



Bridging Research and Innovation in Endoscopic Submucosal Dissection: A Publication-Patent Analysis Revealing Systematic Clinical-Technical Divergence

İ Hüseyin DEMİR¹, İ Serhat BURMAOĞLU², İ Tuncay AKAR³, İ Adnan YAMANOĞLU⁴, İ Sezgin VATANSEVER⁵

¹İzmir Katip Çelebi University Faculty of Economics and Administrative Sciences, Department of Health Management, İzmir, Türkiye

²İzmir Katip Çelebi University Faculty of Economics and Administrative Sciences, Department of Data Science and Analytics, İzmir, Türkiye

³Muğla Sıtkı Koçman University, Institute of Health Sciences, Department of Health Management, Muğla, Türkiye

⁴İzmir Katip Çelebi University Faculty of Medicine, Division of Internal Medical Sciences, Department of Emergency Medicine, İzmir, Türkiye

⁵İzmir Katip Çelebi University Faculty of Medicine, Department of Internal Medical Sciences, İzmir, Türkiye

Cite as: Demir H, Burmaoğlu S, Akar T, Yamanoğlu A, Vatansever S. Bridging research and innovation in endoscopic submucosal dissection: a publication-patent analysis revealing systematic clinical-technical divergence. Forbes J Med 2026;7:128-39

ABSTRACT

Objective: Clinical research and technological innovation are assumed to evolve in tandem through market-mediated knowledge diffusion, yet systematic evidence for their alignment in medical device domains remains limited. This study examines the relationship between the production of scientific evidence and the commercial development of devices for endoscopic submucosal dissection, a minimally invasive treatment for gastrointestinal neoplasms, which has experienced uneven global diffusion.

Methods: We employed a dual-corpus computational approach to analyze a stratified sample of 2,470 publications [60% highly cited, 20% recent (2023-2025), and 20% randomly selected] and a corpus of 4,526 device patents (2010-2025), using Word2Vec embeddings and Latent Dirichlet Allocation to identify thematic structures. Cross-corpus semantic similarity measured the alignment between clinical research priorities and commercial development trajectories. Results indicate a substantial publication-patent divergence (divergence score =0.792), with no topic pair exceeding 0.5 in semantic similarity (an original, study-specific metric that has not yet been externally validated).

Results: Most publication topics (88.9%) lack corresponding patent activity, while most patent topics (83.3%) develop independently of academic discourse. Seven innovation gaps emerged in which well-documented clinical complications lack technological solutions, particularly in the management of delayed complications, the prevention of strictures, and the control of pain. Temporal analysis showed an inverse relationship between patenting and publication activity ($r=-0.949$, lag=-5 years), but this finding should be interpreted cautiously given the small number of annual observations ($n=16$) and the absence of multiple-lag correction; Granger causality tests indicated no significant bidirectional influence. Only 16.1% of patents demonstrate a high citation impact; these are predominantly in materials science.

Conclusion: The findings challenge assumptions of efficient knowledge diffusion in medical device markets, revealing parallel knowledge streams with weak connectivity. The dual-corpus framework offers a generalizable methodology for assessing research-innovation integration across medical specialties.

Keywords: Endoscopic submucosal dissection, knowledge diffusion, patent-publication analysis, medical device innovation, clinical-technological alignment

Received: 24.07.2026

Accepted: 16.09.2026

Publication Date: 02.10.2026

Corresponding Author:

Hüseyin DEMİR

İzmir Katip Çelebi University Faculty
of Economics and Administrative
Sciences, Department of Health
Management, İzmir, Türkiye

✉ huseyin.demir@ikc.edu.tr

ORCID ID: 0000-0002-8990-7228

INTRODUCTION

Endoscopic submucosal dissection (ESD), developed in Japan in the late 1990s, represents a paradigm shift in minimally invasive treatment of gastrointestinal neoplasms.^{1,2} The technique overcomes limitations of endoscopic mucosal resection, particularly the piecemeal resection required for lesions exceeding 20 mm and the associated high recurrence rates.³⁻⁵ ESD enables

en bloc resection of large superficial lesions while preserving organ function, offering complete resection rates exceeding 90% for early gastric neoplasms, with 5-year survival approaching those of surgical resection.⁶⁻⁸ Despite growing acceptance as a standard treatment for superficial gastric neoplasms, esophageal squamous cell carcinoma, and select colorectal lesions, the diffusion of ESD beyond Japan has been uneven.⁹⁻¹¹



Copyright© 2026 The Author(s). Published by Galenos Publishing House on behalf of Buca Seyfi Demirsoy Training and Research Hospital. This is an open access article under the Creative Commons AttributionNonCommercial 4.0 International (CC BY-NC 4.0) License.

This heterogeneity reflects patterns documented in knowledge diffusion theory: geographical localization of tacit expertise, knowledge filters impeding commercialization, and institutional mechanisms constraining technology transfer.^{12,13} ESD's technical complexity generates substantial tacit knowledge that resists codification and is transmitted primarily through hands-on training.¹⁴ Technical demands are reflected in complication rates of 3.5% for gastric, 3.3% for esophageal, and 4.6% for colorectal procedures, with operator experience critically determining safety.^{15,16} In parallel, a robust commercial ecosystem has emerged, featuring sophisticated instrumentation: electrosurgical knives, injection systems, and hemostatic devices.¹⁷ Medical device innovation, characterized by iterative cycles between clinical use and technological refinement and involving complex interactions among physicians, engineers, regulatory bodies, and manufacturers, differs fundamentally from pharmaceutical development.¹⁸⁻²⁰ This co-evolutionary dynamic suggests that alignment between clinical evidence and commercial device development should emerge naturally, yet evidence from medical device markets reveals frequent misalignment.^{21,22}

Integrated patent-publication databases now provide an unprecedented opportunity to systematically examine clinical-technical alignment. Combined with natural language processing and unsupervised machine learning, these databases enable innovation intelligence that transcends single-corpus limitations.²³ Despite a rich literature on ESD clinical outcomes, comprehensive macro-level analysis connecting publications to patents remains limited. This study addresses clinical-technical integration in ESD through three research questions, the first of which is: (RQ1) To what extent do clinical research themes align with commercial device development priorities? (RQ2) Which specific clinical complications lack corresponding technological solutions? (RQ3) What temporal relationship exists between publication activity and patenting trends, and does evidence suggest bidirectional knowledge transfer? We employ a dual-corpus computational approach, integrating 2,470 publications and 4,526 patents, to quantify alignment and identify structural knowledge filters impeding translation, thereby contributing a generalizable framework with implications for clinicians, training programs, research funders, device manufacturers, and policymakers.

