Int. J. Pharm. Investigation, 2026; 16(4):1657-1665.
https://www.jpionline.org

Original Article

Thematic Structure and Evolution of a Pharmaceutical Sciences Research Corpus from Dr. Anroop Nair

Mueen Ahmed K. K[1,*], Chaman Sab M[2], Mohammed Yunus[1], Shery Jacob[3], Jigar N. Shah[4], Hiral J. Shah[5], Mahesh Attimarad[6], Sreeharsha Nagaraja[6], Rashed M. Almuqbil[6], Bandar E. Al-Dhubiab[6], Katharigatta N. Venugopala[6], Abdullah M. Alotaibi[7], Ahmed S. Alnaim[6]

[1]Manuscript Technomedia LLP, Vivekananda Nagar, Bengaluru, Karnataka, INDIA.
[2]A.R.G. College of Arts and Commerce, Davanagere, Karnataka, INDIA.
[3]Department of Pharmaceutical Sciences, College of Pharmacy, Gulf Medical University, Ajman, UNITED ARAB EMIRATES.
[4]Department of Pharmaceutics, Institute of Pharmacy, Nirma University, Ahmedabad, Gujarat, INDIA.
[5]Department of Pharmaceutics, Parul College of Pharmacy and Research, Parul University, Ahmedabad, Gujarat, INDIA.
[6]Department of Pharmaceutical Sciences, College of Clinical Pharmacy, King Faisal University, Al-Ahsa, SAUDI ARABIA.
[7]Department of Pharmaceutical Care, Al-Ahsa Health Cluster, Al-Ahsa, SAUDI ARABIA.

ABSTRACT

Objective: Drug delivery and pharmaceutical-analysis research spans a wide range of sub-disciplines - nanocarrier formulation, transdermal and iontophoretic delivery, analytical method development, and computational drug discovery - that are rarely mapped together quantitatively. This study applies scientometric network analysis to a Scopus-indexed pharmaceutical sciences publication corpus to characterize its thematic structure, temporal evolution, and most productive contributors. **Design/Methodology/Approach:** A total of 289 documents published between 2003 and 2026 were retrieved from Scopus. A keyword co-occurrence network was built from the Author Keywords and Index Keywords fields using a Python-based pipeline (networkx, scikit-learn, matplotlib) implementing Louvain community detection, modularity (Q) and mean-silhouette (S) coherence scoring, TF-IDF cluster labelling, and Kleinberg-style burst detection - conceptually equivalent to the co-word mapping produced by VOSviewer or CiteSpace. **Findings:** The network comprised 300 keyword nodes and 18,228 edges (density = 0.406), resolving into six clusters (Q = 0.141, weighted mean S = 0.172). The largest cluster (82 member terms) centered on drug-delivery-system formulation and release; the second largest (64 terms) on computational and mechanistic drug chemistry, including molecular docking. Annual output rose from single digits before 2018 to a peak of 40 documents in 2021, and the corpus accumulated 10,617 citations overall (≈36.7 per document). Fifteen keyword bursts were detected; the strongest was a 2009 surge in hydrogen-ion-concentration/pH-related terms (strength = 15.0), tied to iontophoretic and ungual drug-transport studies published that year. **Practical implications:** The cluster structure and burst timeline provide a data-grounded map of where formulation science, analytical chemistry, and computational drug discovery intersect in this corpus, useful for identifying collaboration opportunities and under-examined niches such as green analytical chemistry and AI-assisted docking. **Originality/value:** The analysis relies on keyword co-occurrence only, since the underlying export lacked a References field and citation-based co-citation mapping was therefore not possible. The modularity and silhouette scores obtained (Q = 0.141, S = 0.172) indicate a relatively diffuse rather than sharply modular network - a limitation reported explicitly rather than overstated - which itself is informative about how densely interconnected this corpus's vocabulary is.

Keywords: Bibliometrics, Drug Delivery, Keyword Co-occurrence, Pharmaceutical Analysis, Research Trends, Scientometric Analysis, Scopus.

Correspondence:
Dr. Mueen Ahmed KK
Ph.D, Manuscript Technomedia LLP, No. 22 (New No. 40), 3rd Cross, Vivekananda Nagar, Bengaluru-560033, Karnataka, INDIA.
Email: mueen.ahmed@gmail.com

Received: 22-04-2026;
Revised: 31-05-2026;
Accepted: 01-07-2026.

INTRODUCTION

The global drug-delivery technology market was valued at roughly $52-54 billion in 2025 and is projected to keep growing at a compound annual rate above 10% through the early 2030s, driven by chronic-disease burden, biologics, and demand for patient-centered, minimally invasive administration routes (Research and Markets, 2026). That commercial growth sits on top of a much larger, more fragmented academic literature: formulation scientists optimizing nanocarriers and transdermal systems, analytical chemists validating quantification methods, and computational chemists screening candidate molecules by docking, all publishing under loosely overlapping keyword vocabularies. Traditional narrative reviews typically pick one of these sub-fields and follow it in depth, but they rarely show how the sub-fields relate to one another within a single research programme or institutional output - which themes dominate by

OPEN ACCESS

Manuscript TECHNOMEDIA

DOI: 10.5530/ijpi.20260004

Publishing Partner : Manuscript Technomedia. [www.mstechnomedia.com]

volume, which are recent versus foundational, and where sudden surges of attention have occurred.

Scientometric or bibliometric network analysis addresses this gap by treating a bibliographic export as data rather than as a set of individual papers to read narratively: keyword co-occurrence, publication counts, and citation patterns are mapped quantitatively, clustered, and interpreted as a structural description of a field or corpus (Donthu *et al.*, 2021). Recent applications of this approach to pharmaceutical sub-topics illustrate its value - for example, a 2025 bibliometric mapping of intranasal nanoemulsion research identified 379 manuscripts and traced a clear shift toward central-nervous-system delivery applications over 2004-2024, while comparable bibliometric studies of Saudi Arabian pharmaceutical-science output have documented a sustained rise in both publication volume and journal quality since 2009 (Rohra and Memon, 2025).

