Semantic Echo Pathways (SEP): Tracing How Medical Language Propagates and Transforms

Charu Karakkaparambil James, Marcio Monteiro, Sophie Fellenz
RPTU University Kaiserslautern-Landau
Kaiserslautern, Germany
charu@cs.uni-kl.de, marcio.monteiro@cs.rptu.de, fellenz@cs.uni-kl.de

Abstract

We introduce Semantic Echo Pathways (SEP), a new approach for modeling the cross-domain evolution of medical language. Using continual neural topic models (CoNTM) trained separately on scientific literature, clinical notes, and public health-related data, we track linguistic drift and identify points where concepts change meaning. We propose three novel metrics: Cross-Domain Drift Score, Temporal Echo Lag, and Semantic Mutation Patterns to quantify how medical language travels between the scientific, clinical, and public domain. Applications to evolving concepts such as "long COVID", diagnostic category changes reveal previously undocumented patterns of medical-semantic evolution. Our results bridge computational modeling with the human-centered perspectives of medical humanities, offering clear, domain-aware maps of how medical language shifts across time and domains, and combining quantitative analysis with linguistic and clinical insight.

1 Introduction

Medical language evolves continuously as knowledge, practice, and public understanding change over time. As medical concepts circulate across scientific literature (Canese and Weis, 2013; Marta, 2015), clinical practice (Peterson and Liu, 2021; Chute, 2000), and public discourse, they often reappear with shifts in emphasis, framing, or meaning, a phenomenon we refer to as semantic echo. For example, concepts introduced in biomedical research may later be adopted in clinical notes with abbreviated terminology and eventually appear in public discourse through more narrative or experiential language (Wilce, 2009).

Concrete examples of semantic echo are common in medicine. Disease terminology frequently changes as understanding deepens or social norms shift: terms such as manic depression or senile dementia have been replaced by bipolar disorder and Alzheimer's disease (Leucht et al., 2025) to reflect improved clinical precision and reduced stigma. More recently, the term long COVID emerged from patient communities on social media before being adopted in scientific publications and clinical guidelines, illustrating how public discourse can influence formal medical language (Perego et al., 2020; Callard and Perego, 2021). Similar patterns can be observed in the adoption of person-first language (e.g., "person with diabetes") and in shifts in metaphorical framing, such as describing cancer as a "battle" or immunity as a "shield".

Understanding these dynamics is increasingly important. Research on semantic drift in computational linguistics has advanced substantially, with methods that track meaning change over time in general corpora, such as incremental clustering approaches (Periti et al., 2025). However, these methods rarely address domain transitions or medical language specifically. In parallel, medical NLP research has focused largely on adapting large language models to biomedical and clinical tasks, emphasizing performance, bias, and interpretability (Doneva et al., 2024). What remains underexplored is the systematic modeling of how medical language evolves across both time and domains.

We introduce Semantic Echo Pathways (SEP), an interpretable framework for modeling how medical language evolves across domains such as scientific literature, clinical practice, and public-health discourse. Using continual neural topic models (CoNTM) (James et al., 2025), SEP traces cross-domain concept propagation with temporal lag and semantic change, and quantifies these dynamics through three metrics: Cross-Domain Drift Score, Temporal Echo Lag, and Semantic Mutation Patterns. While we use topic-based representations for interpretability and longitudinal analysis, SEP is model independent and can be applied with alternative semantic representations like LLM embeddings. Our results connect computational modeling

Proceedings of the 1st Workshop on Linguistic Analysis for Health (HeaLing 2026), pages 55–66

with linguistic, clinical, and public perspectives to reveal the temporal life cycle of medical language.

Our work makes the following key contributions:

- We present *Semantic Echo Pathways (SEP)*, a new computational framework and the first method capable of tracing how medical language emerges, spreads, and evolves across scientific, clinical, and public-health domains.
- Three new metrics for quantifying medical language flow: Cross-Domain Drift Score (CDDS), Temporal Echo Lag (TEL), Semantic Mutation Patterns (SMP).
- Empirical analysis demonstrating that medical concepts undergo structured, domain-dependent semantic change as they propagate across the scientific, clinical, and public domains.

The paper is organized as follows. We discuss related work in Section 2. Section 3 presents the SEP framework, including model and metric definitions. Section 4 presents datasets, experimental setup, quantitative results, and qualitative insights. Section 5 provides the conclusion.

2 Related Work

In this section, we examine prior work on medical language evolution, cross-domain and dynamic topic models in healthcare.

2.1 Medical Language Evolution Over Time

Medical discourse is not static; it evolves continuously as scientific knowledge advances, clinical practices shift, and societal attitudes toward illness transform (Nazi and Peng, 2024; Bretonnel Cohen and Demner-Fushman, 2014; Wang et al., 2018). Studies in medical linguistics and NLP have examined terminology change, concept redefinition, and framing effects in biomedical and clinical text, often focusing on specific phenomena such as disease renaming, diagnostic refinement, or stigma reduction (Leucht et al., 2025). More recently, computational approaches have explored semantic drift and meaning change in general-purpose corpora (Periti et al., 2025), but these methods are rarely applied to medical language or evaluated across distinct healthcare domains.

While these studies establish that medical language is dynamic, most analyses are limited to a single domain (e.g., scientific literature or clinical notes) and do not model how concepts propagate or transform as they move between scientific, clinical, and public contexts. As a result, cross-domain semantic relationships and temporal dependencies remain largely unexplored.

We review existing cross-domain and dynamic topic modeling approaches applied in healthcare in the following section.

2.2 Cross-Domain and Dynamic Topic Models in Healthcare

Topic modeling provides an interpretable framework for uncovering latent thematic structure in large text corpora by representing topics as distributions over words. Topic modeling has been widely applied to healthcare text, including clinical notes, biomedical literature, and patient-generated content. Dynamic topic models have been used to study temporal changes within individual domains, such as clinical records or scientific publications (Ye et al., 2024; Sheng et al., 2023; Martinis et al., 2024; Guillén-Pacho et al., 2024), while static models are commonly applied to social media and online health forums (Sheng et al., 2023; Dolatabadi et al., 2023; Paul and Dredze, 2014; Porturas and Taylor, 2021; Zengul et al., 2023).