METHODS

Data Acquisition

Publication and patent data were retrieved from Lens.org.²⁴ The publication dataset encompassed 17,830 peer-reviewed articles identified through queries combining “endoscopic submucosal dissection,” “ESD procedure,” and “submucosal dissection technique” with organ-specific terms (2010-2025). The patent dataset comprised 5,096 utility patents identified using technology-focused queries. A stratified sampling strategy was implemented to select publications: 60% highly cited, 20% recent

(2023-2025), and 20% randomly selected publications, yielding 2,470 documents. “Highly cited” was operationalized as the total citation count within the top decile for its publication year according to Lens.org indexing. The exact numbers of records screened into and retained within each of the three tiers, by year, and the procedure used to resolve overlap between tiers (e.g., when a record qualified as both highly cited and recent) were reported. The tiered design was informed by general bibliometric sampling conventions; however, it did not in itself validate the specific sampling design. A dedicated methodological reference on stratified or citation-tiered sampling in bibliometric research was also provided.²⁵

The exclusion workflow was documented by reporting the number of records excluded at each stage and the corresponding reasons for exclusion, including duplicate records, non-English-language records, records with unavailable abstracts, records with fewer than 20 tokens after preprocessing, and records of non-primary-research document types. For patent records, the unit of analysis was explicitly defined as granted patents. To avoid multiple counting of the same invention across jurisdictions, patent records were deduplicated, collapsing records to Lens.org simple-family identifiers where applicable.

Text Preprocessing

All documents underwent a comprehensive NLP pipeline that integrated the title and abstract fields. Research demonstrates that abstracts capture 80-90% of semantic information for topic modeling in medical domains.²⁶ The preprocessing sequence included conversion to lowercase, removal of URLs and emails, removal of numeric characters, tokenization, removal of stopwords, and WordNet lemmatization. Documents with fewer than 20 tokens were excluded.²⁷⁻²⁹

Semantic Clustering and Topic Modeling

Word2Vec models were trained separately for publications and patents using the CBOW architecture (100-dimensional vectors, context window of 5, minimum word frequency of 5).³⁰ K-means clustering was applied to document embeddings, and the optimal number of clusters was determined using the elbow method, silhouette coefficients, and Davies and Bouldin³² indices.³¹ Visualization was performed using t-SNE (perplexity =30, 1,000 iterations).³³ Probabilistic topic modeling employed latent Dirichlet allocation (LDA), with the optimal number of topics determined by maximization of the CV coherence score across k=2-12 (publications) and k=2-10 (patents).^{34,35} Final models: 9 publication topics and 6 patent topics (15 passes, 400 iterations).

Innovation Divergence Quantification

The two LDA models were fitted independently on their respective corpora (publications, patents), each producing a topic-term probability distribution over its own vocabulary. A common vocabulary was constructed (e.g., as the union or intersection

of terms retained after preprocessing in both corpora, or as the top-N TF-weighted terms per topic), and each topic's term-probability vector was renormalized to this vocabulary before computing cosine similarity, producing a 9×6 (publication-topic × patent-topic) similarity matrix, i.e., 54 pairwise comparisons. Topic labels (e.g., "procedural technique," "electrosurgical tools") were assigned by independent review of the top 15 terms per topic and were applied separately to the publication and patent models, since the two were not jointly estimated.

Cross-corpus alignment was assessed by extracting topic-term probability matrices from both LDA models, identifying shared vocabulary (typically 40-60% overlap), and computing cosine similarity for all publication-patent topic pairs. Topic pairs with similarity ≥ 0.5 were classified as strongly aligned; publication topics with maximum similarity < 0.3 were designated "orphan" topics. The divergence score was calculated as: Divergence = 1-(mean maximum similarity across all topics).

Classification into "strongly aligned" versus "orphan" categories used fixed descriptive similarity thresholds (≥ 0.5 , < 0.3) rather than a per-pair null-hypothesis significance test, so a Bonferroni correction does not directly apply to these categorical cut-offs. When inferential statistics (e.g., permutation-based p values) are also reported for any of the 54 topic-pair comparisons, a multiple-comparison correction should be applied, and the corrected results should be reported alongside the raw similarity values.

Temporal Dynamics

Annual time series spanning 2010-2025 were constructed for both corpora. Cross-correlation analysis examined lead-lag relationships across lags of -10 to +10 years on the detrended series. Granger causality testing employed a vector autoregression with lags of 1-3 years ($\alpha=0.05$). Because the underlying series comprise only 16 annual observations (2010-2025), the VAR and Granger causality estimates are based on a sample that is too small for reliable inference; this constraint constitutes a major limitation, and the results should be interpreted as exploratory rather than confirmatory. Granger causality was evaluated using the sum of squared residuals F-test applied to the detrended publication and patent series for lags 1 through 3. Because this implementation tests each specified lag independently rather than selecting a single lag via an information criterion and does not perform an internal stationarity test or post-estimation residual diagnostics, no formal stationarity or residual diagnostic tests were conducted on the underlying series in the present analysis. This is acknowledged as an additional methodological limitation, and we recommend that a formal stationarity assessment and residual diagnostics be incorporated into future work to strengthen the reliability of causal inference. The effective number of observations available at each of the three tested lags is presented, together with 95% confidence intervals for the reported correlation and F-statistics.³⁶ Compound annual growth rates (CAGR) were calculated for both time series.

Complication-Device Mapping

A complication taxonomy encompassing five categories (bleeding, perforation, stricture, infection, pain) and encompassing a device taxonomy (dissection knives, hemostasis devices, injection systems, closure devices, cautery instruments) were established. Keyword-based classification identified mentions of complications and of device types in abstracts. A complication-device solution matrix was used to cross-tabulate these categories and reveal innovation gaps. Technology maturity was assessed by patent age (< 5 , 5-10, > 10 years) and forward citation counts (low < 5 , moderate 5-10, high ≥ 10). All analyses were performed using Python 3.12 with a fixed random seed [42] for reproducibility. Informed consent was not required for this study, as it involved the bibliometric analysis of publicly available bibliographic data and did not include human participants or the collection or processing of personally identifiable information. This study was based exclusively on publicly available publication data retrieved from established bibliometric databases. No human participants were involved, and the study did not entail primary data collection or any form of intervention. As the data analyzed were already publicly accessible and did not contain identifiable personal information or sensitive data, institutional review board or ethics committee approval was not required for this bibliometric analysis.

Findings

Semantic Structure and Thematic Organization

Detailed findings regarding the semantic structure and thematic organization are presented in Figure 1a-d. Word2Vec clustering identified $k=2$ as the optimal number of clusters for publications (silhouette = 0.188) and $k=3$ as the optimal number of clusters for patents (silhouette = 0.157). Publication Cluster 0 (51.3%) emphasized procedural-technical content; Cluster 1 (48.7%) focused on oncological outcomes. Patent clustering demonstrated pronounced segregation: Cluster 0 (38.2%, mechanical innovations), Cluster 1 (33.5%, injection/lifting technologies), and Cluster 2 (28.3%, materials science/coatings). When visualized in the same semantic space, these clusters show minimal spatial overlap, providing early evidence of thematic divergence. The silhouette coefficients obtained for both clustering solutions were low (0.188 for publications; 0.157 for patents), indicating weak cluster separation in the Word2Vec embedding space. These Word2Vec + K-means clusters, therefore, provide only preliminary, low-confidence evidence of thematic structure and should be interpreted as complementary to, and subordinate to, the LDA topic model results, which form the basis of the main divergence analysis.