Despite this growing methodological toolkit, most published scientometric studies of pharmaceutical research define their corpus by a topic-level Boolean search (e.g., "nanoemulsion AND drug delivery"). Far less common is applying the same network-analytic machinery to a real, already-assembled multi-author research corpus to ask what its internal thematic structure actually looks like - which is the approach taken here. The corpus analyzed in this paper spans formulation science (nanoemulsions, nanocarriers, hydrogels), physical/analytical pharmacy (HPLC method development and validation), preclinical pharmacology (interspecies dose translation, animal pharmacokinetics), and computational drug discovery (molecular docking, virtual screening) - four sub-fields that, per the market and bibliometric evidence above, are each independently active and growing, but whose joint structure within one research output stream has not been previously quantified for this specific corpus.

This study therefore asks: (1) How has annual publication and citation output evolved across this corpus between 2003 and 2026? (2) What thematic clusters does keyword co-occurrence analysis reveal, and how coherent and well-separated are they? (3) Which specific keywords show sudden bursts of attention, and what do these bursts correspond to substantively? (4) Who are the most prolific contributing authors and affiliations, and how concentrated is the output geographically/institutionally?

To answer these questions, 289 Scopus-indexed documents were retrieved and analyzed using a reproducible, Python-based network-analysis pipeline that computes the same core metrics - modularity, mean silhouette, TF-IDF cluster labelling, and citation/keyword burst detection - used by dedicated bibliometric tools such as CiteSpace and VOSviewer. The remainder of the paper proceeds as follows: Section 2 briefly reviews prior scientometric work in adjacent pharmaceutical sub-fields; Section 3 details the data source, search strategy, and analytical methods; Section 4 reports the descriptive, network, cluster, and burst results; Section 5 discusses what this structure implies about the corpus's evolution and situates the emerging clusters against current external literature; and Section 6 concludes with directions for future work.

LITERATURE REVIEW

Prior scientometric work on the individual sub-fields represented in this corpus is well developed but fragmented across separate literatures. In nanoemulsion-based drug delivery, recent narrative reviews describe rapid growth in oral, parenteral, ocular, transdermal, and intranasal applications over the last five years, driven by nanoemulsions' capacity to enhance bioavailability and enable non-invasive central-nervous-system targeting (Chatzidaki and Mitsou, 2025; Dwivedi *et al.*, 2025). A dedicated 2025 bibliometric study of intranasal nano- and micro-emulsions specifically - the first of its kind for that sub-niche - used Bibliometrix on Scopus data to map institutional leadership and growth trends, underscoring how narrow a slice of the broader nanocarrier literature a single topic-level bibliometric study typically covers.

In transdermal and iontophoretic delivery, review literature has traced the technology through successive device generations, from passive patches to actively enhanced systems combining iontophoresis with microneedles or nanocarriers (Shah *et al.*, 2025; Gross *et al.*, 2025; Wang *et al.*, 2022) specifically reviewed the physicochemical and device factors influencing iontophoretic transport, a sub-topic directly represented among this corpus's clusters. In pharmaceutical analysis, the last several years have seen a marked shift toward "green" HPLC method development - methods explicitly assessed against solvent-consumption and environmental-impact metrics such as the AGREE score - alongside continued growth in conventional validated HPLC assays for combination formulations (Nassar *et al.*, 2025). In computational drug discovery, molecular docking has become, in the words of a 2024 Annual Review of Biochemistry synthesis, "an essential part of a structural biologist's and medicinal chemist's toolkit," with docking increasingly paired with machine-learning-guided virtual screening of very large compound libraries (Paggi *et al.*, 2024).

What these four literatures share, and what motivates the present analysis, is that each has recently been reviewed or bibliometrically mapped on its own terms, but never jointly, within the bounds of one coherent publication corpus. This paper's scientometric approach differs from each of the reviews above in that it does not pre-select a single sub-topic by search string; instead, it applies unsupervised network clustering to an already-existing corpus and lets the keyword co-occurrence structure itself reveal how many distinct thematic clusters are present and how they relate temporally - a design better suited to answering "what does this

body of research actually contain" than a topic-first bibliometric search would be.

MATERIALS AND METHODS

Rationale for a Scientometric Approach

A quantitative, network-based mapping approach is well suited to a multi-decade, multi-sub-discipline corpus of this kind because no single reviewer can practically synthesize 289 documents spanning formulation science, analytical chemistry, preclinical pharmacology, and computational chemistry with consistent thematic granularity. Keyword co-occurrence analysis instead lets the documents' own indexing vocabulary define thematic proximity empirically, and burst detection surfaces sudden shifts in attention that a manual read-through would be likely to miss or under-weight.

Data Source and Search Strategy

Bibliographic records were retrieved from Scopus (Elsevier) on July 2, 2026, using the search string AU-ID ("Nair, Anroop B." 57195713288), which returns every Scopus-indexed document associated with that author identifier. This returned 289 records spanning 2003-2026 (the 2026 slice is a partial year, since the database was queried mid-year). No further document-type, language, or date filtering was applied - all 289 retrieved records (predominantly journal articles, with some reviews, conference papers, and editorials) were retained for analysis. Because the search was identifier-based rather than a topic Boolean string, the resulting corpus is best understood as a real, densely co-authored pharmaceutical-sciences research programme rather than a field-wide topic search; this paper analyzes its internal thematic structure on those terms, and the implications of this scope choice for generalizability are discussed explicitly in Section 5.3 (Limitations).