Existing extensions of topic models typically address either temporal evolution (James et al., 2025; Rahimi et al., 2024; Gupta et al., 2020; Grootendorst, 2022; Dieng et al., 2019) or cross-domain variation (Osnabrügge et al., 2023; Bao et al., 2013; Akash and Chang, 2025), but not both. Dynamic models are usually confined to a single corpus, whereas cross-domain or cross-corpus models assume static settings without explicit temporal dynamics. As a result, scientific, clinical, and public healthcare texts are generally modeled independently, limiting cross-domain semantic alignment. In particular, prior work does not capture how medical concepts propagate and transform across domains over time.

Our work on Semantic Echo Pathways (SEP) addresses this gap by using a Continual Neural Topic Model (CoNTM) (James et al., 2025) to jointly model cross-domain topic structure with temporal dynamics. To the best of our knowledge, no existing topic modeling framework simultaneously captures both cross-domain alignment and topic evolution in the medical domain, motivating the need for a multi-domain, temporally grounded approach.

3 Methodology

To examine how medical concepts emerge and evolve across scientific, clinical, and public domains, we develop Semantic Echo Pathways (SEP), a computational framework that integrates continual neural topic modeling, cross-domain alignment, and new metrics for linguistic change. SEP reconstructs how medical language moves through the healthcare ecosystem and how its meaning shifts over time. The following sections outline SEP's components, including domain-specific models, cross-domain alignment, and metrics for semantic drift and propagation.

3.1 Semantic Echo Pathways (SEP)

Semantic Echo Pathways is a unified framework for modeling how medical concepts originate, diffuse, and transform across scientific, clinical, and public corpora over time.

It consists of three components: (1) domain-specific continual neural topic models, (2) cross-domain topic alignment, and (3) new quantitative metrics designed to capture cross-domain semantic drift and propagation patterns. Together, these components allow us to reconstruct the pathways through which medical language travels and evolves.

3.2 Domain-Specific Continual Topic Models

To capture linguistic evolution within each domain, SEP trains a separate continual neural topic model (CoNTM) (James et al., 2025) for each domain, $d \in \{s, c, p\}$, corresponding to the scientific, clinical, and public domains. Neural topic models are used to discover the hidden thematic structure in a collection of documents by representing topics as distributions over words. In CoNTM, the documents arrive in continuous time slices, with each slice characterized by slightly varying topics. These topics are interconnected through a global parameter, allowing for minor temporal adjustments to the global topics at each time step t.

For each domain, CoNTM learns a sequence of topic-word distribution that evolves over time. Specifically, each model maintains a global and local topic parameter.

Global topic parameter: $\hat{\phi}_d^{\text{global}} \in \mathbb{R}^{K \times V}$,

Local topic parameters :$\hat{\phi}_{d,t}^{\text{local}} \in \mathbb{R}^{K \times V}$,

where K is number of topics, V is vocabulary size for domain d, t is temporal index (e.g., year). By modeling each domain independently, SEP isolates the internal linguistic evolution specific to scientific, clinical, and public domain before comparing them.

Continual Global Update Rule SEP inherits the continual update rule (Algorithm 2, Step 10) from CoNTM (James et al., 2025):

$$\hat{\phi}_d^{\text{global}} \leftarrow (1 - \rho_t)\, \hat{\phi}_d^{\text{global}} + \rho_t \hat{\phi}_{d,t}^{\text{local}},$$

where the forgetting-controlled learning rate is:

$$\rho_t = \frac{1}{(\tau_0 + t)^{\kappa}}, \quad \kappa \in (0.5, 1], \tau_0 \geq 0.$$

This matches CoNTM's continual learning setting, where κ controls forgetting and smoothness over time. In SEP, each domain has its own temporal dynamics:

$$\rho_{d,t} = \frac{1}{(\tau_{0,d} + t)^{\kappa_d}}$$

This allows different domains to evolve at different velocities (e.g., Twitter evolves faster than PubMed).

Generative Process Using CoNTM For each domain, SEP adopts the Dirichlet VAE (Burkhardt and Kramer, 2019) inference used in CoNTM. For a given document, the topic proportions are modeled as a latent variable:

$$z \sim \text{Dirichlet}\left(\alpha_\theta(w)\right).$$

The word likelihood is computed using a product of experts softmax,

$$p(\mathcal{D}_n = w \mid z, \hat{\phi}^{\text{glo}}, \Delta\hat{\phi}_t^{\text{loc}}) = [\sigma(g(\hat{\phi}^{\text{glo}}, \Delta\hat{\phi}_t^{\text{loc}}) \cdot z)]_w,$$

For SEP, this means each domain-level CoNTM models its own linguistic evolution.

3.3 Cross-Domain Topic Alignment

While CoNTM ensures temporal alignment within each domain, the Semantic Echo Pathways (SEP) framework extends this to cross-domain alignment across scientific, clinical, and public corpora. At each timestamp t, we denote the topics for the three domains as:

$$\hat{\phi}_{s,k}(t), \hat{\phi}_{c,j}(t), \hat{\phi}_{p,m}(t)$$

where k, j, m index topics in their respective domains.

Topic Embedding Construction To compare topics across domains, SEP projects each topic into a unified semantic space using pretrained medical word embeddings (e.g., BioWordVec (Zhang et al., 2019), ClinicalBERT token embeddings (Alsentzer et al., 2019)). Each topic is embedded as

$$\mathbf{e}_{d,k}(t) = \sum_{v=1}^{V_d} \hat{\phi}_{d,k}(t)[v]\mathbf{w}_v,$$

where each word $\mathbf{w}_v$ is the pretrained medical embedding for word v. This produces dense semantic vectors that capture conceptual meaning and make topics comparable across heterogeneous datasets (scientific abstracts, clinical documentation, public social media).