LDA coherence optimization identified 9 topics in publications (coherence = 0.465, $k=9$) and 6 topics in patents (coherence = 0.463, $k=6$). Figure 2a-d provide a comprehensive overview of the key topics and their associated word clouds. This dimensional asymmetry—nine clinical topics mapping onto six patent topics—represents a structural obstacle to knowledge transfer.

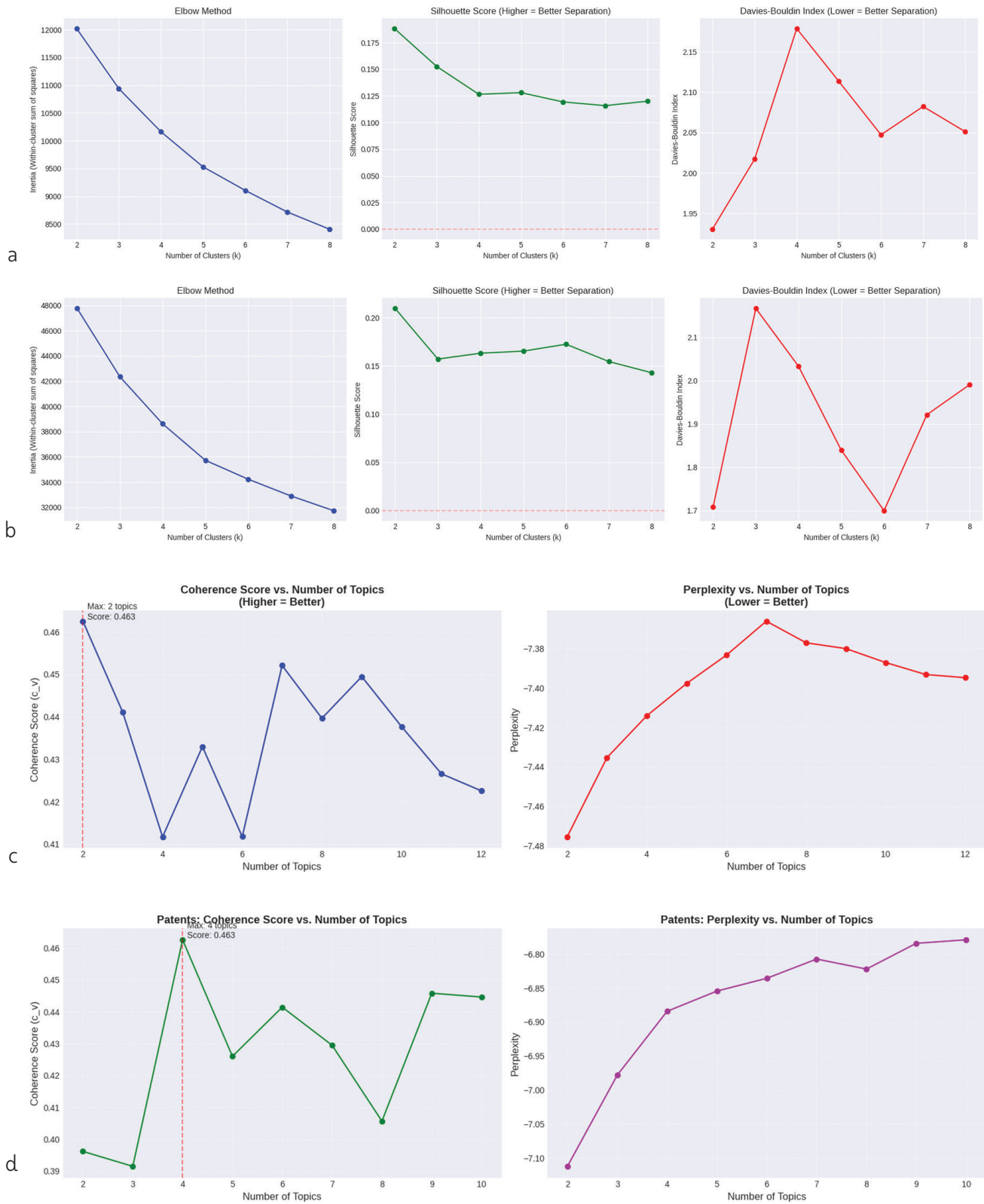


Figure 1. Cluster optimization metrics for publications (a) and patents (b); LDA topic coherence optimization curves for publications (c) and patents (d)

LDA: Latent Dirichlet allocation

Publication topics spanned procedural technique (T0), patient outcomes (T1), cancer characteristics (T2), technique refinement (T3), adverse events (T4), comparative studies (T5), training/competency (T6), anatomical applications (T7), and classification systems (T8). Patent topics covered: device mechanics (T0), electrosurgical tools (T1), material compositions (T2), injection systems (T3), imaging/navigation (T4), and closure/sealing (T5).

Innovation Landscape Divergence (RQ1)

The publication-patent alignment matrix quantified semantic similarity across all 54 possible topic pairs. Figure 3a-e provide a detailed overview of the innovation landscape divergence. No topic pair achieved strong alignment (≥ 0.5), and only two achieved moderate alignment: publication T3 (technique refinement) aligned with patent T1 (electrosurgical tools) at 0.316—the maximum observed—while publication T4 (adverse events) aligned with patent T5 (closure devices) at 0.289. Eight of nine publication topics (88.9%) showed no meaningful alignment (score < 0.3) with any patent topic, including patient outcomes (maximum 0.18), training (maximum 0.19), and classification systems (maximum 0.17). Reciprocally, five of the six patent topics (83.3%) develop independently of academic research. The divergence score of 0.792 [Divergence = $1 - (0.316 + 0.289 + \dots + 0.282) / 15 = 1 - 0.208$] is an original metric introduced in this study and has not yet been validated or benchmarked against other studies or device domains, and is descriptively consistent with a substantial clinical-technical misalignment.

Temporal Innovation Dynamics (RQ3)

Figure 4a-c illustrate the temporal dynamics of innovation in detail. ESD publications increased from 89 (2010) to a peak of 241 (2015) and then declined sharply (CAGR = -21.5%), reaching 82 by 2024—a 66% reduction from the peak. In contrast, patent filings grew steadily (CAGR = +8.9%) from 138 (2010) to 479 (2025) without reaching a plateau. The crossover point occurred in 2015; by 2025, patents outnumbered publications by 7.4:1. This inverse relationship persisted across all individual topics: all nine publication topics declined after the 2015-2018 peaks, whereas all six patent topics maintained robust growth through 2025.

Cross-correlation analysis revealed the strongest negative correlation at lag = -5 years ($r = -0.949$, $p < 0.001$). Given that only 16 annual observations underlie this analysis and that 21 lags (-10 to +10 years) were examined, this correlation should be interpreted with caution: it does not, by itself, establish that patent growth temporally precedes publication decline, and no correction for the 21 lag-wise comparisons has yet been applied. Because the corresponding Granger causality tests were not significant in both directions, this cross-correlation result is presented as a descriptive, hypothesis-generating association rather than evidence of a lead-lag causal relationship. A weak positive correlation at lag = +3 years ($r = 0.187$, $p = 0.342$) and a near-zero contemporaneous correlation ($r = -0.143$, $p = 0.476$) indicate that publication trends

do not predict patent activity. Granger causality tests yielded nonsignificant results in both directions (patents $\rightarrow$ publications: $F = 1.83$, $p = 0.17$; publications $\rightarrow$ patents: $F = 1.24$, $p = 0.31$), suggesting a shared response to external confounding factors rather than a direct causal influence.

