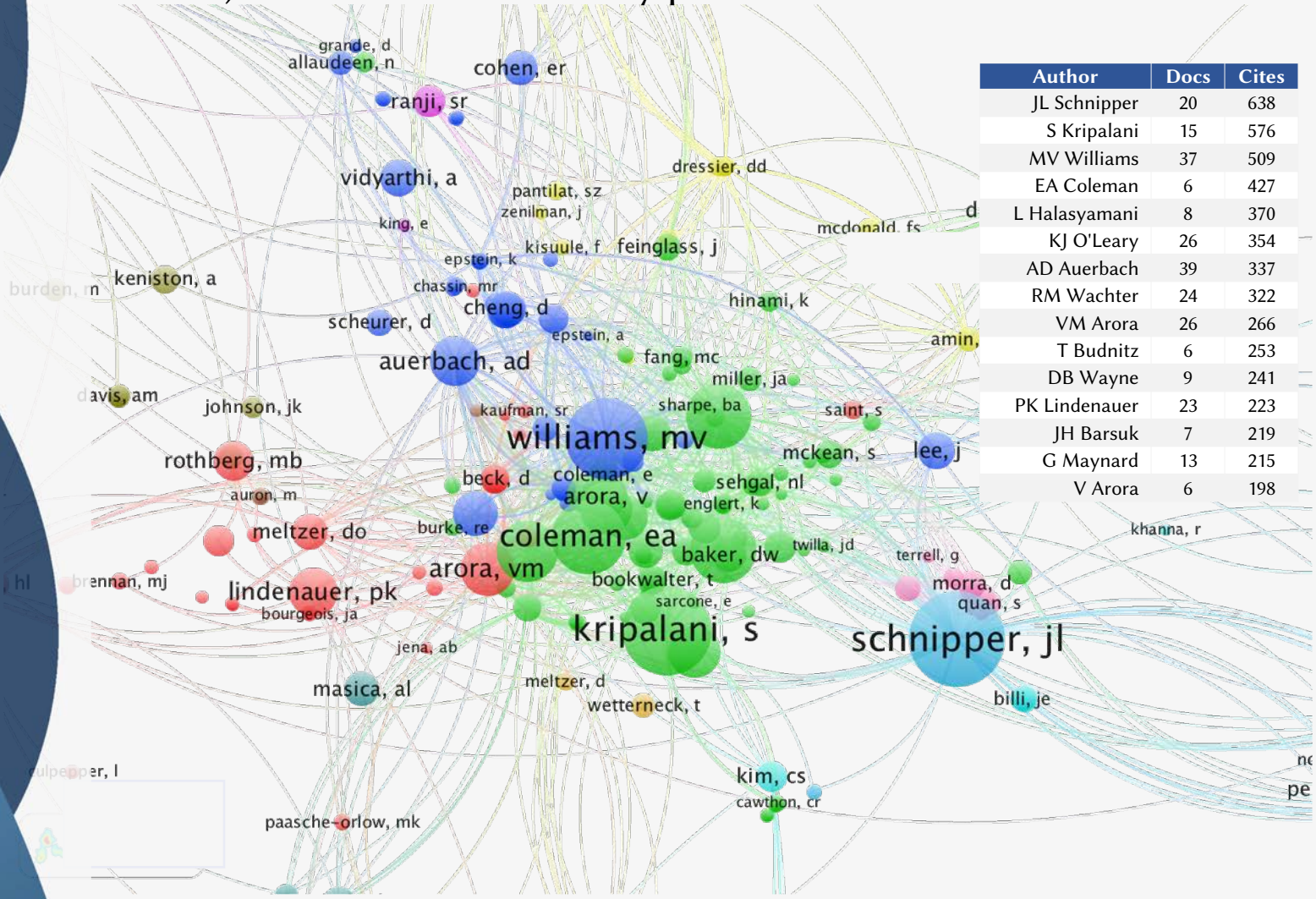
- Historical journal participation (# papers, # citations, cites per item, co-authorship groups)
- H-index normalized for career maturity (i.e. m-index dated to first publication in literature)

For editor selection:

- Centrality in citation network space
- Link association with author groups or topics

Author-Level

Author relationships (in terms of both co-authorship and citation direction) are visualized to identify potential board members or editors.



Author	Docs	Cites
JL Schnipper	20	638
S Kripalani	15	576
MV Williams	37	509
EA Coleman	6	427
L Halasyamani	8	370
KJ O'Leary	26	354
AD Auerbach	39	337
RM Wachter	24	322
VM Arora	26	266
T Budnitz	6	253
DB Wayne	9	241
PK Lindenauer	23	223
JH Barsuk	7	219
G Maynard	13	215
V Arora	6	198

Get in touch

consult@coronisgroup.com

http://coronisgroup.com

coronis group