Cross-Domain Similarity Matrix After embedding topics into vectors $\mathbf{e}_{d,k}(t)$, SEP constructs a similarity matrix for topic sets across domain pairs:

$$M_{d_1,d_2}(i,j) = \cos\left(\mathbf{e}_{d_1,i}(t), \mathbf{e}_{d_2,j}(t)\right)$$

This matrix tells us how likely topic i in domain d_1 corresponds to topic j in domain d_2. For example, scientific Topic 3 might align with clinical Topic 7 or public Topic 12 might align with scientific Topic 5. This matrix is the foundation for the matching step below.

Hungarian Optimal Matching Produce a one-to-one, globally optimal alignment between topics. Once we have the similarity matrix M, SEP needs to produce a mapping between topics of the two domains. However, the greedy matching often leads to suboptimal assignments by taking the highest similarity pairs one by one. But we need global optimality, not the local choices. Thus, SEP applies the Hungarian algorithm (Kuhn, 1955) (a classic bipartite graph matching algorithm) to minimize overall semantic distance:

$$\pi_{d_1,d_2}(t) = \arg\min_{\pi} \sum_{i=1}^{K} \left(1 - M_{d_1,d_2}(i, \pi(i))\right),$$

where π is a permutation representing the topic-to-topic assignment. For each topic i in domain d_1, the algorithm finds the best corresponding topic $\pi(i)$ in domain d_2. The objective function minimizes the total semantic dissimilarity, and the result is the unique optimal assignment linking the two topic sets. This mapping enables SEP to trace the semantic pathway of a topic across: Scientific, Clinical, and Public domain over time.

After all three steps, this allows SEP to reconstruct cross-domain semantic trajectories. That is, how a concept starts in one domain, when it appears in another, and how its meaning changes. This is the foundation for SEP's drift scores, echo lags, and mutation pathway detection.

3.4 Novel Metrics for Cross-Domain Semantic Analysis

To quantify cross-domain linguistic evolution, SEP introduces three new metrics that have no existing equivalents in the literature. These metrics capture how medical concepts: (1) drift semantically across domains, (2) propagate with temporal delay, (3) mutate into new or divergent meanings.

3.4.1 Cross-Domain Drift Score (CDDS)

After SEP aligns topics across domains using the Hungarian algorithm, each topic in the source domain d_1 is mapped to exactly one corresponding topic in the target domain d_2. Let $\pi_{d_1,d_2}(i)$ denote the index of the topic in domain d_2 that is matched to topic i in domain d_1.

The Cross-Domain Drift Score (CDDS) measures how much the meaning of an aligned topic pair changes as the concept moves from one domain to another. To compute this, SEP first represents each topic as a semantic embedding vector.

$$\text{CDDS}(d_1 \to d_2, i) = \begin{array}{c} \| e_{d_1,i}(t_k) - \\ e_{d_2,\pi_{d_1,d_2}(i)}\left(t'_k\right) \| \end{array}$$

where $\mathbf{e}_{d_1,i}(t_k)$ is the first domain topic at time t_k, and $\mathbf{e}_{d_2,\pi_{d_1,d_2}(i)}(t'_k)$ corresponds to aligned topic embedding in domain d_2. A higher CDDS indicates, introduction of new metaphors, appearance of different symptoms or concerns, recontextualization through media or public discourse.

This metric is crucial for detecting semantic drift, especially when scientific terminology is adopted and transformed by the public.

3.4.2 Temporal Echo Lag (TEL)

Temporal Echo Lag (TEL) quantifies the delay with which a medical topic transitions from one domain to another. For a given term, we define TEL as the difference between the earliest timestamp at which the term appears in an originating domain and the earliest timestamp at which it appears in a receiving domain.

By computing TEL across domain pairs (e.g., Scientific→Clinical, Scientific→Public, Clinical→Public), SEP provides a structured way to measure how quickly concepts propagate across institutional, professional, and public domains. This metric is especially important in rapidly evolving medical contexts, such as pandemic outbreaks where delays in terminology adoption may correspond to gaps in knowledge transmission, inconsistent messaging, or misalignment between expert and lay understandings of disease. For domain d_1 and domain d_2, and topic i:

$$TEL\,(d_1 \to d_2, i) = t^{\text{first}}_{d_2,i} - t^{\text{first}}_{d_1,i}.$$

where $t^{\text{first}}_{d_2,i}$ is the first year the topic appears in domain d_2, and $t^{\text{first}}_{d_1,i}$ is the first year the topic appears in domain d_1. If TEL > 0, it means the concept spread forward, TEL < 0 means topic originates in receiving domain, then influences origin. TEL = None, means topic does not appear in one domain. TEL provides a direct measure of concept dissemination speed like, how fast do scientific terms appear in public conversation?, and Do clinicians adopt new terminology later than researchers?

TEL also reveals whether domains evolve in parallel or whether one consistently leads. Negative TEL values highlight cases where patients or the public introduce concepts earlier.

Operationalization of Topic Appearance In our setting, topics are represented by their top-N words, which are fixed. We define the first appearance of a topic in a domain as the earliest time step at which at least one of its representative words appears in the corpus with a document frequency exceeding the preprocessing threshold. Concretely, a topic is considered present at a given time step if at least one of its top-N words occurs in at least $\tau\%$ of documents within that time slice, where τ corresponds to the minimum document frequency used during corpus preprocessing ($\tau = 0.1\%$; see Table 5). This operationalization mitigates sensitivity to isolated or noisy mentions while preserving early semantic signals. We empirically observe that TEL estimates are stable across reasonable choices of N and τ.

3.4.3 Semantic Mutation Pattern (SMP)

Understanding how medical concepts transform as they circulate across scientific, clinical, and public discourse requires more than quantifying when they appear (TEL) or how much they change (CDDS). To capture the qualitative structure of linguistic evolution, we introduce Semantic Mutation Patterns (SMP), a taxonomy of cross-domain propagation pathways that characterize the ways concepts emerge, diverge, refract, and feed back across the healthcare system. SMP operationalizes the intuition that medical language evolves analogously to biological entities: concepts mutate, branch, converge, or remain isolated, reflecting the influence of institutional norms, communicative pressures, and societal interpretation.