Complication Management Technology Landscape (RQ2)

Figure 5a-e provide a detailed overview of the complication management technology landscape. The most frequently mentioned complications were bleeding (486, 19.7%) and perforation (469, 19.0%), followed by stricture (270, 10.9%), infection (194, 7.9%), and pain (64, 2.6%). Reported values reflect the number of publications in which a given complication keyword was detected. Publication-to-patent ratios revealed striking mismatches (publication mentions per corresponding patent; e.g., bleeding: 486 publication mentions vs. 136 patents ≈ 3.6 publications per patent): bleeding 3.6:1, perforation 6.2:1, stricture 2.8:1, infection 2.2:1, and pain 9.1:1. Closure devices were most prevalent ($n = 508$), followed by dissection knives ($n = 452$), injection systems ($n = 423$), hemostasis devices ($n = 222$), and cautery instruments ($n = 20$).

Seven innovation gaps were identified from the complication-device solution matrix:

Gap 1-Delayed Bleeding Management (Critical): The 12 studies targeting delayed hemorrhage lack integration with clinical prediction tools that could guide prophylactic interventions.

Gap 2-Stricture Prevention in Circumferential Resection (Critical): Eighty-seven of 98 stricture patients focus on post-stricture dilation rather than prevention. Tissue engineering approaches remain absent from the patent landscape.

Gap 3-Pain Control for Extended Procedures (Moderate-High): Only 7 patents address sedation monitoring without targeting pain sources; 64 clinical mentions produce a 9.1:1 gap ratio.

Gap 4-Submucosal Fibrosis Navigation (High): Twenty-three tissue characterization patents detect fibrosis but provide no actionable dissection solutions.

Gap 5-Perforation Risk Prediction (Moderate): Seventy-six patents address closure; only 8 address prediction. Integrated predictive models guiding real-time decision-making are absent.

Gap 6-Mucosal Defect Size Estimation (Low-Moderate): Twenty measurement patents lack decision-support algorithm integration.

Gap 7-Post-ESD Infection Management (Low): Despite the second-best ratio (2.2:1), innovations targeting specific ESD infection pathways remain absent.

These gaps reveal a systematic pattern: commercial innovation gravitates toward acute, intra-procedural, device-amenable problems while neglecting delayed complications, prevention (rather than treatment), and patient-centered endpoints.

A technology-maturity assessment revealed that 46.7% of patents were recent (<5 years), 35.9% were established (5-10 years), and 17.4% were mature (>10 years), indicating active innovation rather than market consolidation. Paradoxically, 74.4% demonstrated low citation impact (<5 forward citations), with only 16.1% achieving high impact (≥ 10 citations). High-impact patents were concentrated in materials science (48%), electrosurgical energy (31%), and mechanical innovations (21%), suggesting clinical impact depends more on enabling technology platforms than on individual device designs.

Integration and Synthesis

Three critical patterns emerge: simultaneous high divergence (0.792), declining publications (-21.5% CAGR), and rising patents (+8.9% CAGR) reveal a fundamental disconnection, wherein commercial innovation proceeds independently of academic research—these opposite-directional trends contradict knowledge diffusion theory's prediction of increasing convergence in mature domains. Second, even the most-studied complications remain inadequately addressed by technology (ratios of 3.6:1 for

bleeding, 6.2:1 for perforation), challenging the assumption that market forces efficiently translate documented clinical needs into innovation opportunities. Third, the predominance of recent, low-impact patents (46.7% are <5 years old; 74.4% have <5 citations) indicates incremental rather than disruptive advancement.

The ESD innovation ecosystem suffers from three interconnected pathologies: misaligned priorities, insufficient translation, and high noise-to-signal ratio, which warrant strategic intervention rather than reliance on market forces alone.

DISCUSSION

This study provides the first comprehensive integration of publication and patent data to characterize the clinical and technical innovation landscape in ESD. The near-complete absence of strong alignment between publication and patent topics (zero pairs with >0.5 similarity among 54 combinations) provides empirical confirmation of the "knowledge filter": institutional and informational barriers separating clinical researchers from industrial engineers.^{37,38}

