
THE LYME DISEASE CONTROVERSY: AN AI-DRIVEN DISCOURSE ANALYSIS OF A QUARTER CENTURY OF ACADEMIC DEBATE AND DIVIDES

SUBMITTED FOR REVIEW

 **Teo Susnjak***,  **Cole Palffy**,  **Tatiana Zimina**,  **Nazgul Altynbekova**
School of Mathematical and Computational Sciences, Massey University, Albany, New Zealand

 **Kunal Garg**,  **Leona Gilbert**
Te?ted Oy, Jyv?skyl? , Finland

April 4, 2025

ABSTRACT

The scientific discourse surrounding Chronic Lyme Disease (CLD) and Post-Treatment Lyme Disease Syndrome (PTLDS) has evolved over the past twenty-five years into a complex and polarised debate, shaped by shifting research priorities, institutional influences, and competing explanatory models. This study presents the first large-scale, systematic examination of this discourse using an innovative hybrid AI-driven methodology, combining large language models with structured human validation to analyse thousands of scholarly abstracts spanning 25 years. By integrating computational techniques with expert oversight, we developed a quantitative framework for tracking epistemic shifts in contested medical fields, with applications to other content analysis domains. Our analysis revealed a progressive transition from infection-based models of Lyme disease to immune-mediated explanations for persistent symptoms, a shift that has been particularly pronounced in high-impact clinical and immunology journals. At the same time, research supporting CLD has remained largely confined to hypothesis-driven publications, indicating a persistent asymmetry in how competing perspectives are disseminated and legitimised. The investigation into thematic trends further highlighted the enduring complexity of Lyme disease diagnostics and evolving research focus on therapeutic controversies, even as institutional alignment with PTLDS perspectives continues to grow. This study offers new empirical insights into the structural and epistemic forces shaping Lyme disease research, providing a scalable and replicable methodology for analysing discourse. The findings have implications for policymakers, clinicians, and communication strategists, emphasising the need for more equitable research funding, standardised diagnostic criteria, and improved patient-centred care models. This research also underscores the value of AI-assisted methodologies in social science and medical research by systematically quantifying discourse evolution, offering a foundation for future studies examining other contested conditions and controversies.

Keywords Lyme disease controversy · Chronic Lyme Disease (CLD) · Post-Treatment Lyme Disease Syndrome (PTLDS) · Medical controversy · AI in medical research · Large Language Models in academic analysis · Stance detection in medical literature · Lyme disease academic discourse · Science and Technology Studies · Social construction of knowledge

1 Introduction

It has been estimated that every year, Lyme disease affects hundreds of thousands in North America and Europe [1, 2], while for over a quarter of a century, the medical and scientific communities have been sharply divided over

*Corresponding author: t.susnjak@massey.ac.nz

the persistent effects of Lyme disease on patients after standard antibiotic treatments [3–9]. Although most patients recover fully, approximately 25% of patients [10–14] continue to experience symptoms like fatigue, pain, and cognitive difficulties, sparking a debate about the nature of these persistent and debilitating health issues [15–19]. This discourse has polarised into two major schools of thought: one asserts that these symptoms are the result of a post-infectious syndrome that does not involve ongoing bacterial infection [20, 21], while the other posits that a subset of Lyme disease cases may become chronic [3–6], and might require prolonged antibiotic treatment due to signs of a persistent infection [10, 22]. Research indicates that how Lyme disease is perceived is influenced by power dynamics in healthcare, patient advocacy, and media discussions, highlighting conflicts between medical experts and public knowledge [23–25].

These conflicting viewpoints have resulted in a substantial but contentious body of academic work, with researchers, healthcare providers, and patient advocacy groups frequently taking opposing positions [26–29]. Mainstream medical bodies, like the Infectious Diseases Society of America (IDSA), argue that post-treatment symptoms experienced by a subset of patients after completing standard antibiotic therapy can be attributed to what they define as Post-Treatment Lyme Disease Syndrome (PTLDS) [10]. The IDSA suggests that symptoms like fatigue and cognitive issues are probably due to immune responses or tissue damage, not ongoing infection [5]. Conversely, organisations like the International Lyme and Associated Diseases Society (ILADS) advocate for recognising chronic Lyme disease (CLD), contending that ongoing infection may be responsible for these symptoms [30]. ILADS recommends extended antibiotic regimens, pointing to contested evidence of patient improvement [31]. While deeply rooted in scientific inquiry, this debate has also been shaped by patient experiences, public advocacy, and extensive media attention, further complicating efforts to reach a consensus [27, 32–34]. To that end, this scientific controversy is also intricately linked to sociopolitical and economic influences, encompassing insurance reimbursement systems, the regulation of alternative medicine, and the stigmatisation of disputed illnesses [35, 36].

At the heart of it, the Lyme disease controversy exemplifies the sociology of medical knowledge, where grassroots patient movements contest biological authority, promoting alternative diagnosis and treatment approaches [37, 38]. Internet and social media have significantly shaped the narrative around Lyme disease [39–42]. Patients who feel unheard by traditional medicine have discovered online communities to share experiences and explore alternative treatment options. These patients, who perceive conventional healthcare as dismissive, have increasingly sought refuge in online groups, where accounts of medical neglect are validated, and alternative illness models proliferate [43, 44]. These digital platforms serve as counter-publics that challenge prevailing scientific narratives, exemplifying what researchers in Science and Technology Studies (STS) call “epistemic resistance” to dominant biomedical paradigms [45, 46]. The proliferation of conflicting medical assertions inside online Lyme disease forums has underscored the influence of digital platforms on health beliefs and patient choices [47, 48]. These platforms have amplified the voices of those advocating for chronic Lyme disease [49]. Still, according to opposing voices, they have also facilitated the spread of misinformation, further complicating and sharpening the discourse [49]. As a result, the controversy surrounding Lyme disease has extended beyond medical journals into mainstream media [39, 41], shaping public perception and influencing policy decisions.