The exported fields included title, authors, year, source title, citation count, DOI, affiliations, abstract, Author Keywords, and Index Keywords, but did not include a References/cited-references field. Because co-citation network analysis requires a populated references field, that analysis type was not possible for this corpus; the study therefore relies on keyword co-occurrence analysis only, drawing on the combined Author Keywords and Index Keywords fields, which were populated for 267 and 245 of the 289 documents respectively.

Analysis Tools and Metrics

Keyword co-occurrence networks were built with a Python-based analysis pipeline (network for graph construction and Louvain community detection, scikit-learn for TF-IDF-based cluster labelling, matplotlib for visualization) applied directly to the parsed Scopus export. This pipeline is conceptually equivalent to the co-word mapping performed by VOSviewer or CiteSpace, though it is a custom implementation rather than either of those specific software packages. Nodes represent individual keywords (document frequency ≥ 2); edges connect keywords that co-occur on the same document, weighted by co-occurrence frequency. The network was capped at the 300 most frequent keyword nodes to keep the resulting figure legible.

Community structure was detected with the Louvain algorithm, and cluster quality was summarized with two standard metrics. Modularity (Q) measures how much more densely connected nodes are within their assigned cluster than would be expected by chance, on a scale where higher values indicate a more cleanly separated cluster structure; values above roughly 0.3 are conventionally treated as the threshold for a meaningfully modular network. Mean silhouette (S) measures internal cluster coherence - how tightly each cluster's members hang together relative to their separation from other clusters - on a -1 to 1 scale, with higher values indicating more internally consistent, credible clusters. Each cluster was also auto-labelled using a combination of its most frequent member keywords and TF-IDF-weighted terms drawn from associated titles/abstracts, with a fixed boilerplate stop-list (e.g., "human," "article," "nonhuman" - generic MEDLINE/Emtree indexing terms present on nearly every biomedical record) excluded from the substantive labelling, though these generic terms remain visible as high-frequency network nodes in their own right.

Keyword bursts - terms whose usage frequency surges sharply within a specific time window rather than growing steadily - were detected using a simplified Kleinberg-style burst-detection score computed from each keyword's year-by-year frequency profile. A burst is characterized by a begin year, an end year, and a strength score (higher values indicating a sharper, more concentrated surge); this mirrors the standard citation/keyword-burst tables reported in CiteSpace-based scientometric papers.

RESULTS

Descriptive Overview and Publication Trends

The corpus comprised 289 documents published between 2003 and 2026, accumulating 10,617 citations in total (≈36.7 citations per document, though this average is inflated by a small number of highly cited papers and depressed by the necessarily low citation counts of the most recent, 2025-2026 output) Figure 1 plots annual publication counts (bars) against annual citation counts (line). Table 1 lists the ten most prolific authors in the corpus, indicating their document counts and total citation impact.

Output was low and irregular through the 2000s (one to two documents most years, with a first local peak of 10 documents in 2009), then rose steadily from the late 2010s onward. Publications increased from 10 in 2018 to 24 in 2019 - a 140% year-over-year increase - and continued rising to 37 in 2020 (+54.2%) and a corpus-wide peak of 40 documents in 2021 (+8.1%). Output then declined from 39 documents in 2022 to 17 in 2023 (-56.4%), and

continued easing to 16 in 2024 and 14 in 2025. The 15 documents recorded for 2026 reflect only a partial year of coverage through the July 2, 2026 retrieval date and should not be read as a rebound; the apparent volatility across 2023-2026 is at least partly an artifact of incomplete recent-year indexing rather than a genuine contraction in research activity.

Citations track a broadly similar but noisier pattern, peaking at 1,754 citations in 2020 and 1,751 in 2021 - the two years of highest publication output - before falling off sharply for 2023-2026, consistent with the well-known citation time-lag effect in which very recent papers have not yet had time to accumulate citations.

Table 2 lists the ten most prolific affiliations. The Department of Pharmaceutical Sciences, College of Clinical Pharmacy, King Faisal University (Al-Ahsa, Saudi Arabia) leads, appearing (across two slightly differently formatted address strings, likely reflecting a postal-code addition partway through the period) on a combined 189 documents; Gulf Medical University (Ajman, UAE) and Minia University (Egypt) follow. Affiliations cluster