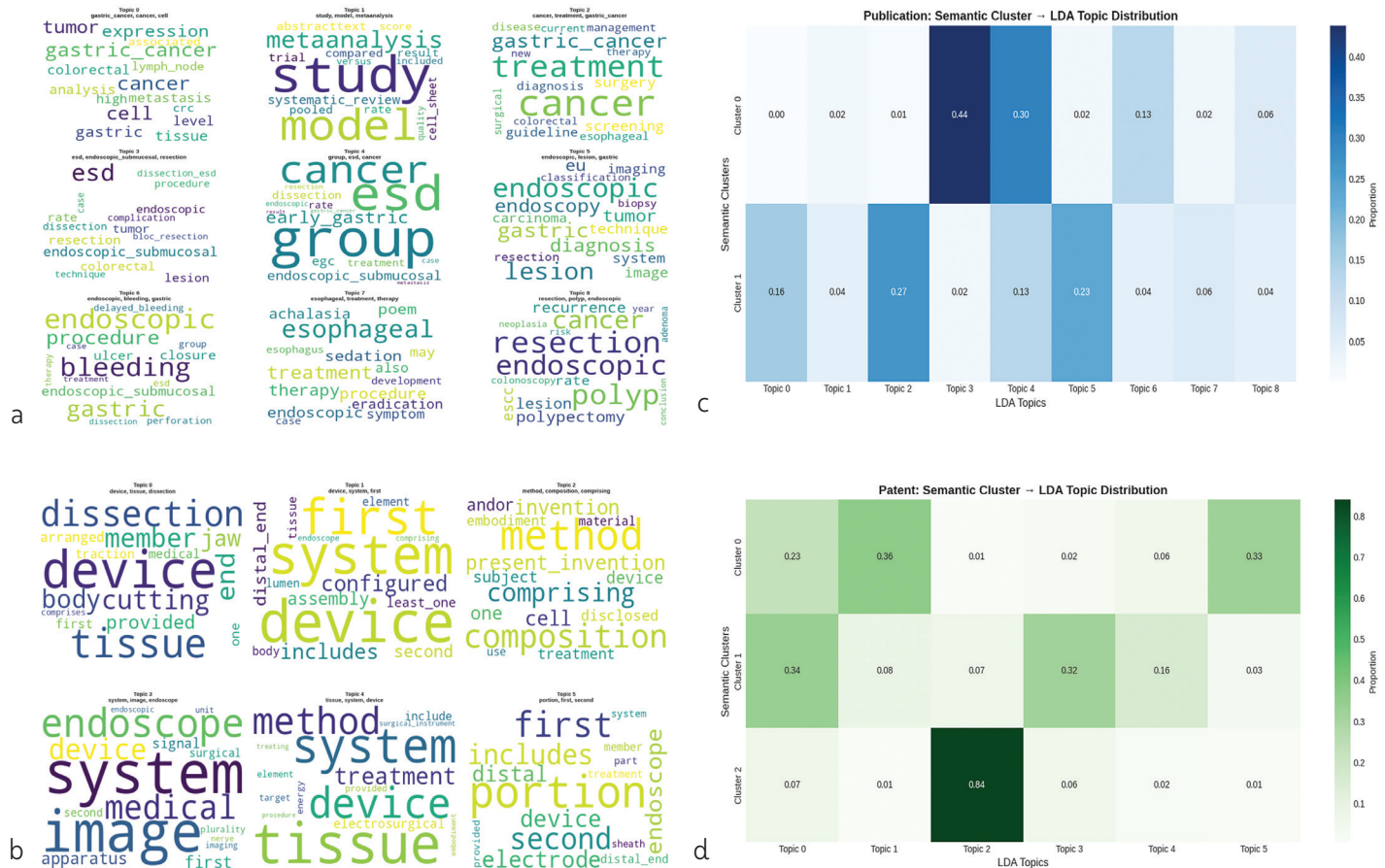


Figure 2. Topic word clouds for publication (a) and patent topics (b), Cluster-topic distribution heatmaps for publications (c) and patents (d)

LDA: Latent Dirichlet allocation

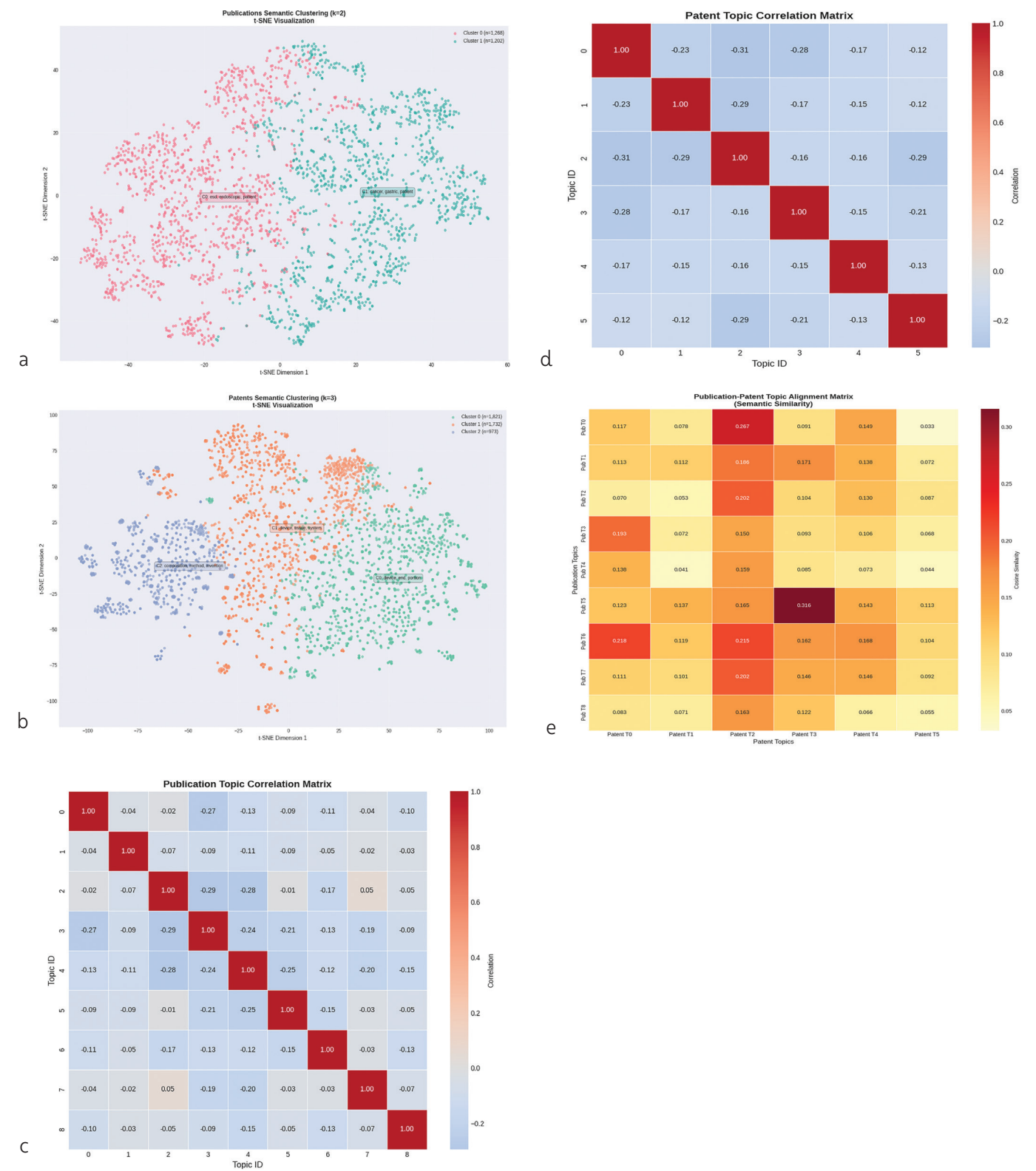


Figure 3. Within-corpus topic correlation matrices for publications (a) and patents (b), t-SNE visualization of semantic clustering for publications (c) and patents (d), and Publication-patent topic alignment matrix (e)