To appreciate the current state of the discourse surrounding this controversy, it is useful to chart its history. Figure 1 captures the main themes over the decades. During the mid-to-late 1970s, an atypical outbreak of arthritis in youngsters from rural Connecticut resulted in the early identification of what would later be termed Lyme disease [52]. Shortly after its formal diagnosis, clinicians noted that several patients exhibited persistent symptoms—such as arthralgia, tiredness, and neurological complications—despite receiving prescribed antibiotic treatment [53]. Medical researchers determined that Lyme disease is an illness triggered by the bacterium *Borrelia burgdorferi*, transmitted through the bites of infected ticks, which acquire it from animals like mice and deer they feed on [54]. A hallmark of the disease in humans is a bullseye rash known as erythema migrans. Without treatment, the disease can lead to a variety of symptoms, such as joint pain and swelling, meningitis, partial facial paralysis, cognitive impairments, fatigue, headaches, and heart complications [55, 56].

In the subsequent period from the 1990s to the early 2000s, differing scientific perspectives emerged and led to opposing factions about disease chronicity and treatment, coinciding with the rise of patient advocacy and online communities that enriched public discourse. During the period from the mid-2000s to the 2010s, tensions escalated as researchers, clinicians, and patient organisations became increasingly polarised, while media coverage intensified sensationalism surrounding the chronic Lyme disease controversy [20, 26, 50, 51]. At this time, patient organisations began to assert that the treatment protocols established by the IDSA had been employed to refuse medical insurance for those seeking extended antibiotic courses. At the same time, it also penalised physicians who prescribed these treatments [26]. The discourse became inflammatory with terms like “Axis of Evil” used to describe physicians prescribing prolonged antibiotics, “specialty laboratories” offering alternative tests, and the internet’s role in promoting “Lyme hysteria” [50], reflecting the intensifying acrimony and the frustration within the mainstream medical community. While it has been suggested [51] that Lyme disease was receiving an over-proportionate exposure in media coverage, the claim was

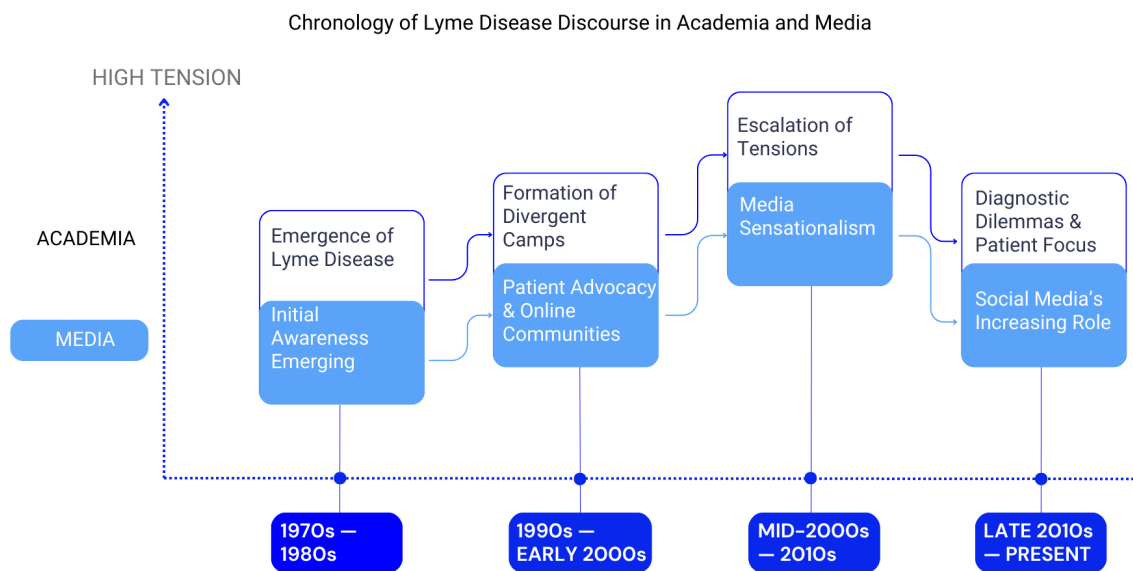


Figure 1: A broad and consolidated outline of the discourse themes and tensions on the Lyme disease controversy in media and academia over time as reported in literature [20, 26, 50, 51].

that the coverage often perpetuated this perspective, portraying those advocating for CLD as misguided or promoting unscientific practices.

The cooling in tensions began to emerge only from the late 2010s, when the focus shifted to diagnostic problems and patient-centred perspectives, with social media playing a substantial role in shaping public understanding, activism, and patient support [57]. As the debate continued, there has been a further shift in focus in recent years, emphasising the complexities of Lyme disease diagnostics and patients' subjective experiences. Studies now point out the challenges with current blood tests, especially in the advanced stages of the disease [54]. Researchers acknowledge the need for improved diagnostic tools to address potential underdiagnosis, particularly in cases with atypical symptoms or less common *Borrelia burgdorferi* species [58]. Qualitative studies using interviews and ethnographic methods offer valuable insights into the experiences of individuals with persistent symptoms attributed to Lyme disease [15]. These studies explore the challenges of navigating a complex medical system while facing scepticism and dismissal from some healthcare providers. What began as a relatively straightforward debate over diagnosis and treatment has expanded into a complex and often polarised controversy that touches on issues of medical authority, patient autonomy, and the role of the media in shaping public understanding of health[59], deserving of a comprehensive investigation.