Formally, for a topic i, we define a concept trajectory:

$$\Gamma(i) = \left\{ \left(d, t^{\text{first}}_{d,i}, \mathbf{e}_{d,t}(i)\right) : d \in \{\text{Sci}, \text{Clin}, \text{Pub}\} \right\}$$

where each tuple captures (i) the domain of occurrence, (ii) the time of first emergence of topic, and (iii) the semantic signature of the topic derived from pre-trained embedding. By analyzing the order of first occurrences together with the directional magnitudes of semantic drift, SEP categorizes trajectories into one of five mutation patterns:

(1) Linear Echo A unidirectional propagation from scientific research to clinical adoption and subsequently to public communication. This pattern reflects hierarchical knowledge transfer and is characterized by strictly increasing timestamps $t_{\text{Sci}} < t_{\text{Clin}} < t_{\text{Pub}}$ and moderate CDDS across transitions.

(2) Reverse Echo A patient or public originating pattern in which a topic first appears in public domain, then enters clinical documentation, and eventually scientific literature (e.g., "long COVID"). Reverse Echoes challenge top-down models of medical communication and often correspond to negative TEL values, indicating bottom-up conceptual innovation.

(3) Loop Echo A bidirectional propagation in which a topic originates in scientific literature, is reframed or recontextualized by public domain, and then reappears within science in an altered form. This pattern reflects feedback loops between societal interpretation and institutional language, typically exhibiting asymmetric CDDS reflecting substantial semantic reframing.

(4) Divergent Echo A branching pattern in which the topic splits into distinct semantic variants across domains. Here, CDDS values differ significantly between domain pairs (e.g., Scientific→Clinical vs. Scientific→Public), revealing conceptual fragmentation and domain-specific reinterpretation.

5) Dead Echo A degenerate pattern in which the topic remains confined to a single domain, suggesting a lack of perceived relevance, insufficient dissemination, or terminological isolation.

Together, these patterns form an interpretable taxonomy of linguistic evolution in healthcare, showing how terminology adapts across domains and enabling SEP to trace the circulation of meaning and potential sources of misunderstanding.

4 Experiments

This section evaluates the Semantic Echo Pathways (SEP) framework and presents details of the datasets used. Experimental details are provided in Appendix C.

4.1 Datasets

To study how medical language evolves across scientific, clinical, and public settings, we use three distinct datasets that reflect different modes of communication.

Scientific sources include PubMed abstracts (1980–2025) (Canese and Weis, 2013), representing over 35 million biomedical research texts. These provide formal institutional language and long-term terminology evolution, serving as a baseline for early conceptual emergence.

Clinical sources consist of clinical notes (Naudet, 2025), de-identified hospital documentation, and physician forums (e.g., SERMO, Doximity). They reflect how practitioners adopt, reinterpret, or operationalize concepts in real-world medical practice.

Public sources include the Pubhealth dataset (Kotonya and Toni, 2020), a publicly available dataset designed for explainable fact-checking in the public health domain. The dataset is released under the MIT License, allowing research use. Additional details in Appendix A.

4.2 Quantitative Results

In this section, we quantify cross-domain semantic change and compare cross-domain drift against natural intra-domain temporal evolution.

4.2.1 Quantifying Semantic Drift from Scientific to Public Medical Language

The distribution of Cross-Domain Drift Scores shows a clear ordering of semantic drift across domain transitions. Topics moving from scientific to clinical discourse exhibit the lowest drift, followed by increased drift from clinical to public discourse, with the largest drift observed when concepts propagate directly from scientific literature to public discussion. This pattern supports the central hypothesis of the Semantic Echo Pathways framework that semantic transformation intensifies as concepts move farther from their point of origin.

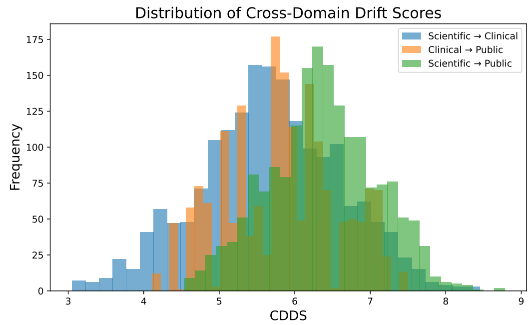

Figure 1: Distribution of Cross-Domain Drift Scores (CDDS) across scientific, clinical, and public domains, showing increasing semantic drift as medical concepts propagate from scientific literature to clinical practice and public discourse.

The scientific to clinical transition displays a relatively compact and left-shifted distribution, indicating that clinical discourse largely preserves scientific meaning while adapting concepts for applied use. In contrast, the clinical-to-public transition shows higher drift and greater variance, reflecting increased reinterpretation driven by public-facing communication. The scientific-to-public pathway exhibits the highest drift and widest spread, suggesting substantial semantic transformation when concepts bypass clinical mediation.

Overall, the observed distributions demonstrate that semantic drift is not uniform across domains but instead follows a structured progression driven by the communicative role of each domain. The results support the central claim of SEP that medical concepts accumulate semantic change as they propagate through the healthcare communication pipeline, with the greatest transformations occurring when concepts enter public discourse without professional mediation.

4.2.2 Semantic Drift vs Propagation Delay

Figure 2 illustrates the relationship between Cross-Domain Drift Scores (CDDS) and echo lag (in years) for three domain transitions: scientific→clinical, clinical→public, and scientific→public. The scatter distribution shows a general positive association between propagation delay and semantic drift, indicating that concepts experiencing longer delays across domains tend to display greater semantic divergence.