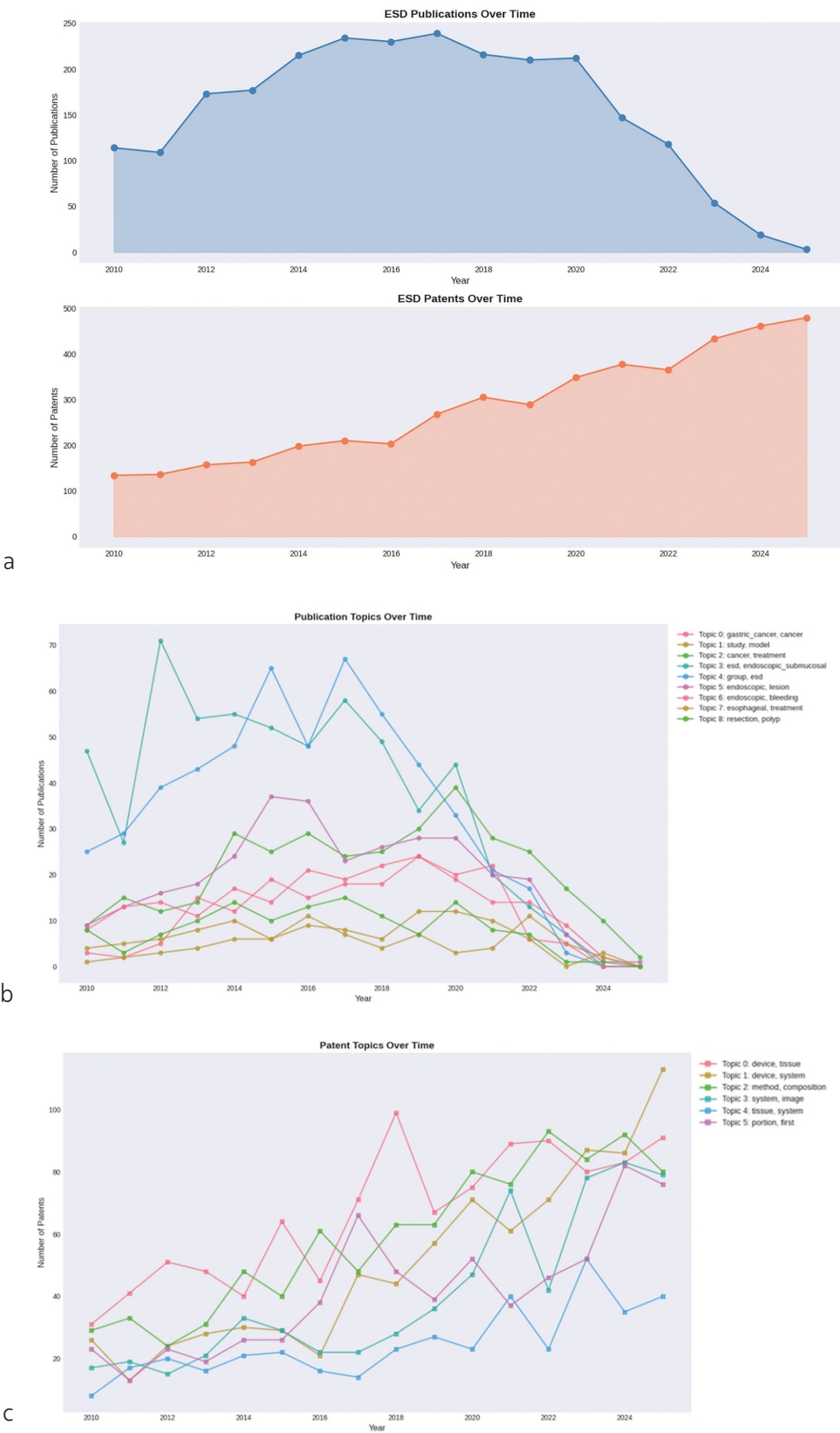


Figure 4. Temporal trends (a), individual topic evolution over time for publications (b) and patents (c)
ESD: Endoscopic submucosal dissection

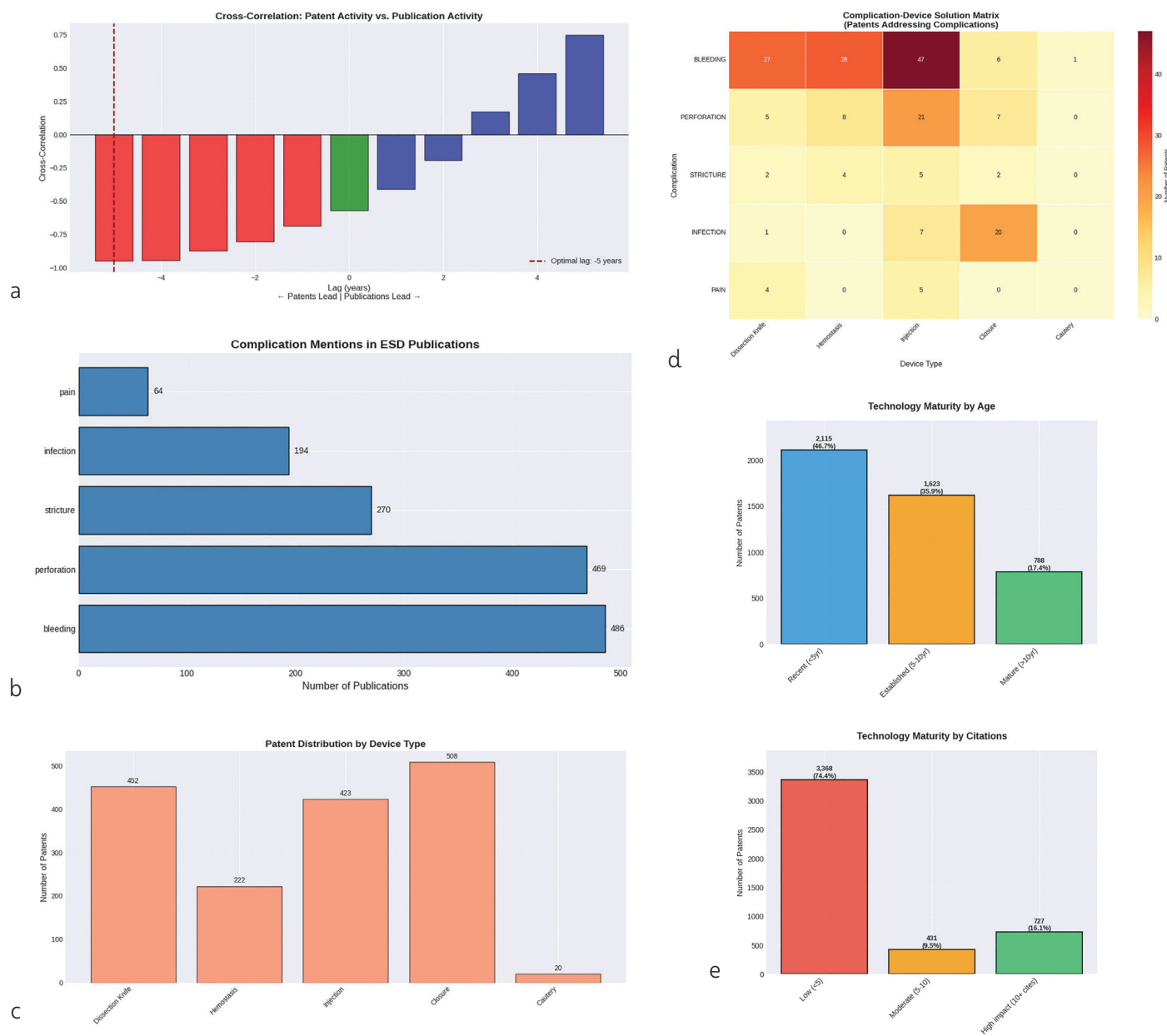


Figure 5. Cross-correlation (a) and complication mentions in ESD Publications (b), patent distribution across five device categories (c), complication-device solution matrix (d), and technology maturity metrics (e)
ESD: Endoscopic submucosal dissection

Several mechanisms explain this divergence: a temporal mismatch between publication cycles and patent-filing strategies; divergent incentive structures (clinical novelty vs. competitive advantage in device specifications); proprietary commercial development that renders substantial innovation invisible to the published literature; and a geographical concentration of ESD expertise in Japan during 2010-2018, resulting in parallel trajectories.^{13,39,40}

Innovation Gaps and Knowledge Filter Mechanisms

The complication-solution matrix reveals heterogeneous technology availability, consistent with uneven absorptive capacity. For bleeding (486 publication mentions, 136 patents, 3.6:1 ratio), the technological emphasis on prevention rather than reactive management may reflect industry risk assessment, although both needs exist. Perforation exhibits more severe gaps (6.2:1): the

predominance of injection-based prevention over closure devices permits two interpretations: either effective knowledge transfer that prioritizes prevention or a failure to develop post-perforation repair technologies. The seven complication-device combinations exhibiting zero patent coverage represent a complete blockage of the knowledge filter. Pain management (9:1:1 ratio) represents the clearest gap: procedural pain remains outside the absorptive capacities of academic and commercial ecosystems.