This persistent discourse on CLD and PTLDS is crucial to health communication, as it highlights the downstream effects of academic research on media narratives, which in turn affects public attitudes and, ultimately, behaviours regarding their health choices. It is equally important to consider the public's trust in scientific expertise and healthcare institutions in the face of stark disagreements, whose erosion of credibility can also result in various health behaviours, including non-compliance with standard treatments, seeking unverified or potentially harmful remedies, and withdrawing from conventional healthcare systems [60]. Since such medical conflicts seldom exist solely within the realm of scientific discourse but are instead socially produced through the interplay of scientific institutions, policymakers, media, and the public [54, 55, 61, 62], Lyme disease is, therefore, a prime example of a contested sickness, illustrating how medical ambiguity generates competing knowledge claims and varied treatment paradigms [24].

This controversy, therefore, is important and extends beyond scientific inquiry, shaping public discourse and healthcare behaviours. Media narratives, spanning news reports, documentaries, and digital platforms, often amplify conflicting perspectives, influencing public attitudes toward diagnosis and treatment [29, 30]. As a result, patients navigating uncertainty may turn to alternative sources when they perceive mainstream reporting as dismissive, potentially leading to non-adherence to conventional medical guidance or the pursuit of extended antibiotic regimens [63, 64]. A critical dimension of this debate is the erosion of trust in scientific expertise and healthcare institutions. Conflicting guidelines

from organizations such as IDSA and ILADS contribute to uncertainty, influencing patient behaviours that range from scepticism toward standard treatments to reliance on unverified or experimental therapies [28, 60]. This complex interplay between medical uncertainty, media influence, and public trust underscores the need for strategic health communication approaches that integrate evidence-based insights with patient experiences.

Contribution and novelty

Despite the wealth of research and debate, no comprehensive and systematic synthesis has been conducted to map the evolution of academic discourse on Lyme disease over the past 25 years [65]. The goal of this study has been to fill this gap, provide practical insights into the development of the scientific discourse on this controversy, and identify overarching trends. The key contribution of this study lies in examining over a thousand relevant academic studies spanning the past quarter-century. The novelty is centred around leveraging the latest advancements in AI technologies, employing the most sophisticated LLMs to date to perform stance and viewpoint detection expressed within these studies, allowing the extraction of both explicit positions and more nuanced sentiments contained in these texts surrounding the controversy over CLD and PTLDS. Furthermore, we developed a novel hybrid AI-driven methodology that enabled us to automate the processing of large volumes of text to map the discourse of this controversy and its evolution over time. Our work not only conducted a comprehensive analysis of how viewpoints have shifted over time and how different journals have platformed differing positions, but we also extracted major themes charting their development over time through to the present status of the debate. This timely integration of computational intelligence with medical research provides unprecedented insights into this long-standing controversy and establishes a robust AI-driven methodology to address complex debates in healthcare literature at scale.

2 Related Works

In recent years, research on Lyme disease has examined its clinical, epidemiological, and sociocultural dimensions. These studies have ranged from broad narrative and scoping reviews investigating overarching challenges to systematic reviews offering targeted insights into specific aspects of diagnosis, treatment, and disease mechanisms. The literature covered several key areas: treatment and management, pathology and disease mechanisms, diagnostic controversies and challenges, epidemiology and public health as well as the sociological perspective.

2.1 Treatment and Management

The efficacy of antibiotic treatment for PTLDS has been a persistent issue where early research questioned the rationale for prolonged antimicrobial therapy. Lantos [66] conducted a systematic review that evaluated the role of chronic co-infections such as *Babesia*, *Anaplasma*, and *Bartonella*, concluding that no compelling evidence supported their role in PTLDS or CLD. This finding challenged earlier assertions that lingering symptoms were due to persistent infections requiring long-term antibiotic regimens. Subsequent studies reinforced these findings. Rebman and Aucott [67] reviewed PTLDS from a mechanistic perspective and argued that persistent symptoms were more likely linked to immune dysfunction and neural sensitisation rather than ongoing infection, suggesting a shift from pathogen-focused treatments. Dersch et al. [68] further challenged the antibiotic paradigm, finding no statistically significant benefit of antimicrobial therapy on quality of life, cognition, or depression while reporting an increased incidence of adverse effects. Beyond clinical efficacy, the risks of overtreatment became more prominent in other reviews. Sébastien et al. [69] conducted a systematic review and found that misdiagnosis of PTLDS was widespread, ranging from 80% to 100% of suspected cases being incorrectly classified, leading to unnecessary and sometimes harmful antibiotic treatment. In addition, Mattingly and Shere-Wolfe [70] evaluated the economic burden of Lyme disease and revealed substantial healthcare costs associated with misdiagnosis and overtreatment. At the same time, Van Hout [71] highlighted that despite the increasing global incidence of Lyme disease, pharmaceutical investment in its treatment was lacking. Additionally, the authors claimed that the evidence and international guidelines for managing CLD remained conflicting and controversial, posing challenges to public health policy and clinical practice.