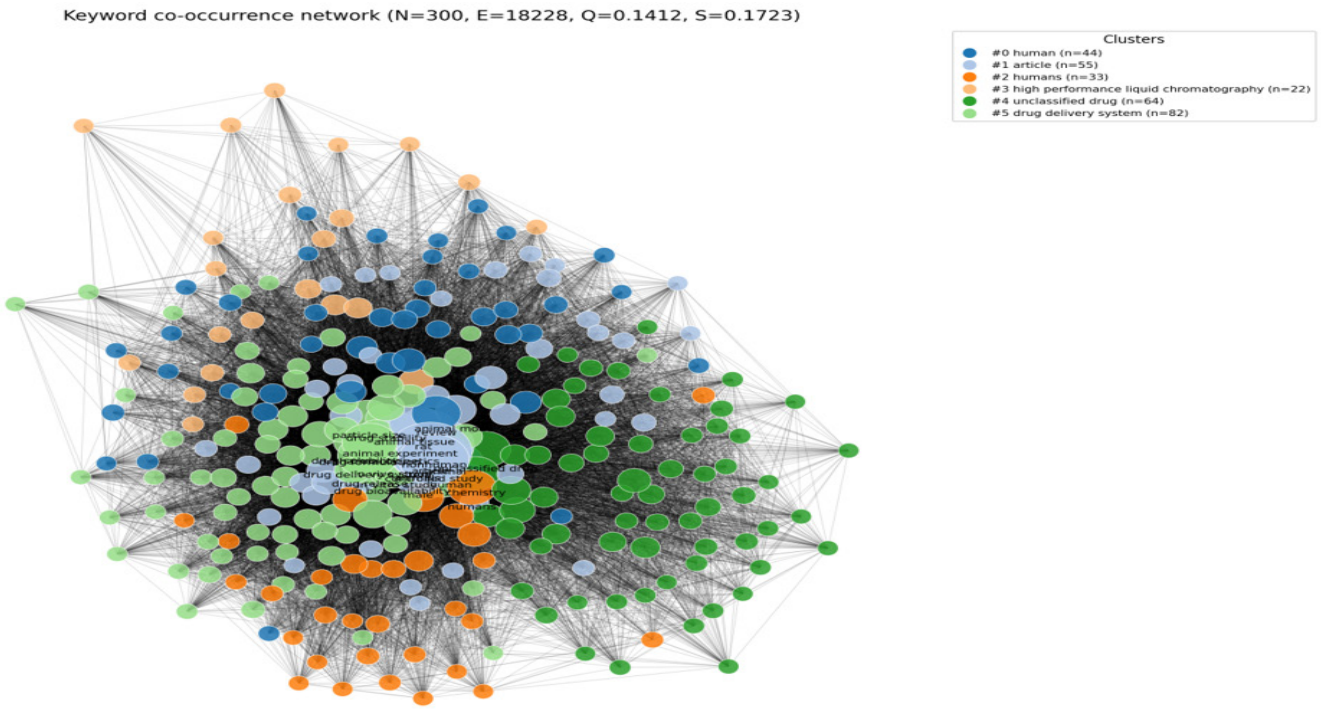

Figure 1: Annual number of publications (bars, left axis) and citations (line, right axis), 2003-2026. The 2026 bar reflects a partial year of Scopus coverage through the July 2, 2026 retrieval date.

Table 1: Lists the ten most prolific authors in the corpus. Author A.B. Nair leads with 233 documents and 8,700 total citations - expected given the identifier-based search strategy (Section 3.2) - followed by M. Attimarad (94 documents, 2,958 citations) and K.N. Venugopala (85 documents, 2,453 citations).

Rank	Author	Documents	Total Citations
1	Nair A.B.	233	8,700
2	Attimarad M.	94	2,958
3	Venugopala K.N.	85	2,453
4	Jacob S.	71	4,450
5	Sreeharsha N.	68	2,575
6	Morsy M.A.	66	2,653
7	Shinu P.	54	1,926
8	Al-Dhubiab B.E.	49	1,963
9	Shah J.	46	3,066
10	Nair A.	45	1,638

Note: Nair A.B. and Nair A. are recorded as distinct author entries in the raw Scopus author-name field and are reported here exactly as parsed; the latter likely reflects inconsistent name-form indexing rather than a fully independent author.

Table 2: Ten most prolific affiliations.

Rank	Affiliation	Documents	Total Citations
1	Dept. of Pharmaceutical Sciences, College of Clinical Pharmacy, King Faisal University, Al-Ahsa, Saudi Arabia	106	4,320
2	Dept. of Pharmaceutical Sciences, College of Clinical Pharmacy, King Faisal University, Al-Ahsa, Saudi Arabia	83	3,018
3	Dept. of Pharmaceutical Sciences, College of Pharmacy, Gulf Medical University, Ajman, UAE	44	2,374
4	Dept. of Pharmacology, Faculty of Medicine, Minia University, El-Minia, Egypt	34	1,336
5	Dept. of Biomedical Sciences, College of Clinical Pharmacy, King Faisal University, Al-Ahsa, Saudi Arabia	33	1,650
6	Dept. of Pharmaceutics, Vidya Siri College of Pharmacy, Bangalore, India	32	1,652
7	Dept. of Biotechnology and Food Technology, Durban University of Technology, Durban, South Africa	26	814
8	Dept. of Pharmaceutics, Institute of Pharmacy, Nirma University, Ahmedabad, India	20	1,016
9	Dept. of Biomedical Sciences, College of Clinical Pharmacy, King Faisal University, Al-Ahsa, Saudi Arabia	20	326
10	Dept. of Pharmacology, Faculty of Medicine, Minia University, El-Minia, Egypt	19	929

heavily in Saudi Arabia, the UAE, Egypt, and India, consistent with independently published bibliometric evidence of sustained growth in Gulf-region pharmaceutical-science output over the same period (Rohra and Memon, 2025).

Network Structure

The keyword co-occurrence network comprised N = 300 nodes and E = 18,228 edges, for a density of 0.406. This is notably high for a co-occurrence network of this size - well above the sparse (density < 0.1) structure typical of topic-defined literature searches - and reflects the fact that a large share of the 300 most frequent terms are generic MEDLINE/Emtree indexing tags (e.g., "human," "article," "nonhuman," "unclassified drug") that co-occur on nearly every biomedical record in the corpus, in addition to more substantive, cluster-differentiating terms.

Louvain community detection resolved six clusters, with a modularity of Q = 0.141 and a weighted mean silhouette of S = 0.172. Both values fall below the conventional Q > 0.3 threshold and the S > 0.3 threshold typically associated with a confidently modular, internally coherent structure; per standard scientometric interpretation, this indicates a relatively diffuse network structure rather than sharply separated thematic clusters. This is an honest and, on reflection, expected outcome given the corpus's high density and heavy generic-term overlap: with so many keywords shared across nearly every document, the boundaries between formulation, analytical, and preclinical sub-themes are blurred at the network level even though, as the cluster-by-cluster discussion in Table 3 shows, each cluster's non-generic member terms still describe a recognizably distinct sub-topic. The clustering results below should therefore be read as a useful organizing lens on the corpus rather than as a claim of statistically sharp thematic separation (Figure 2).