Temporal Dynamics and Ecosystem Maturation

The inverse temporal relationship ($r=-0.949$ at -5 year lag), which should be interpreted cautiously given the small number of annual observations underlying it, contradicts conventional expectations. When combined with non-significant Granger causality (both directions $p>0.05$), these results suggest a shared response to external factors rather than a direct causal influence. The decline in publications following the 2015 peak may reflect the maturation of clinical evidence: as ESD's efficacy became well-established, diminishing returns reduced incentives to publish while commercial pressures intensified, consistent with innovation lifecycle theory. Changing funding priorities, global diffusion dynamics, and regulatory changes represent additional potential drivers. The convergence of all nine publication topics toward near-zero activity by 2024-2025 suggests systemic factors rather than topic-specific dynamics.

Technology Maturity and Innovation Quality

The paradox of quantitatively active yet qualitatively incremental innovation (46.7% recent patents and 74.4% low-citation) is consistent with findings that healthcare innovation evolves through asynchronous, institutionally fragmented pathways. Technical constraints of endoscopic platforms, regulatory incentives favoring 510(k) clearance for substantially equivalent devices, and reimbursement structures that do not differentiate based on device sophistication collectively create a bounded innovation space. The concentration of high-impact patents in materials science (48%) rather than in mechanical innovations (21%) suggests that novel chemistry is a more influential approach, aligning with knowledge diffusion patterns where breakthrough innovations emerge from cross-disciplinary transfer. The 510(k) pathway, reimbursement structure, and cross-disciplinary transfer mechanisms proposed above are offered as plausible explanatory hypotheses consistent with the observed patterns. This study measures textual similarity and temporal co-variation between publication and patent corpora and does not directly measure regulatory pathway choice, reimbursement policy, technology adoption, or clinical benefit; therefore, these mechanisms have not been directly tested here and should be confirmed with regulatory and health-economics data in future work.

Implications for Practice and Policy

In clinical practice, the technology inventory enables evidence-based device selection and identifies capability gaps; when

complications lacking robust solutions—particularly delayed bleeding and stricture formation—clinicians should emphasize refinement of operator technique over device-centric approaches. For training programs, the 46.7% proportion of recent patents suggests curricula that emphasize fundamental principles over device-specific techniques, while the 17.4% of mature patents identify stable foundations for core skills. Regarding research funding, targeted translational investment could bridge gaps in stricture prevention and pain management through institutional mechanisms that facilitate academic-industrial knowledge transfer. From an innovation policy perspective, the predominance of incremental innovation suggests that current regulatory and reimbursement structures do not sufficiently incentivize high-impact innovation, warranting coordinated action across 510(k) pathways and bundled reimbursement structures.

Study Limitations

Limitations include citation-based sampling, which introduces recency bias for 2023-2025 publications; reliance on abstracts rather than full-text, although prior research demonstrates 80-90% semantic information capture; an English-language patent-family focus that may underrepresent Japanese domestic filings, mitigated by Lens.org's PCT coverage; and a 16-year temporal scope potentially insufficient for detecting longer innovation cycles. Additional limitations identified during peer review include: the Granger causality and vector autoregression analyses rest on only 16 annual observations (2010-2025), which is a very small sample for reliable VAR/Granger estimation; therefore, temporal-causal conclusions should be treated as exploratory; the cross-correlation analysis examined 21 lags on a 16-point series without a formal multiple-comparison correction; therefore, the lag =5-year result requires cautious interpretation and, ideally, a corrected significance threshold; the "divergence score" is a study-specific metric that has not been externally validated or benchmarked against other medical device domains; complication and device classification relied on keyword matching whose precision and recall against manually coded labels has not yet been independently verified; this study measures textual similarity and temporal co-variation only, and does not directly measure technology adoption, clinical benefit, regulatory pathway choice [e.g., 510(k) status], reimbursement outcomes, or knowledge transfer; therefore, causal or mechanistic interpretations of the observed associations remain hypotheses for future testing; and the Word2Vec + K-means clustering solutions showed weak cluster separation (silhouette =0.188 for publications, 0.157 for patents) and should be regarded as a preliminary, lower-confidence complement to the LDA-based topic analysis rather than independent confirmation of it. Future directions include full-text analysis, author-inventor linkage, longitudinal device family tracking, comparative specialty analysis, patient outcome linkage, and geographic comparative analysis. Future work should include a formal validation study of the keyword-based

classification pipeline and, where feasible, present the annual publication and patent counts, topic labels and key terms, the full 9×6 similarity matrix, complication-device cell counts, and patent maturity figures as supplementary tables rather than solely as figure panels, to improve readability and reproducibility.

CONCLUSION

This study demonstrates that clinical research and commercial innovation in ESD operate as parallel, largely independent streams. The 0.792 divergence score, zero strongly-aligned topic pairs, seven innovation gaps, and inverse temporal dynamics collectively reveal an innovation ecosystem requiring deliberate integration mechanisms rather than relying on market-mediated knowledge diffusion. The predominance of incremental innovation (74.4% low-impact patents), despite active filing rates (46.7% recent), suggests that current incentive structures favor patent proliferation over transformative advancement.

For stakeholders, this analysis provides actionable intelligence: clinicians receive evidence-based device inventories; training programs gain guidance on application priorities; research funders identify gaps for targeted investment; device manufacturers understand competitive landscapes and unmet needs; and regulators recognize where approval pathways may insufficiently incentivize needed innovation. The integrated publication-patent methodology offers a generalizable framework for innovation intelligence across procedure-based medical specialties. The path forward requires coordinated action: research funding mechanisms that incentivize technology evaluation; regulatory pathways that balance evidence requirements with innovation incentives; reimbursement structures that reward superior outcomes; and institutional models that facilitate author-inventor overlap.

Ethics

Ethics Committee Approval: This study was based exclusively on publicly available publication data retrieved from established bibliometric databases. No human participants were involved, and the study did not entail primary data collection or any form of intervention. As the data analyzed were already publicly accessible and did not contain identifiable personal information or sensitive data, institutional review board or ethics committee approval was not required for this bibliometric analysis.

Informed Consent: Informed consent was not required for this study, as it involved the bibliometric analysis of publicly available bibliographic data and did not include human participants or the collection or processing of personally identifiable information.

Footnotes

Authorship Contributions

Concept: H.D., S.B., T.A., A.Y., S.V., Design: H.D., S.B., T.A., A.Y., S.V., Data Collection or Processing: H.D., S.B., T.A., A.Y., S.V., Analysis or

Interpretation: H.D., S.B., T.A., A.Y., S.V., Literature Search: H.D., S.B., T.A., A.Y., S.V., Writing: H.D., S.B., T.A., A.Y., S.V.

Conflict of Interest: No conflict of interest was declared by the authors.