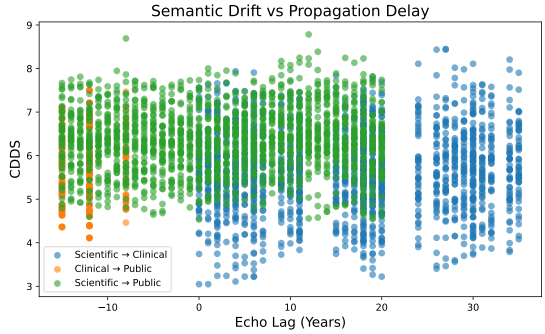

Figure 2: Relationship between Cross-Domain Drift Score (CDDS) and echo lag for aligned topics across scientific (PubMed), clinical (clinical trials), and public (Pubhealth) domains, showing increased semantic drift with longer propagation delays, particularly in the transition to public discourse.

In the scientific→clinical transition, echo lags begin at zero and extend into substantial positive delays, indicating that clinical uptake follows scientific publication with varying temporal gaps. Concepts that take longer to enter public discourse tend to show stronger semantic transformation. This indicates that delayed public diffusion is often accompanied by reinterpretation, simplification, or reframing processes that are less constrained by professional or institutional standards.

The scientific→public transition further reinforces this pattern, displaying consistently high CDDS values across a wide range of echo lags. This suggests that when knowledge travels directly from scientific contexts into public discourse, semantic drift tends to be substantial, especially as temporal delay increases.

Overall, the figure supports the central premise of SEP that semantic drift and temporal delay are connected. Longer echo lags correspond to greater semantic divergence, particularly in transitions involving public discourse. These findings underscore that semantic change is shaped not only by domain differences but also by the temporal dynamics of knowledge propagation.

In contrast, the clinical-to-public transition shows a clearer upward trend in CDDS with increasing echo lag. Points corresponding to longer delays cluster at higher drift values, indicating that concepts that take longer to reach public discourse are more likely to undergo substantial semantic change. This pattern aligns with the idea that delayed public uptake often involves reinterpretation, simplification, or reframing driven by media narratives rather than professional constraints.

Overall, the figure provides empirical support for SEP's central claim that temporal delay and semantic drift are closely linked. Longer echo lags are associated with greater semantic divergence, particularly in the transition to public discourse.

4.2.3 Cross-Domain Drift vs. Natural Temporal Drift

Figure 3 demonstrates a clear separation between cross-domain semantic drift and natural intra-domain temporal drift. Across all domain transitions, the magnitude of semantic change observed when topics move between domains substantially exceeds the drift arising from normal topic evolution over time within a single domain. This indicates that semantic change in medical language is not merely a consequence of gradual temporal dynamics, but is strongly amplified by transitions across communicative contexts.

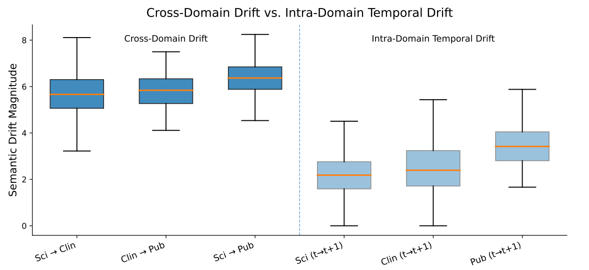

Figure 3: Boxplot comparison of semantic drift magnitudes for cross-domain topic transitions (Scientific→Clinical, Clinical→Public, Scientific→Public) and intra-domain temporal evolution (t→t+1) within each domain. Cross-domain transitions exhibit substantially greater semantic drift than natural within-domain topic evolution, particularly for pathways involving public discourse.

Among the cross-domain pathways, the Scientific→Public transition exhibits the largest semantic disruption, with the highest median drift and variance. This supports the SEP hypothesis that concepts transferred directly from scientific literature into public discourse undergo stronger reinterpretation in the absence of clinical mediation.

The Scientific→Clinical transition shows more moderate drift: while clearly exceeding natural scientific temporal change, it remains lower than both Clinical→Public and Scientific→Public transitions. This pattern reflects the stabilizing role of clinical discourse, which largely preserves scientific meaning while adapting terminology for practice.

Temporal baselines further indicate that public discourse is intrinsically less stable than scientific

Domain	Year	Topic Keywords
Scientific	1990	negative, transmitted, infection, positive, infections, outbreak, transmission, hepatitis, staphylococcus, gram
Clinical	1999	protease, tetanus, ritonavir, viral, mass, infected, reverse, inhibitor, gamma, transcriptase
Public	2010	white, government, private, press, mayor, radiation, vaccination, conversation, revenue, federal

Table 1: Cross-domain semantic trajectory of the *infection* concept (K=50). The term emerges first in scientific, subsequently adopted in clinical, and later formalized in the public domain, illustrating a Linear Echo semantic mutation pattern.

Metric	Value
TEL (Scientific → Clinical)	9 years
TEL (Clinical → Public)	11 years
CDDS (Scientific → Clinical)	0.473
CDDS (Clinical → Public)	0.621
CDDS (Scientific → Public)	0.663
Semantic Mutation Pattern	Linear Echo

Table 2: Quantitative analysis of the semantic echo of the concept *infection* across domains. Results show increasing temporal delay and semantic drift as the concept propagates from scientific to clinical to public discourse, following a linear echo pattern.

or clinical language, with higher intra-domain drift. This instability creates greater potential for semantic variation even prior to cross-domain transfer. To conclude, the Medical concepts change a little as time passes, but they change a lot when they move from science to clinics or to the public.

4.3 Deep Dive: One Concept, All Three Domains

Table 1 shows that the concept infection follows a Linear Echo semantic mutation pattern, with first emergence in scientific discourse (1990), followed by delayed adoption in clinical documentation (1999) and substantially later appearance in public discourse (2010). The large Temporal Echo Lags ($TEL_{Sci \to Clin} = 9$, $TEL_{Clin \to Pub} = 11$) indicate a slow hierarchical diffusion of terminology across domains, consistent with institutionalized knowledge transfer rather than bottom-up conceptual innovation.

Despite this orderly propagation, the Cross-Domain Drift Scores reveal substantial semantic transformation at each transition. The drift from scientific to clinical discourse (CDDS = 0.473) reflects a shift from pathogen-centric framing focused on transmission, outbreak dynamics, and specific infectious agents to a clinically operational perspective emphasizing treatment mechanisms and biomedical intervention, as indicated by keywords such as protease, ritonavir, and reverse transcriptase.