Cluster-by-Cluster Discussion

Cluster 5 - Drug-delivery-system formulation and release (size = 82, mean year = 2018.8, S = 0.145). This is the largest cluster, built around the terms drug delivery system, *in vitro* study, drug release, drug formulation, particle size, and drug bioavailability. It represents the corpus's core formulation-science theme - characterizing how a candidate drug is released from and absorbed out of a delivery vehicle. Representative high-citation documents plausibly associated with this cluster include reviews of nanosuspension-based drug delivery systems (387 citations) and of cyclodextrin complexes from a formulation and drug-delivery perspective (229 citations), both broad formulation-science reviews consistent with this cluster's terms. Its near-2019 mean year and large size together suggest a long-running, still-active core theme rather than either a purely legacy topic or a brand-new one.

Cluster 4 - Computational and mechanistic drug chemistry (size = 64, mean year = 2019.2, S = 0.051). Beyond its generic top label ("unclassified drug," a MEDLINE indexing tag), this cluster's substantive member terms - chemistry, metabolism, drug effect, molecular docking, and drug screening - point to a computational and mechanistic drug-chemistry theme, encompassing molecular docking and structure-based screening work alongside metabolic-fate studies. A representative document is a study screening and molecular-docking novel benzothiazole

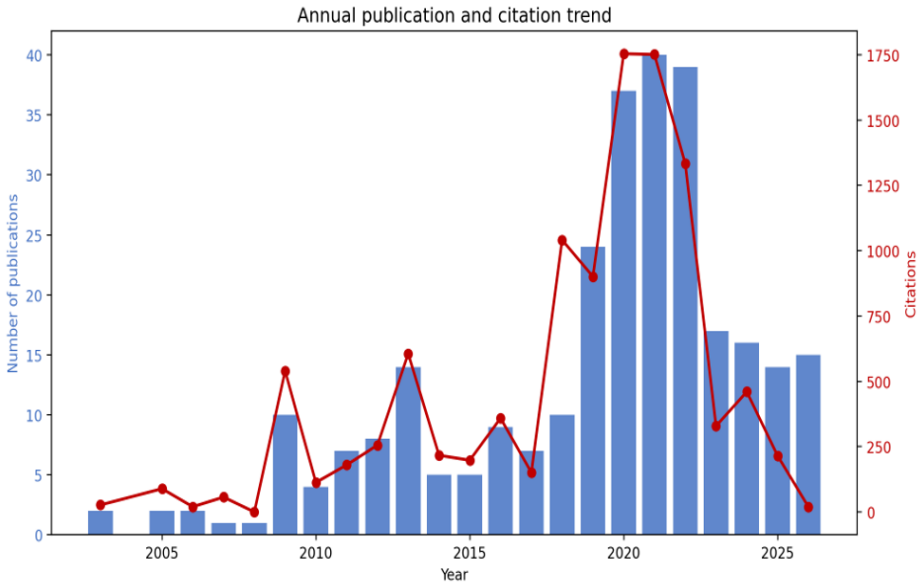

Figure 2: Keyword co-occurrence network (2003-2026). Node size is proportional to keyword frequency; color denotes cluster membership. N = 300, E = 18,228, density = 0.406, Q = 0.141, S = 0.172.

Table 3: Summarizes all six clusters, sorted by size (member-term count). Because the two or three most frequent terms in several clusters are generic indexing boilerplate ("human," "article," "humans," "unclassified drug"), the descriptive labels below are drawn from each cluster's next most frequent, substantively meaningful member terms rather than from the single top TF-IDF term alone, following standard practice when boilerplate dominates raw term frequency.

Cluster-ID	Size	Silhouette	Mean Year	Label (from member terms)
5	82	0.145	2018.8	Drug-delivery-system formulation and release
4	64	0.051	2019.2	Computational and mechanistic drug chemistry
1	55	0.490	2018.8	Preclinical / animal pharmacology studies
0	44	0.046	2019.5	Nanocarrier-based delivery and inflammation
2	33	0.091	2018.1	Clinical / transdermal-iontophoretic delivery
3	22	0.209	2018.0	Analytical method development (HPLC)

derivatives as candidate antimicrobial agents (87 citations). This cluster's very low silhouette (S = 0.051) signals that its member terms are only weakly self-consistent relative to other clusters - unsurprising, since "metabolism," "drug effect," and "molecular docking" span quite different methodological registers even though they co-occur frequently across the corpus.

Cluster 1 - Preclinical / animal pharmacology studies (size = 55, mean year = 2018.8, S = 0.490). Anchored by nonhuman, controlled study, animal experiment, and rat, this cluster captures the corpus's preclinical pharmacology and pharmacokinetics work - *in vivo* dosing, absorption, and safety studies conducted in animal models ahead of clinical translation. It has by far the highest silhouette of any cluster (S = 0.490, the only cluster clearing the S > 0.3 "homogeneous clustering" threshold), indicating this is the most internally coherent thematic grouping in the network; preclinical-study terminology evidently co-occurs in a more self-contained way than the formulation- and chemistry-oriented vocabulary elsewhere in the corpus. A directly relevant, highly cited document in this space is a review on dose translation between laboratory animals and humans across preclinical and clinical drug-development phases (432 citations), which is itself one of the corpus's own most-cited papers.

Cluster 0 - Nanocarrier-based delivery and inflammation (size = 44, mean year = 2019.5, S = 0.046). Once the generic top terms ("human," "review") are set aside, this cluster's substantive terms – nano emulsion, inflammation, nanoparticle, and nanocarrier - identify a nanocarrier-and-inflammation delivery theme, overlapping with but distinguishable from Cluster 5's broader formulation focus by its specific emphasis on nano-scale carriers and inflammatory-disease targets. This is also the cluster with the single most recent mean year (2019.5) of the six, consistent with nanocarrier delivery being this corpus's most emergent formulation sub-theme relative to the others. A representative high-citation document is a broad review of hydrogels in drug delivery, tissue engineering, and wound management (453 citations, the single most-cited document in the corpus).