2.2 Pathology and Disease Mechanisms

The biological mechanisms underlying PTLDS and CLD have also remained an area of significant debate over the past few decades. Marques [11] provided an early framework for distinguishing between different patient groups diagnosed with CLD, recognising that many individuals lacked objective evidence of active *Borrelia burgdorferi* infection. This categorisation laid the groundwork for later investigations into the nature of persistent symptoms. Borgermans et al. [72] expanded upon the early framework by [11] exploring CLD as a multifaceted clinical entity. The review by Borgermans et al. [72] suggested that CLD remained poorly understood, with ongoing debates regarding its definition, diagnosis,

and treatment. Mac et al. [16] further contributed to this discussion by conducting a systematic review that documented the long-term effects of Lyme disease, reporting that patients frequently experienced fatigue, musculoskeletal pain, and cognitive impairment, though the exact mechanisms remain uncertain. Beyond symptom classification, recent works have examined the complex biological interactions underlying *Borrelia burgdorferi* persistence and disease progression. Bamm et al. [73] provided an integrative review of *Borrelia burgdorferi* biology, host-pathogen interactions, and immune evasion strategies, emphasising that the spirochete's ability to modulate host responses may contribute to prolonged symptoms even after standard treatment. Their findings reinforced earlier hypotheses that PTLDS symptoms could stem from immune dysregulation and chronic inflammation rather than ongoing infection. Focusing on the neuropsychiatric dimensions of Lyme disease, Brackett et al. [74] linked infection to increased risks of cognitive decline, anxiety, and depression. Recently, Bobe et al. [58] reviewed the progress and understanding of Lyme disease in the five years preceding their study and explored the role of immune dysregulation and potential autoimmune triggers, arguing that PTLDS symptoms were more likely driven by sustained inflammation rather than persistent infection.

2.3 Diagnostic Controversies and Challenges

A longstanding source of clinical and research debate has also surrounded the challenges in diagnosing Lyme disease. Brunton et al. [57] systematically reviewed stakeholder perspectives and found widespread dissatisfaction with existing diagnostic tools. Their study highlighted the disconnect between clinician scepticism and patient experiences, which, according to the authors, frequently led to diagnostic uncertainty and strained doctor-patient relationships. Studies have investigated the specific limitations of current diagnostic methods, and thus, the diagnosis of Lyme disease remains contentious, which is marked by significant challenges in clinical practice and research. Diagnostic uncertainty frequently arises from dissatisfaction with existing testing methods, highlighting marked discrepancies between clinical and patient experiences [57]. Conventional diagnostic tests, primarily the two-tiered approach of enzyme immunoassay (ELISA) followed by immunoblotting, have notable limitations, especially in early disease stages, leading to delayed treatments and misdiagnoses [75, 76]. Additionally, regional variation in *Borrelia* genospecies poses substantial obstacles, as standard assays often fail to detect less common strains, further complicating diagnosis [76]. These diagnostic challenges extend into the debate surrounding PTLDS and CLD. Divergent perspectives among medical communities exacerbate the controversy, notably seen in conflicting guidelines from influential organisations such as the IDSA and the ILADS. Such conflicts can result in misdiagnoses, inappropriate therapies, and decreased trust in healthcare institutions [59, 77, 78]. Furthermore, geographical variations in disease presentation, including Lyme-like illnesses without confirmed local *Borrelia* infections, add layers of complexity to accurate disease identification and management [77]. Studies have noted that resolving these diagnostic and therapeutic challenges necessitates more accurate biomarkers and standardised diagnostic protocols to improve early detection and patient outcomes [75].

2.4 Epidemiology and Public Health

Other research reviews on the study of Lyme disease have focused more on the driving environmental factors, whereas Stone et al. [79] took an ecological perspective, identifying climate change and tick habitat expansion as key factors influencing Lyme disease incidence. Dong et al. [80] conducted a global meta-analysis estimating that 14.5% of the population had been exposed to *Borrelia burgdorferi*, with the highest seroprevalence in Central Europe (20.7%) and Eastern Asia (15.9%). This study built on earlier epidemiological assessments, such as those by Van Hout [59], who also identified climate change and tick population dynamics as key factors in the disease's increasing range. Meanwhile, other studies [81] focused on the economic burden imposed by Lyme disease, concluding that it is significant, particularly in the US, and to that end, justifying further research efforts in disease control and management. Mattingly and Shere-Wolfe [70] also evaluated the financial burden of Lyme disease, highlighting significant healthcare costs and productivity losses. Bobe et al. [58] noted in the U.S. that federal funding for Lyme disease research remained disproportionately low relative to its public health impact, with increasing reliance on private philanthropic contributions.

2.5 Sociological Perspective

Peretti-Watel et al. [8] examined Lyme disease as a case study in medical controversy, showing how public mistrust in health authorities fuelled competing narratives about the disease's prevalence and management and exacerbated by ongoing conflicts between the IDSA and ILADS organizations, which has created ambiguity concerning optimal practice guidelines [10, 28]. Olechnowicz et al. [82] introduced a sociological perspective by examining how perceptions of vulnerability influence individual behaviours and trust in information sources regarding Lyme disease. Their study validated a novel vulnerability scale, demonstrating that emotional discomfort and perceived susceptibility shape engagement in preventive behaviours, highlighting the role of psychology and public trust in shaping disease management strategies.