The transition from clinical to public discourse exhibits even greater semantic drift (CDDS = 0.621), culminating in the largest divergence between scientific and public representations (CDDS = 0.663). Notably, the public-domain topic associated with *infection* is dominated by institutional and policy-related terms such as government, mayor, federal, and vaccination. This suggests that in public discourse, the concept of infection becomes embedded within broader political and administrative narratives, moving further away from its original biomedical framing.

Taken together, this case study illustrates that Linear Echo propagation does not imply semantic stability. Even when concepts move slowly and hierarchically across domains, they may undergo pronounced semantic erosion or recontextualization, particularly upon entering public discourse. This result underscores the importance of modeling both temporal delay and semantic drift: long echo lags alone do not preserve meaning, and public adoption may coincide with significant conceptual transformation.

Beyond individual examples, we observe that topics in each domain maintain comparable topic coherence, diversity, and overall topic quality over time (Karakkaparambil James et al., 2024). In particular, public-domain topics exhibit good coherence scores, suggesting that broader or more heterogeneous vocabularies reflect semantic broadening rather than topic collapse. This supports the interpretation that CDDS captures genuine semantic drift between domains rather than noise introduced by unstable or poorly formed topics.

Datasets	Coherence (TC)	Diversity (TD)	Quality (TQ)
PubMed	0.647	0.969	0.627
Clinical	0.724	0.981	0.711
Pubhealth	0.578	0.975	0.563

Table 3: Average topic coherence (TC), topic diversity (TD), and topic quality (TQ) for topics learned independently in each domain.

Taken together, the qualitative inspection of aligned topic pairs and the observed stability of topic quality (see Table 3) and temporal topic smoothness (TTS) (see Figure 4) indicate that cross-domain topic matching is semantically meaningful, and that CDDS reflects real divergence in how concepts evolve across expert and public domains.

4.3.1 Sensitivity to Number of Topics

The number of topics K affects how easy it is to interpret topic changes across domains. We com-

Domain	Year	Topic Keywords
Scientific	1990	indirect, sera, negative, positive, tested, virus, titers, strains, isolates, infection
Clinical	1999	marrow, cells, infected, lymphocytes, blood, stem, white, donor, viral, needle
Public	2010	story, study, patients, autism, test, novel, preliminary, subjects, compare, results

Table 4: Cross-domain semantic trajectory of the *infection* concept (K=20). The term emerges first in scientific, and is subsequently adopted in clinical, and later formalized in the public domain, illustrating a Linear Echo semantic mutation pattern.

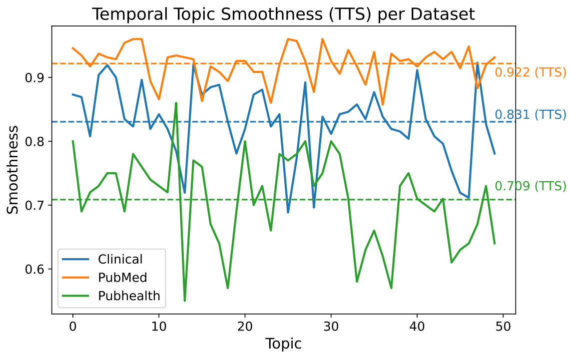

Figure 4: The figure shows the temporal topic smoothness values for the Clinical, PubMed, and Pubhealth datasets across topics. Solid lines represent the smoothness scores per topic, while dashed horizontal lines indicate the average smoothness for each dataset.

pare models with $K = 50$ (used in our main experiments) and $K = 20$. When K is smaller, detailed medical topics (e.g., *infection*) are merged into broader topics, which changes the topic labels.

For *infection*, both models recover a Linear Echo pattern (1990→1999→2010) with stable TEL (9 and 11 years), indicating robust temporal structure. However, semantic drift differs: with K=50, CDDS increases progressively (0.4728→0.6209→0.6632), reflecting clearer downstream transformation, especially in public discourse. In contrast, K=20 yields broader, less distinct topics and weakening Scientific→ Public drift (0.5249), partially masking domain-specific divergence.

Overall, fewer topics produce broader themes but weaker cross-domain alignment, while higher topic granularity enables more precise tracking of semantic evolution. We therefore use K=50 in our main experiments.

5 Conclusions

In this work, we introduced SEP, a cross-domain semantic evolution framework that integrates continual neural topic modeling with novel linguistic analytics to examine how medical concepts propagate and transform across scientific, clinical, and public domain. Building on the temporal continuity enforced by CoNTM, SEP extends topic evolution analysis beyond a single domain by introducing Temporal Echo Lag (TEL) to quantify delays in cross-domain concept adoption, Cross-Domain Drift Score (CDDS) to measure semantic reframing across vocabularies, and Semantic Mutation Patterns (SMP) to characterize the qualitative pathways through which concepts evolve. Together, these components offer a unified and interpretable methodology for mapping the lifecycle of medical terminology as it travels across expert and lay communities.

Our analysis highlights that the evolution of healthcare language is neither linear nor uniform: concepts may diffuse hierarchically from scientific research to clinical practice to public communication, but they may also originate within patient communities, diverge into multiple domain-specific interpretations, or undergo secondary reframing as they re-enter scientific discourse. These findings underscore the importance of treating medical language as a dynamic, multi-voiced phenomenon shaped by institutional processes, professional norms, and social perception.

Beyond providing a computational toolkit, SEP contributes a theoretical lens for understanding linguistic coordination across the healthcare ecosystem. By identifying where semantic drift or delayed adoption occurs, SEP can inform more effective health communication, identify sources of misunderstanding or misinformation, and highlight patient-led innovations in conceptualization. As the volume and velocity of medical information continue to accelerate, especially during public health crises, frameworks such as SEP offer valuable guidance for aligning terminology across stakeholders, supporting equitable knowledge dissemination, and fostering more coherent, transparent, and responsive communication in medicine.