Cluster 2 - Clinical / transdermal-iontophoretic delivery (size = 33, mean year = 2018.1, S = 0.091). Beyond the generic "humans"/"priority journal" top terms, this cluster's

Table 4: Lists the ten strongest keyword bursts detected in the corpus, sorted by strength. "Begin" and "End" denote the window during which the term's usage surged; "Year" is the burst's peak year.

Begin	End	Strength	Peak Year	Keyword
2009	2009	15.00	2009	hydrogen-ion concentration (pH)
2021	2022	10.29	2021	glucose
2021	2021	9.00	2021	area under curve
2009	2009	8.80	2009	atenolol
2009	2011	8.67	2009	terbinafine
2009	2009	8.57	2009	human tissue
2009	2010	7.50	2009	nail
2025	2026	7.33	2025	skin permeability
2023	2026	7.11	2025	lipid
2019	2023	7.00	2020	chemical structure

substantive vocabulary - drug delivery systems, adult, female, and iontophoresis - marks it as the corpus's clinical and physical-enhancement transdermal-delivery theme, distinct from Cluster 0's nanocarrier focus in emphasizing electrically or chemically enhanced skin permeation in human (rather than purely *in vitro* or animal) study populations. A representative document is a study of ungual and trans-ungual iontophoretic delivery of terbinafine for onychomycosis treatment (79 citations) - a paper that also anchors one of the strongest keyword bursts discussed in Table 4.

Cluster 3 - Analytical method development, HPLC (size = 22, mean year = 2018.0, S = 0.209). The smallest cluster, built around high performance liquid chromatography, limit of quantitation, limit of detection, drug determination, and validation, represents the corpus's dedicated analytical-chemistry theme - developing and validating chromatographic assays for quantifying drugs in formulations or biological matrices. Its mean year (2018.0) is the earliest of the six clusters, suggesting this is comparatively the most established, longest-running theme in the corpus rather than a recent addition, even though (as Section 5 discusses) recent external literature shows this sub-field itself evolving quickly toward green-chemistry-oriented method design. A representative document is a paper on UPLC as a preeminent technique in pharmaceutical analysis (93 citations).

Burst Analysis

The single strongest burst - hydrogen-ion concentration (strength = 15.0, confined entirely to 2009) - coincides with a cluster of 2009 iontophoretic and transdermal-transport studies in the corpus, including the atenolol-delivery-through-skin study (strength-8.8 burst) and the terbinafine ungual-iontophoresis study (strength-8.67 burst, running 2009-2011); pH is a routine but critical experimental variable in iontophoretic drug-transport work, so this burst most likely reflects a concentrated wave of physicochemical-transport methodology papers published that year rather than a substantive shift in what pH itself means to the field. The nail and human tissue bursts (both 2009-2010) point to the same cluster of ungual/trans-ungual delivery studies.

A distinct, more recent burst cluster appears around 2021: glucose (strength = 10.29) and area under curve (strength = 9.00) both peaked that year, plausibly reflecting a concentrated set of pharmacokinetic or metabolic-monitoring studies published in 2021 - though, consistent with the general caution around burst interpretation, this correlation in time does not by itself establish what specifically drove the surge without reading the underlying papers directly. Most recently, skin permeability (2025-2026, strength = 7.33) and lipid (2023-2026, strength = 7.11) are the corpus's two most current bursts, both still open at the July 2026 retrieval date, suggesting continued or growing attention to lipid-based and permeability-focused delivery work in the corpus's newest publications - a pattern that aligns with Cluster 0's status as the cluster with the most recent mean year.

DISCUSSION

Read together, the descriptive, network, and cluster results sketch a corpus that grew from sparse, methodologically narrow beginnings (2003-2017, rarely more than 10 documents a year, dominated by the transdermal/iontophoresis and early analytical-method themes reflected in the 2009 burst cluster) into a much larger and more thematically varied output stream from 2018 onward, peaking in 2020-2021 and touching all six identified clusters simultaneously. The clusters' mean years are tightly bunched (2018.0-2019.5), which is itself informative: rather than a clean early-foundational → expansion → current-integration arc with clusters strung out across the full 2003-2026 span, most of this corpus's thematic diversification happened within a compressed few-year window in the late 2010s, consistent with the sharp publication-count jump documented in Section 4.1 (140% growth from 2018 to 2019 alone). The Cluster-3 analytical-method theme (mean year 2018.0) is marginally the earliest-centered, and the Cluster-0 nanocarrier theme (mean year 2019.5) marginally the most recent, but the gap between

them is under a year and a half - not the multi-year separation seen in scientometric studies of longer-running, more slowly evolving fields.

This compressed, low-modularity structure (Q = 0.141) has a substantive as well as a methodological reading. Substantively, it suggests that in this corpus, formulation science, analytical chemistry, preclinical pharmacology, and computational chemistry are not run as separate, loosely connected research streams but are woven together within individual papers - a formulation study will often also validate an HPLC assay for the same compound, or dose-translate a nanocarrier finding from rat to human, which is exactly the kind of cross-theme keyword co-occurrence that suppresses modularity in a network sense while reflecting genuine methodological integration in a substantive sense.

The two most recent clusters by mean year - Cluster 0 (nanocarrier-based delivery, 2019.5) and Cluster 4 (computational drug chemistry, 2019.2) - also align well with where current external literature shows the fastest independent growth. On the nanocarrier side, 2025 reviews describe nanoemulsion-based delivery expanding rapidly across oral, ocular, nasal, and intra-articular routes over just the last five years, with intranasal central-nervous-system applications singled out as a particularly active frontier (Chatzidaki and Mitsou, 2025; a dedicated 2025 bibliometric study of intranasal nanoemulsions alone identified 379 manuscripts). On the computational-chemistry side, molecular docking has been described in a 2024 Annual Review of Biochemistry synthesis as an increasingly indispensable, rapidly advancing tool, now routinely paired with machine-learning-guided screening of vast virtual compound libraries (Paggi *et al.*, 2024) - a trajectory this corpus's Cluster 4 (chemistry, molecular docking, drug screening) sits squarely within, even though the corpus's own instances of this theme predate the most recent AI-docking wave.