Studies by Pascal et al. [49] and Uzzell et al. [83] emphasised the impact of media and public health communication on societal attitudes, portraying Lyme disease as either a disregarded epidemic or an exaggerated controversy, highlighting its depiction as a socially constructed phenomenon shaped by media narratives, patient advocacy, medical uncertainty, and institutional prejudices. Puppo et al. [84] and Hinds and Sutcliffe [85] brought to focus the distinction between conventional medical institutions and alternative “Lyme-literate” viewpoints, highlighting the knowledge disparity between orthodox and heterodox discourses, while Rebman et al. [15], Baarsma et al. [86] conducted concurrent research on patient experiences, highlighting issues on medical legitimacy, diagnostic ambiguity, and the psychological ramifications of disputed illness status. The increasing confluence of medical authority, patient empowerment, and online activism as examined in [87–89], demonstrated how self-advocacy organisations challenge institutional authority while promoting scientific fragmentation. Meanwhile, Bloor et al. [90] demonstrated the influence of scientific uncertainty and policy inconsistency on institutional decision-making and advocacy around Lyme disease. These studies jointly contributed towards perceiving that Lyme disease is not merely a medical ailment but a politicised and socially contentious illness, highlighting the broader conflicts among research, politics, and patient-centred healthcare. To that end, it has been postulated [54, 55, 61, 62] that medical conflicts seldom exist solely within the realm of scientific discourse; instead, they are socially produced through the interplay of scientific institutions, policymakers, media, and the public. Lyme disease is thus a case in point, illustrating how medical ambiguity generates competing knowledge claims and varied treatment paradigms [24].

Pascal et al. [49] demonstrated how media discourse has significantly contributed to the perception of Lyme disease as a societal issue, bolstering biomedical scepticism and patient activism. Puppo et al. [84] illustrated that Lyme-literate medical professionals had established alternative epistemic networks that challenge prevailing medical paradigms and promote unconventional treatment protocols. Moreover, internet platforms have revolutionised the discourse surrounding Lyme disease, serving as “knowledge enclaves” as characterised by Brown [91], where scientific credibility is reinterpreted through collective patient experiences rather than peer-reviewed research [92]. This corresponded with extensive sociological research about disseminating health-related misinformation in digital contexts, reinforcing health attitudes that deviate from conventional medical guidelines [48, 93].

2.6 Study Research Questions

While previous reviews have addressed the clinical, epidemiological, and sociocultural dimensions of Lyme disease, they have not systematically synthesised how the academic discourse on CLD and PTLDS has evolved. This study fills that gap by employing a large-scale, AI-driven approach to map thematic trends and stance distributions. It sought to identify and track shifts in scholarly perspectives, leading to the following key research questions:

- **(RQ1)** How has the academic discourse on CLD and PTLDS evolved over the past 25 years regarding research volume, thematic focus, and stance distribution?
- **(RQ2)** How do journal specialisation and editorial focus influence the representation of perspectives on CLD and PTLDS in the peer-reviewed literature?
- **(RQ3)** What are the dominant thematic structures within the Lyme disease debate, and how do they correspond to competing explanatory models and levels of scientific consensus?

3 Theoretical and Methodological Grounding in Science and Technology Studies (STS) Framework

This study employed a Science and Technology Studies (STS) framework [94] in conjunction with computational AI models to analyse the Lyme disease controversy. STS offers a critical lens for understanding how social, cultural, and historical factors shape scientific knowledge, particularly in biomedicine and contested illnesses [46, 95–97]. For complex and contested illnesses like Lyme disease, an STS framework is particularly valuable in enabling the examination of the interplay of social dynamics, power relations, and the very construction of medical knowledge itself, while pairing this approach with the latest AI-automation and reasoning technologies for information extraction at scale. Key principles and concepts from STS that guided our overall approach in formulating the technical aspects of our methodology in Section 4, include:

- **Selective Social Construction (STS Principle):** STS highlights that social shaping varies across scientific fields. While physics is empirically constrained, biomedicine, addressing complex human systems, is more open to social interpretation. Even within biomedicine, diagnostic categories (PTLDS vs. CLD) are more socially negotiated than, for example, the molecular biology of *Borrelia burgdorferi*.

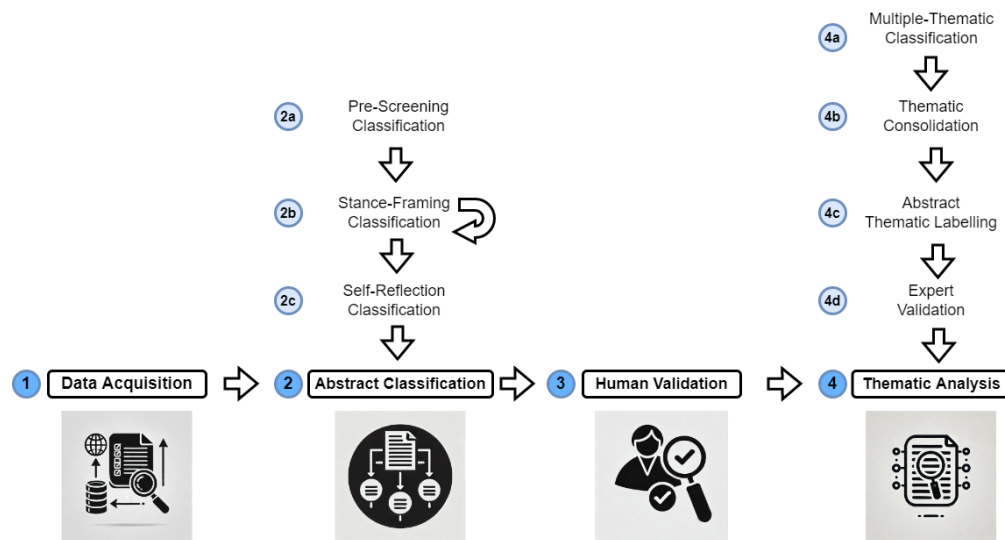


Figure 2: Overview of the steps comprising the proposed hybrid AI-driven content analysis methodology.