Financial Disclosure: The authors declared that this study received no financial support.

REFERENCES

1. Fukuzawa M, Gotoda T. History of endoscopic submucosal dissection and role for colorectal ESD. *Gastrointest Interv.* 2012;1:30-5.
2. Yamamoto H. Endoscopic submucosal dissection—current success and future directions. *Nat Rev Gastroenterol Hepatol.* 2012;9:519-29.
3. Yamamoto H. Endoscopic submucosal dissection of early cancers and large flat adenomas. *Clin Gastroenterol Hepatol.* 2005;3:S74-6.
4. Elkholy S, Awad A, Haggag H, Essam K, Yousef K, El Sherbiny M. Different endoscopic submucosal dissection techniques: a tailored approach. *VideoGIE.* 2024;10:277-84.
5. Bourke MJ, Neuhaus H, Bergman JJ. Endoscopic submucosal dissection: indications and application in western endoscopy practice. *Gastroenterology.* 2018;154:1887-900.e5.
6. Harlow C, Sivananthan A, Ayaru L, Patel K, Darzi A, Patel N. Endoscopic submucosal dissection: an update on tools and accessories. *Ther Adv Gastrointest Endosc.* 2020;13:2631774520957220.
7. Lambin T, Rivory J, Wallenhorst T, et al. Endoscopic submucosal dissection: how to be more efficient? *Endosc Int Open.* 2021;9:E1720-30.
8. Draganov PV, Wang AY, Othman MO, Fukami N. AGA institute clinical practice update: endoscopic submucosal dissection in the United States. *Clin Gastroenterol Hepatol.* 2019;17:16-25.e1.
9. Matsui N, Akahoshi K, Nakamura K, Ihara E, Kita H. Endoscopic submucosal dissection for removal of superficial gastrointestinal neoplasms: a technical review. *World J Gastrointest Endosc.* 2012;4:123-36.
10. Mendes RR, Barreiro P, Mascarenhas A, Franco AR, Carvalho L, Chagas C. Endoscopic submucosal dissection for resections larger than 10 cm: outcomes from a portuguese center. *GE Port J Gastroenterol.* 2023;31:33-40.
11. Kim DH, Jin BC, Oh HH, et al. Outcomes of colorectal endoscopic submucosal dissection according to the size of colorectal neoplasm: a HASID multicenter study. *In Vivo.* 2024;38:2009-15.
12. Autant-Bernard C, Fadaïro M, Massard N. Knowledge diffusion and innovation policies within the European regions. *Res Policy.* 2013;42:196-210.
13. Bahar D, Hausmann R, Hidalgo CA. Neighbors and the evolution of the comparative advantage of nations. *J Int Econ.* 2014;92:111-23.
14. Barbour JA, O'Toole P, Suzuki N, Dolwani S. Learning endoscopic submucosal dissection in the UK: barriers, solutions and pathways for training. *Frontline Gastroenterol.* 2020;12:671-6.
15. Odagiri H, Yasunaga H. Complications following endoscopic submucosal dissection for gastric, esophageal, and colorectal cancer: a review of studies based on nationwide large-scale databases. *Ann Transl Med.* 2017;5:189.
16. Kuroha M, Shiga H, Kanazawa Y, et al. Factors associated with fibrosis during colorectal endoscopic submucosal dissection: does pretreatment biopsy potentially elicit submucosal fibrosis and affect endoscopic submucosal dissection outcomes? *Digestion.* 2021;102:590-8.
17. Toyonaga T, Man-i M, East JE, et al. 1,635 Endoscopic submucosal dissection cases in the esophagus, stomach, and colorectum:

- complication rates and long-term outcomes. *Surg Endosc.* 2013;27:1000-8.
18. Kaplan AV, Baim DS, Smith JJ, et al. Medical device development: from prototype to regulatory approval. *Circulation.* 2004;109:3068-72.
 19. Chatterji AK, Fabrizio KR, Mitchell W, Schulman KA. Physician-industry cooperation in the medical device industry. *Health Aff (Millwood).* 2008;27:1532-43.
 20. Lettl C, Rost K, von Wartburg I. Why are some independent inventors 'heroes' and others 'hobbyists'? The moderating role of technological diversity and specialization. *Res Policy.* 2009;38:243-54.
 21. Cohen D, Billingsley M. Europeans are left to their own devices. *BMJ.* 2011;342:d2748.
 22. Murray F, Stern S. Do formal intellectual property rights hinder the free flow of scientific knowledge? *J Econ Behav Organ.* 2007;63:648-87.
 23. Vajjala S, Majumder B, Gupta A, Surana H. Practical natural language processing. Sebastopol: O'Reilly Media; 2020.
 24. The Lens. The Lens: Free & open patent and scholarly literature search [Internet]. Available from: <https://about.lens.org/>
 25. Zupic I, Čater T. Bibliometric methods in management and organization. *Organizational Research Methods.* 2015;18:429-72.
 26. Syed S, Spruit M. Full text or abstract? Examining topic coherence scores using latent Dirichlet allocation. In: 2017 IEEE International Conference on Data Science and Advanced Analytics (DSAA); 2017; Tokyo, Japan. IEEE; 2017. p. 165-74.
 27. Bird S, Klein E, Loper E. Natural language processing with python. Sebastopol: O'Reilly Media. 2009.
 28. Miller GA. WordNet: a lexical database for English. *Commun ACM.* 1995;38:39-41.
 29. Maier D, Waldherr A, Miltner P, et al. Applying LDA topic modeling in communication research. *Commun Methods Meas.* 2018;12:93-118.
 30. Mikolov T, Chen K, Corrado G, Dean J. Efficient estimation of word representations in vector space. *arXiv.* 2013.
 31. Rousseeuw PJ. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *J Comput Appl Math.* 1987;20:53-65.
 32. Davies DL, Bouldin DW. A cluster separation measure. *IEEE Trans Pattern Anal Mach Intell.* 1979;1:224-7.
 33. Van der Maaten L, Hinton G. Visualizing data using t-SNE. *J Mach Learn Res.* 2008;9:2579-605.
 34. Blei DM, Ng AY, Jordan MI. Latent Dirichlet allocation. *J Mach Learn Res.* 2003;3:993-1022.
 35. Röder M, Both A, Hinneburg A. Exploring the space of topic coherence measures. *WSDM.* 2015.
 36. Granger CWJ. Investigating causal relations by econometric models and cross-spectral methods. *Econometrica.* 1969;37:424-38.
 37. Braunerhjelm P, Acs S, Audretsch D, Carlsson B. The missing link: knowledge diffusion and entrepreneurship in endogenous growth. *Small Bus Econ.* 2010;34:105-25.
 38. Audretsch DB, Bönte W, Keilbach M. Entrepreneurship capital and its impact on knowledge diffusion. *J Bus Venturing.* 2008;23:687-98.
 39. Bacchiocchi E, Montobbio F. Knowledge diffusion from university and public research. *J Technol Transfer.* 2009;34:169-81.
 40. Baruffaldi SH, Simeth M. Patents and knowledge diffusion: the effect of early disclosure. *Res Policy.* 2020;49:103927.