Acknowledgements

The authors acknowledge support by the BMFTR award 01IS24071A, DFG through FOR 5359 (ID 459419731), TRR 375 (ID 511263698), SPP 2298 (Id 441826958), and SPP 2331 (441958259, 553345933, 466468799), and the Carl-Zeiss Foundation through the initiatives AI-Care and Process Engineering 4.0.

Limitations

Our public-domain corpus is derived from the Pubhealth dataset, which consists of public health–related claims and explanations curated for fact-checking tasks. While this dataset provides structured and thematically focused public discourse, it does not fully capture the breadth of informal, conversational, or community-driven medical communication. Compared to open social media platforms, Pubhealth reflects a more curated and claim-centered form of public health language. Consequently, estimates of Temporal Echo Lag (TEL) and Cross-Domain Drift Score (CDDS) in the public domain should be interpreted as representative of structured public health communication rather than the entire spectrum of participatory discourse.

In addition, topic-based representations are sensitive to the choice of the number of topics K, which determines the granularity of learned concepts. Smaller values of K aggregate semantically related themes into broader topics, potentially inflating measured semantic drift and reducing cross-domain interpretability. We fix $K = 50$ to balance interpretability and alignment stability, though different choices of topic granularity may affect the observed semantic trajectories.

Ethical statement

This study uses three publicly available datasets: PubMed abstracts (scientific domain), clinical trials dataset (clinical domain), and the Pubhealth dataset (public domain)

PubMed abstracts are publicly accessible bibliographic records and contain no patient-level identifiers. The clinical data were previously anonymized and made available for research use.

The public-domain corpus is the Pubhealth dataset, a publicly released benchmark dataset for explainable automated fact-checking in public health. The dataset is distributed under the MIT License and permits research use.

Across all domains, preprocessing removed metadata fields and retained only textual content necessary for aggregate topic modeling. The study does not involve interaction with human participants, intervention, or collection of new personal data. All analyses were conducted on pre-existing datasets in compliance with their respective licenses and applicable research ethics guidelines.

We acknowledge that even publicly available health-related text may concern sensitive topics. Accordingly, we report only aggregated topic-level findings and avoid quoting or reproducing potentially sensitive content.

References

Pritom Saha Akash and Kevin Chen-Chuan Chang. 2025. Understanding cross-domain adaptation in low-resource topic modeling. *arXiv preprint arXiv:2506.07453*.

Emily Alsentzer, John Murphy, William Boag, Wei-Hung Weng, Di Jindi, Tristan Naumann, and Matthew McDermott. 2019. Publicly available clinical bert embeddings. In *Proceedings of the 2nd clinical natural language processing workshop*, pages 72–78.

Yang Bao, Nigel Collier, and Anindya Datta. 2013. A partially supervised cross-collection topic model for cross-domain text classification. In *Proceedings of the 22nd ACM international conference on Information & Knowledge Management*, pages 239–248.

Kevin Bretonnel Cohen and Dina Demner-Fushman. 2014. Biomedical natural language processing.

Sophie Burkhardt and Stefan Kramer. 2019. Decoupling sparsity and smoothness in the dirichlet variational autoencoder topic model. *Journal of Machine Learning Research*, 20(131):1–27.

Felicity Callard and Elisa Perego. 2021. How and why patients made long covid. *Social science & medicine*, 268:113426.

Kathi Canese and Sarah Weis. 2013. Pubmed: the bibliographic database. *The NCBI handbook*, 2(1):2013.

Christopher G Chute. 2000. Clinical classification and terminology: some history and current observations. *Journal of the American Medical Informatics Association*, 7(3):298–303.

Adji B Dieng, Francisco JR Ruiz, and David M Blei. 2019. The dynamic embedded topic model. *arXiv preprint arXiv:1907.05545*.

Elham Dolatabadi, Diana Moyano, Michael Bales, Sofija Spasojevic, Rohan Bhambhoria, Junaid Bhatti, Shyamolima Debnath, Nicholas Hoell, Xin Li, Celine Leng, and 1 others. 2023. Using social media to help understand patient-reported health outcomes of post–covid-19 condition: natural language processing approach. *Journal of Medical Internet Research*, 25:e45767.

Simona Emilova Doneva, Sijing Qin, Beate Sick, Tilia Ellendorff, Jean-Philippe Goldman, Gerold Schneider, and Benjamin Victor Ineichen. 2024. Large language models to process, analyze, and synthesize biomedical texts: a scoping review. *Discover Artificial Intelligence*, 4(1):107.

Maarten Grootendorst. 2022. Bertopic: Neural topic modeling with a class-based tf-idf procedure. *arXiv preprint arXiv:2203.05794*.

Ibai Guillén-Pacho, Carlos Badenes-Olmedo, and Oscar Corcho. 2024. Dynamic topic modelling for exploring the scientific literature on coronavirus: an unsupervised labelling technique. *International Journal of Data Science and Analytics*, pages 1–31.

Pankaj Gupta, Yatin Chaudhary, Thomas Runkler, and Hinrich Schütze. 2020. Neural topic modeling with continual lifelong learning. In *Proceedings of the 37th International Conference on Machine Learning*, ICML'20. JMLR.org.

Matthew Honnibal and Ines Montani. 2017. spaCy 2: Natural language understanding with Bloom embeddings, convolutional neural networks and incremental parsing. To appear.

Charu Karakkaparambil James, Waleed Mustafa, Marius Kloft, and Sophie Fellenz. 2025. Continual neural topic model. *arXiv preprint arXiv:2508.15612*.

Charu Karakkaparambil James, Mayank Nagda, Nooshin Haji Ghassemi, Marius Kloft, and Sophie Fellenz. 2024. Evaluating dynamic topic models. In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 160–176, Bangkok, Thailand. Association for Computational Linguistics.

Neema Kotonya and Francesca Toni. 2020. Explainable automated fact-checking for public health claims. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 7740–7754, Online. Association for Computational Linguistics.

Oliver Kramer and Oliver Kramer. 2016. Scikit-learn. *Machine learning for evolution strategies*, pages 45–53.