- **Empirical Constraint and Real Phenomena (STS Tenet):** Our STS framework acknowledges that social construction does not equate to extreme relativism or a dismissal of empirical reality. Empirical observation, clinical data, and technological applications do constrain scientific theories. In the Lyme context, STS helps us analyse how the interpretation of patient symptoms is contested while acknowledging the underlying reality of patient suffering and the biological basis of Lyme infection as legitimate areas of scientific inquiry.
- **Social Values and Epistemic Commitments in Contested Fields (STS Focus):** Lyme disease controversies highlight paradigm clashes and differing epistemic commitments, not just factual disagreements [98]. Competing values, clinical priorities, economic motivations and institutional affiliations influence research agendas and data interpretation, especially in areas lacking scientific consensus.
- **Discourse Analysis as an STS Methodology (STS Method):** Computationally enhanced discourse and framing analysis were essential to reveal social influences in our study. By systematically analysing language, framing, and thematic patterns in Lyme disease literature, guided by an STS framework, we sought to expose the social processes that have historically shaped the scientific conversation and contributed to the enduring Lyme disease controversy.

The adoption of the STS framework ensured a theoretically grounded and methodologically rigorous analysis of the Lyme disease debate was conducted. The STS framework informed both our technical methodological implementation, in aspects such as prompt engineering outlined in Section 4.2 as well as in the formulation of the thematic analysis of Section 4.4, but also the interpretation of results, allowing us to analyse the Lyme controversy as a socially constituted and knowledge-producing phenomenon, beyond purely biomedical or clinical perspectives. The integration of STS with medical sociology [99], framing theory [100], and patient experience research [15] collectively offered a rich lens and social science framework for understanding this complex and contested illness.

4 Methodology

The technical aspects of our methodology combined the social science framework with cutting-edge AI technologies, specifically focusing on LLMs and their emerging reasoning capabilities that venture well beyond classical machine learning approaches [101]. This methodological choice was motivated by the need to analyse a vast and complex body of text, exceeding the capacity of traditional qualitative methods alone. While qualitative approaches are crucial for in-depth analysis of social meaning, LLMs and their sophisticated reasoning abilities offer a scalable and systematic way to identify semantic patterns and trends across a large corpus of scientific abstracts, enabling analyses of shifts in thematic focus, stance distribution, and journal-level preferences on specific topics that would be challenging through manual review [102]. Increasingly, LLMs are being deployed in research for these kinds of tasks [103]. Since LLMs exceed the capabilities of traditional NLP techniques, their value is found in their ability to surface implicit assumptions and underlying frames within the discourse, which can contribute to a deeper understanding of the social construction of knowledge in this contested medical field [104]. We acknowledge the methodological challenges of using AI in

social science research [105]; therefore, we prioritise transparency² and validation throughout our methodology, aiming to complement and provide new approaches for scalable research rather than replace established qualitative approaches to discourse analysis. Our entire methodological approach can be summarized in four key steps that include data acquisition (Step 1), abstract classification using automated approaches (Step 2), validation of the automation process via human experts (Step 3) and finally, the thematic analysis of the dataset including the human-in-the-loop verification (Step 4). These four stages and the sub-steps are visualized in Figure 2, according to which the remainder of this section is organized.

4.1 Dataset acquisition - Step 1

Step 1 in Figure 2 represents the entire data acquisition process in detail. The data was collected for a time period from 2000 to 2024 using the Publish or Perish (PoP) [106] software for paper search³. Academic databases used for paper searches were Google Scholar, Scopus, PubMed, CrossRef, Web of Science, and Semantic Scholar. Search keyword combinations were used focusing on topics related to chronic Lyme disease and post-treatment Lyme disease syndrome, and the search was restricted to each year individually due to a large number of returned results. The search term *lyme* was used for paper titles. In contrast, various combinations of terms were used for matches across the paper documents comprising *disease, borreliosis, borrelia, chronic, controversy, post-treatment, PTLDS, acute, syndrome, and post-Lyme*. The initial search produced a dataset of 84,140 papers covering a 25-year timeframe; however, a large proportion of abstracts were missing from this initial data collection process. Notably, Google Scholar contributed 41% of all the retrieved records. Figure 3 summarises the proportion of papers acquired by the database. The acquired dataset from PoP comprised the following fields for each paper:

- publication: the name of the publication source.
- authors: the list of authors who contributed to the article.
- year: the year the article was published, from 2000 to 2024.
- type: paper type, i.e. article or review etc.
- abstracts: the text of the abstract for each article - mostly unpopulated from the PoP search results.
- cites: the number of times that the paper has been cited.