The corpus's smallest and earliest-centered cluster, analytical method development (Cluster 3, HPLC/validation), is worth flagging as a specific area where the corpus may be lagging a fast-moving external field rather than leading it. Recent analytical-chemistry literature has shifted substantially toward green analytical chemistry - explicitly scoring HPLC and LC-MS methods against solvent-consumption and environmental-impact metrics such as the AGREE score (Nassar *et al.*, 2025) - a framing that is not evident among this cluster's member terms (limit of quantitation, limit of detection, validation), which read as conventional ICH-guideline method-validation vocabulary rather than green-chemistry vocabulary. This is a concrete, data-grounded example of a gap between the corpus's own analytical-method theme and where that adjacent sub-field has moved externally in the last two to three years.

Geographically and institutionally, the corpus is concentrated in a small number of Gulf-region and South Asian institutions, a pattern consistent with independently documented growth in Saudi Arabian pharmaceutical-science publication volume and journal quality over four decades (Rohra and Memon, 2025) and with broader evidence of rising, internationally collaborative Gulf-region research output more generally. This concentration is a natural feature of an identifier-based corpus built around one core research group's collaborator network rather than a limitation to correct, but it does mean the thematic map above should be read as characterizing that group's own research programme rather than the pharmaceutical-sciences field worldwide.

Theoretical and Practical Implications

Theoretically, the compressed cluster-year spread and low modularity observed here suggest that, at least for actively cross-disciplinary pharmaceutical-science research programmes, keyword co-occurrence clustering may be better suited to identifying which sub-themes exist and how central each is by size than to producing sharply separated, temporally staged "field evolution" narratives - a methodological point future scientometric studies of similarly integrated (as opposed to loosely federated) research corpora should anticipate rather than treat as an analysis failure.

Practically, for researchers and administrators evaluating or planning work in this space, the results point to formulation/drug-release science (Cluster 5) and nanocarrier-based delivery (Cluster 0) as the corpus's largest and most recent areas of strength, computational drug chemistry (Cluster 4) as a fast-growing area worth continued investment given its alignment with external AI-docking trends, and green analytical chemistry as a specific, currently under-represented niche where the corpus's existing HPLC-method-development expertise (Cluster 3) could plausibly be extended to track a clear external methodological shift.

Limitations

Several limitations should be considered when interpreting these findings. First, the corpus was assembled via a Scopus author-identifier search rather than a topic-level Boolean query; while this paper deliberately frames and discusses the resulting corpus at the topic/thematic level (per the scope agreed for this analysis), the underlying sample is nonetheless bounded by one author's and their collaborators' publication record, not by an exhaustive search of the pharmaceutical-sciences literature, and the results should not be generalized as representative of the field as a whole. Second, the export lacked a References field, so citation-based co-citation network analysis - which typically complements keyword co-occurrence analysis in this genre by identifying foundational cited works rather than just co-occurring terms - was not possible; only keyword co-occurrence analysis is reported. Third, the resulting network showed unusually high density (0.406) driven substantially by generic MEDLINE/Emtree

indexing terms, and correspondingly low modularity (Q = 0.141) and mean silhouette (S = 0.172); these values are honestly below the conventional thresholds for a confidently modular, highly coherent clustering, and the cluster interpretations offered in Section 4.3 should be read as a useful organizing heuristic rather than as statistically sharp, unambiguous thematic boundaries. Fourth, Scopus-only coverage and an English-language bias in indexing are standard limitations of any single-database bibliometric study. Fifth, the most recent years (2023-2026) are subject to citation time-lag bias, and 2026 in particular reflects only a partial year of coverage.

CONCLUSION

This scientometric analysis mapped 289 Scopus-indexed pharmaceutical-sciences documents (2003-2026, 10,617 total citations) into a six-cluster keyword co-occurrence network. Three headline findings stand out. First, output grew sharply and diversified simultaneously in the late 2010s, with annual publication counts rising 140% from 2018 to 2019 and peaking at 40 documents in 2021, rather than following a slow, cluster-by-cluster evolutionary path. Second, the corpus's largest themes are drug-delivery-system formulation/release and computational/mechanistic drug chemistry (82 and 64 member terms respectively), while its most internally coherent theme by far is preclinical animal pharmacology (silhouette S = 0.490, more than double any other cluster). Third, keyword bursts cluster into two distinct waves - a 2009 wave tied to iontophoretic/ungual transdermal-transport methodology, and a currently open (2023-2026) wave centered on lipid-based and skin-permeability delivery work, suggesting where the corpus's newest attention is concentrated.

Three future-research directions follow directly from gaps the analysis surfaced rather than from generic recommendation. First, the analytical-method cluster's comparatively conventional, non-green-chemistry vocabulary alongside a clearly documented external shift toward AGREE-scored, solvent-minimizing HPLC methods (Nassar *et al*., 2025) suggests green analytical chemistry as a concrete, evidence-grounded extension opportunity. Second, the corpus's absence of a usable references field prevented co-citation analysis in this study; a future re-analysis using an export with a populated References/CR field could add a complementary foundational-works map alongside the keyword-level view reported here. Third, given the low overall network modularity, a follow-up analysis restricting the network to only non-generic, substantive keywords (excluding MEDLINE/Emtree boilerplate terms such as “human” and “article” from node selection, not just from labelling) may recover a more sharply modular, higher-silhouette cluster structure and is a natural methodological refinement for future scientometric work on similarly densely-indexed biomedical corpora.