Harold W Kuhn. 1955. The hungarian method for the assignment problem. *Naval research logistics quarterly*, 2(1-2):83–97.

Stefan Leucht, Deniz S Yurtseven, Josef Priller, John M Davis, Jim van Os, and Iris EC Sommer. 2025. Towards a name change of schizophrenia: Positive and negative symptoms disorder (pnd). *European Archives of Psychiatry and Clinical Neuroscience*, pages 1–9.

Monica Mihaela Marta. 2015. A brief history of the evolution of the medical research article. *Clujul Medical*, 88(4):567.

Maria Chiara Martinis, Chiara Zucco, Antonio Amodeo, Vincenzo Facente, Francesco Greco, and Mario Cannataro. 2024. Evolution of medical reports over time: an analysis using dynamic topic modeling. In *2024 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, pages 6921–6928. IEEE.

Louis Brulé Naudet. 2025. Clinical trials dataset: Comprehensive clinicaltrials.gov data for semantic analysis.

Zabir Al Nazi and Wei Peng. 2024. Large language models in healthcare and medical domain: A review. In *Informatics*, volume 11, page 57. MDPI.

Moritz Osnabrügge, Elliott Ash, and Massimo Morelli. 2023. Cross-domain topic classification for political texts. *Political Analysis*, 31(1):59–80.

Michael J Paul and Mark Dredze. 2014. Discovering health topics in social media using topic models. *PloS one*, 9(8):e103408.

Elisa Perego, Felicity Callard, Laurie Stras, Barbara Melville-Jóhannesson, Rachel Pope, and Nisreen A Alwan. 2020. Why the patient-made term'long covid'is needed. *Wellcome Open Research*, 5:224.

Francesco Periti, Sergio Picascia, Stefano Montanelli, Alfio Ferrara, and Nina Tahmasebi. 2025. Studying word meaning evolution through incremental semantic shift detection. *Language Resources and Evaluation*, 59(2):1363–1399.

Kevin J Peterson and Hongfang Liu. 2021. An examination of the statistical laws of semantic change in clinical notes. *AMIA Summits on Translational Science Proceedings*, 2021:515.

Thomas Porturas and R Andrew Taylor. 2021. Forty years of emergency medicine research: Uncovering research themes and trends through topic modeling. *The American Journal of Emergency Medicine*, 45:213–220.

Hamed Rahimi, Hubert Naacke, Camelia Constantin, and Bernd Amann. 2024. Antm: aligned neural topic models for exploring evolving topics. In *Transactions on Large-Scale Data-and Knowledge-Centered Systems LVI: Special Issue on Data Management-Principles, Technologies, and Applications*, pages 76–97. Springer.

Bo Sheng, Zheyu Wang, Yujiao Qiao, Sheng Quan Xie, Jing Tao, and Chaoqun Duan. 2023. Detecting latent topics and trends of digital twins in healthcare: A structural topic model-based systematic review. *Digital Health*, 9:20552076231203672.

Yanshan Wang, Liwei Wang, Majid Rastegar-Mojarad, Sungrim Moon, Feichen Shen, Naveed Afzal, Sijia Liu, Yuqun Zeng, Saeed Mehrabi, Sunghwan Sohn, and 1 others. 2018. Clinical information extraction applications: a literature review. *Journal of biomedical informatics*, 77:34–49.

James M Wilce. 2009. Medical discourse. *Annual review of anthropology*, 38(1):199–215.

Hanwen Ye, Tatiana Moreno, Adrianne Alpern, Louis Ehwerhemuepha, and Annie Qu. 2024. Dynamic topic language model on heterogeneous children's mental health clinical notes. *The Annals of Applied Statistics*, 18(4):3165–3184.

Ferhat Zengul, Aysegul Bulut, Nurettin Oner, Abdulaziz Ahmed, Manju Yadav, Hope Gracie Gray, and Bunyamin Ozaydin. 2023. A practical and empirical comparison of three topic modeling methods using a covid-19 corpus: Lsa, lda, and top2vec.

Yijia Zhang, Qingyu Chen, Zhihao Yang, Hongfei Lin, and Zhiyong Lu. 2019. Biowordvec, improving biomedical word embeddings with subword information and mesh. *Scientific data*, 6(1):52.

A Additional Details on Datasets

The three domains shown in Table 5, differ not only in content but also in linguistic function, offering a multi-layered view of medical discourse. Scientific texts emphasize controlled vocabulary, diagnostic precision, and incremental conceptual refinement. Clinical writing blends formal terminology with pragmatic shorthand, reflecting workflow constraints and practitioner-specific conventions. Public forums introduce narrative framing, metaphorical expressions, emotional language, and folk-taxonomies of illness, often revealing early signals of contested or emerging concepts. Together, these domains enable a cross-sectional and longitudinal analysis of semantic drift, providing insight into how medical terminology originates, spreads, and transforms across scientific, professional, and public (Social media) domains.

Dataset	**PubMed**	**Clinical**	**Pubhealth**
Domain	Scientific	Healthcare	Public
No: of Docs	22936327	540541	6935
Vocab Size	6989	7800	4083
min_df	0.1%	0.1%	1.0%
max_df	95%	95%	95%

Table 5: Statistical analysis of corpus data across three domains: scientific (PubMed), clinical (Healthcare), and public domain (Pubhealth).

B Preprocessing

To preprocess the datasets, all text was converted to lowercase, and stopwords and punctuation were removed. Tokenization was performed using SpaCy (Honnibal and Montani, 2017). In addition, words appearing in fewer than or more than the specified minimum and maximum proportions of documents were filtered out using Scikit-learn (Kramer and Kramer, 2016).

C Experiment Setup

The Continual Neural Topic Model (CoNTM) is trained using a learning rate of 0.01 and the Adam optimizer. The dataset is partitioned into 80% training, 10% validation, and 10% test sets. For all experiments, the parameters are set to $\kappa = 0.7$ and $\tau = 1$, and the number of topics is uniformly fixed at 50 across all datasets.