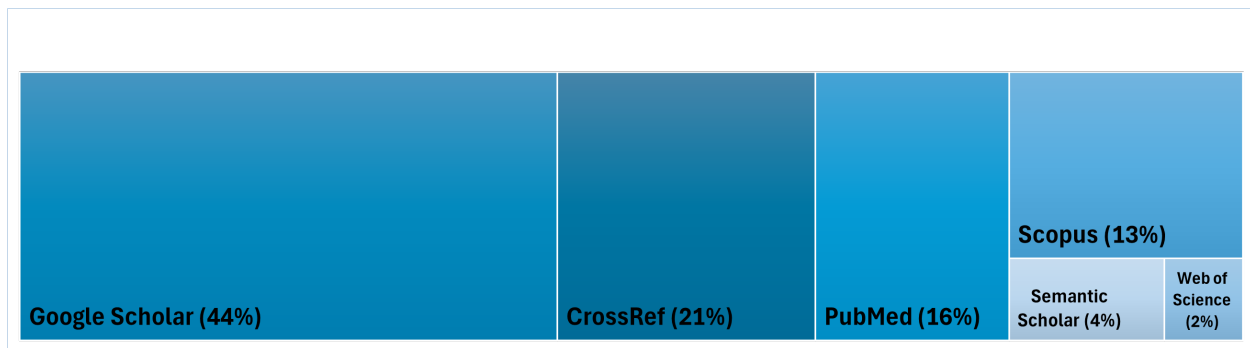


Figure 3: Proportion of retrieved articles per database

Data screening and filtering

Next, the PRISMA flow diagram in Figure 4 details the pre-processing, screening and the filtering process. Records with missing or duplicated DOIs (digital object identifier) were excluded from the dataset. More than 38,000 records were removed, leaving 45,271 records for screening. Subsequently, approximately 27,180 records were excluded due to missing publication names, article titles or abstracts. Python scripts were written to automatically retrieve missing abstracts where possible using APIs and DOIs⁴ resulting in 7,360 previously missing abstracts. Approximately 7,528 records were excluded due to irretrievable abstracts, non-English text, and lacking relevant search terms. Consequently, the resulting dataset comprised 8829 potentially relevant abstracts, requiring more detailed screening and analysis for relevance.

²The datasets, LLM prompts and outputs can be found at <https://github.com/teosusnjak/Lyme-disease-controversy>

³This data collection builds on and extends the earlier work [107].

⁴Scopus API Springer Nature API

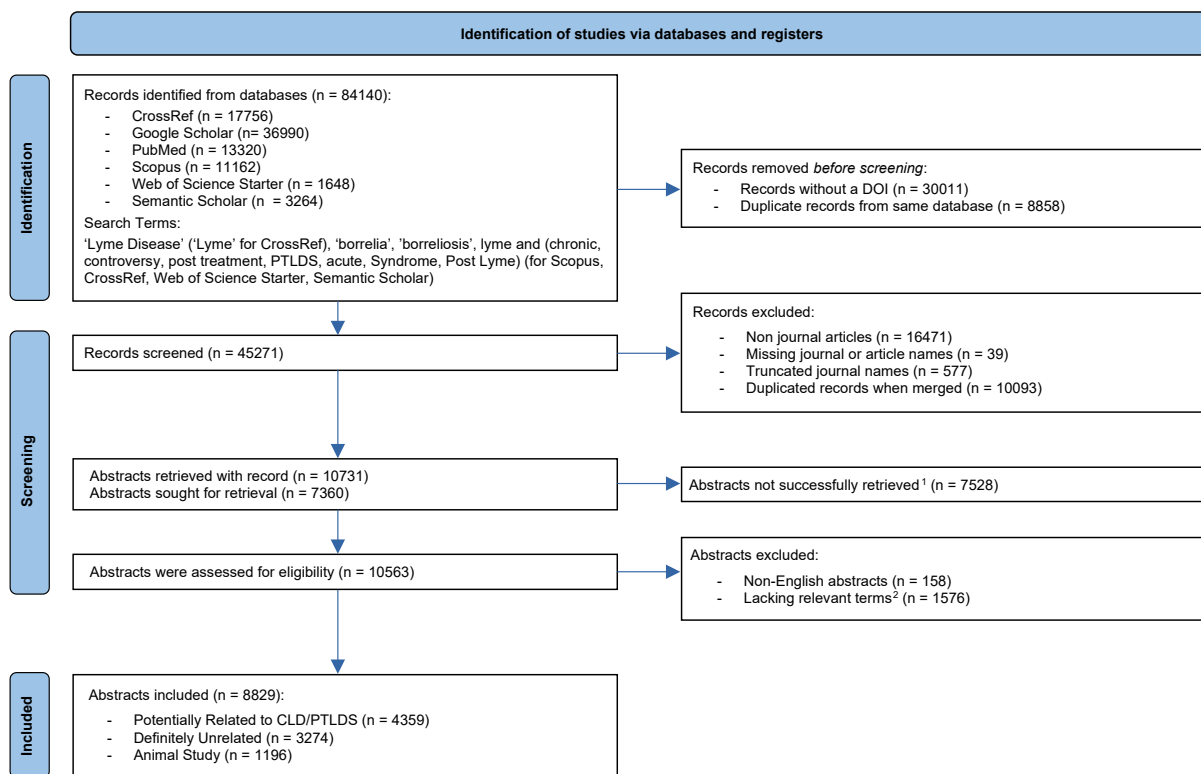


Figure 4: PRISMA flow diagram for new systematic reviews, which included searches of databases and registers only. ¹This number includes 2088 abstracts with less than 300 characters in length which were rejected due to the information content being too low for analysis. ²Lyme, Borrelia*, burgdorferi, Ixodes, Erythema, migrans, tick-borne, tickborne, tick borne.

4.2 Abstract Classification - Step 2

Following the initial screening and filtering process that relied on a straightforward mechanical application of inclusion and exclusion rules, 8829 records remained, which needed a deeper semantic analysis of the content to be relevant to the PTLDS/CLD controversy. Three sub-steps with increasing sophistication were applied. Initially, an automated pre-screening classification (Step 2a) approach was applied leveraging a reasoning LLM (OpenAI's GPT-4o-mini). The purpose of this task was to eliminate the most obviously irrelevant abstracts for which a very high level of confidence in their rejection was achieved while retaining others for a deeper analysis. The aim was to remove all abstracts with a clear focus on animal studies and those that focus on Lyme disease, but hold no relevance to the PTLDS/CLD discourse. Therefore, each abstract was classified into one of three predefined classes: *Potentially Related to CLD/PTLDS*, *Definitely Unrelated*, or *Animal Study*. This approach is supported by literature, where machine learning has been successfully used to classify medical texts for some time [108]. As more sophisticated models have emerged, medical text classification has improved considerably [109], and LLM models have thus been applied to aid in systematic reviews of the literature [110, 111]. Notably, due to the speed and cost, LLMs have even been found to be more accurate overall than expert annotators at classifying texts [112].