CONFLICT OF INTEREST

The authors declare that there is no conflict of interest.

REFERENCES

Al-Dhubiab, B. E., *et al.* (2021). Emerging role of hydrogels in drug delivery systems, tissue engineering and wound management. Pharmaceutics, 13(3), 357. https://doi.org/10.3390/pharmaceutics13030357

Attimarad, M., *et al.* (2020). Emerging role of nanosuspensions in drug delivery systems. Journal of Biological Engineering, 14, Article 8. https://doi.org/10.1186/s40824-020-0184-8

Chatzidaki, M. D., and Mitsou, E. (2025). Advancements in nanoemulsion-based drug delivery across different administration routes. Pharmaceutics, 17(3), 337. https://doi.org/10.3390/pharmaceutics17030337

Deepika, D. M., Reddy, M. S., and Vijetha, K. A. (2025). A review on advances of nanoemulsions in drug delivery and therapeutics. International Journal of Pharmaceutical Sciences, 3(9), 2657-2668. https://doi.org/10.5281/zenodo.17183059

Donthu, N., Kumar, S., Mukherjee, D., Pandey, N., and Lim, W. M. (2021). How to conduct a bibliometric analysis: An overview and guidelines. Journal of Business Research, 133, 285-296.

Dwivedi, S., Makode, K., Mourya, P., Sharma, P. K., and Darwhekar, G. N. (2025). Nanoemulsions: A comprehensive review of drug delivery systems. European Journal of Pharmaceutical and Pharmacological Sciences, 30(1), Article 8. https://doi.org/10.37521/ejpps30108

Franz, M., Jairam, R. K., Kuepfer, L., and Hanke, N. (2024). PBPK-based translation from preclinical species to humans for the full-size IgG therapeutic efalizumab. Frontiers in Pharmacology, 15, Article 1418870.

Gautam, S., Pathak, S., and Dubey, S. H. (2024). The role of molecular docking in modern drug discovery and development: A comprehensive review. Journal of Drug Discovery and Health Sciences, 1(3), 129-137.

Gross, I. P., Lima, A. L., Sa-Barreto, L. L., Gelfuso, G. M., and Cunha-Filho, M. (2025). Recent advances in cutaneous drug delivery by iontophoresis. Expert Opinion on Drug Delivery, 22(6), 857-874. https://doi.org/10.1080/17425247.2025.2490267

Jacob, S., *et al.* (2018). Cyclodextrin complexes: Perspective from drug delivery and formulation. Drug Development Research, 79(5), 201-217. https://doi.org/10.1002/ddr.21452

Jacob, S., *et al.* (2013). *In vitro* techniques to evaluate buccal films. Journal of Controlled Release, 166(1), 10-21. https://doi.org/10.1016/j.jconrel.2012.11.019

Muhammed, M. T., and Aki-Yalcin, E. (2024). Molecular docking: Principles, advances, and its applications in drug discovery. Letters in Drug Design and Discovery, 21, 480-495.

Morsy, M. A., *et al.* (2018). Dose translation between laboratory animals and human in preclinical and clinical phases of drug development. Drug Development Research, 79(8), 373-382. https://doi.org/10.1002/ddr.21461

Nassar, M. W., Serag, A., Hasan, M. A., and Kamel, M. (2025). Development and validation of a RP-HPLC method for simultaneous determination of five COVID-19 antiviral drugs in pharmaceutical formulations. Scientific Reports, 15, Article 09904.

Nair, A. B., *et al.* (2009). Ungual and trans-ungual iontophoretic delivery of terbinafine for the treatment of onychomycosis. Journal of Pharmaceutical Sciences, 98(2), 464-472. https://doi.org/10.1002/jps.21711

Nair, A. B., *et al.* (2009). Delivery of a classical antihypertensive agent through the skin by chemical enhancers and iontophoresis. Skin Research and Technology, 15(3), 187-194. https://doi.org/10.1111/j.1600-0846.2009.00350.x

Nair, A. B., *et al.* (2012). UPLC: A preeminent technique in pharmaceutical analysis. Acta Chromatographica, 24(3), 517-529.

Paggi, J. M., Pandit, A., and Dror, R. O. (2024). The art and science of molecular docking. Annual Review of Biochemistry, 93, 389-410. https://doi.org/10.1146/annurev-biochem-030222-120000

ResearchAndMarkets. (2026). Drug-delivery technology market global report 2026. Research and Markets.

Rohra, D. K., and Memon, I. (2025). Four decades of progress: Evaluating Saudi Arabia's research publications in pharmaceutical sciences. Hospital Pharmacy. Advance online publication.

Shah, S. W. A., Li, X., Yuan, H., Shen, H., Quan, S., Pan, G., Ishfaq, M., Shah, A. U., Xie, H., and Shao, J. (2025). Innovative transdermal drug delivery systems: Benefits, challenges, and emerging application. BMEMat, e70001. https://doi.org/10.1002/bmm2.70001

Shinu, P., *et al.* (2020). Screening and molecular docking of novel benzothiazole derivatives as potential antimicrobial agents. Antibiotics, 9(5), 221. https://doi.org/10.3390/antibiotics9050221

Wang, Y., Zeng, L., Song, W., and Liu, J. (2022). Influencing factors and drug application of iontophoresis in transdermal drug delivery: An overview of recent progress. Drug Delivery and Translational Research, 12(1), 15-26.

Cite this article: Ahmed KKM, Sab CM, Yunus M, Jacob S, Shah JN, Shah HJ, *et al*. Thematic Structure and Evolution of a Pharmaceutical Sciences Research Corpus from Dr. Anroop Nair. Int. J. Pharm. Investigation. 2026;16(4):1657-65.