Each abstract was processed individually via API calls using the classification prompt in Appendix A. The model generated a JSON output for every input containing the abstract's index, a classification, and a confidence score (*High*, *Medium*, or *Low*). Abstracts classified as *Potentially Related to CLD/PTLDS* were retained for further, more refined processing and classification. Low-confidence non-*Potentially Related to CLD/PTLDS* classifications were flagged for additional classification validation to ensure comprehensive coverage and minimise false exclusions. This low-resolution pre-screening classification step effectively categorized the 8,630 abstracts as follows: 4,160 (48.2%) *Potentially Related to CLD/PTLDS*, 3,274 (37.9%) *Definitely Unrelated*, and 1,196 (13.9%) *Animal Study*. Thus, all

Animal Study and *Definitely Unrelated* abstracts receiving a Medium to High classification confidence level were eliminated from further analysis.

Stance-Framing Classification - Step 2b

In Step 2b, the actual classification of the remaining 4,160 abstracts into target categories that support the study's goals commenced. In this step, the authors' implicit or explicitly stated position on the controversy needed to be identified, requiring both sophisticated reasoning and domain knowledge.

From Stance to Frame Detection

Sentiment analysis or opinion mining is often used to interpret an author's feelings on a subject [113]. Frequently, the sentiment is determined by the number and strength of positive and negative words and which parts of speech are used. However, traditional sentiment analysis cannot recognise implicit meaning in text, frequently resulting in incorrect interpretations [113]. Sentiment analysis is ill-suited for the task of this study also because academic articles are expected to employ near-neutral expressions, while domain-specific terms such as 'disease' are neutral and objective in this context, while interpreted negatively by standard sentiment analysis models, thus skewing results for this oft appearing term. Stance detection, on the other hand, aims to use machine learning models to automatically determine the position or attitude expressed in a text towards a specific target concept, event, or entity and seeks to identify whether the text favours, is against, or is neutral towards the target [114]. Recent studies have demonstrated the effectiveness of LLMs in stance detection tasks with high reliability and accuracy [115]. Therefore, this research also leveraged technology for this task. Moreover, in Step 2b, our study conceptualised 'stance' as reflecting a more complex orientation towards the Lyme disease controversy. Drawing on framing theory [100] and discourse analysis [116], we defined 'stance' more broadly as the underlying perspective or interpretive *frame* and underlying perspectives adopted by authors in relation to the CLD/PTLDS debate – not merely positive/negative sentiment. Therefore, this included explicit agreement or disagreement with particular positions and implicit assumptions, preferred modes of reasoning, and the values and priorities emphasised in the authors' arguments. For example, a 'PTLDS-supporting stance' might be characterised by framing persistent symptoms as primarily immune-mediated, relying on epidemiological evidence, and prioritising mainstream clinical guidelines. Conversely, a CLD-supporting stance might frame persistent symptoms as indicative of ongoing infection, emphasise patient narratives and anecdotal evidence, and prioritise alternative treatment approaches. Therefore, by using LLMs for stance or *frame* detection, we aim to identify these cues and more nuanced framing patterns within the academic discourse, revealing the underlying epistemological and ideological dimensions of the Lyme disease controversy. To accomplish this, we engaged with subject experts in an iterative refinement process (highlighted in Step 2b in Figure 2) to design an LLM prompt capable of performing this complex task, which also considers our STS framework. The final classification prompt can be seen in Appendix A, which set definitions, classification criteria and few-shot in-context learning, to classify 4,160 abstracts into the following categories: "Supports PTLDS", "Supports CLD", "Neutral", "Unrelated", or "Animal Study". For each classification, the LLM needed to provide an accompanying justification text explaining the reasoning for its classification decision and a confidence level. An example of two abstracts, together with classification and confidence categories, as well as the justification texts, can be seen in Appendix B in Figures 11 and 10). The results from this classification step yielded a total of 1,033 abstracts of interest that fell into the target categories of "Supports PTLDS", "Supports CLD" or "Neutral".

Self-Reflection Classification - Step 2c

The classifications and the justifications from Step 2b were then reevaluated and reassessed using automated methods and corrected where necessary to improve accuracy. To achieve this, self-reflection prompting technique was used [117]. Self-reflection prompting is a technique where LLMs re-assess and refine their initial outputs to identify and correct errors, improving overall reasoning and decision-making capabilities. This method has significantly improved problem-solving performance in LLMs [117]. After executing this step, we found that only 5.3% (229 out of 4359) of the classifications changed. The largest categories to undergo re-classifications were "Neutral" (98), where 60% flipped to "Supports PTLDS" upon revision, and the "Unrelated" category (64), where 89% subsequently became "Neutral". Only two abstracts underwent a significant revision from initially "Supports CLD" to subsequently "Supports PTLDS", thus confirming the stability and consistency of the classifications and the approach used. Several examples of abstract classifications and their adjustments in both the final label and justification texts can be seen in Appendix A.1.

4.3 Human Validation - Step 3

Following the final classifications, we assessed their reliability and their overall alignment with human expert judgments. To accomplish this, we conducted a comprehensive Inter-Rater Reliability (IRR) analysis [118, 119]. We sought to

compare and establish the degree of agreement between the classifications (1) of our chosen LLM and those of two subject-expert raters, (2) between the two human raters, and (3) between all classifications, including six additional LLMs. Therefore, this evaluation included both pairwise agreement (using Cohen’s Kappa [118]) and multi-rater agreement (using Fleiss’ Kappa [119]). Ultimately, the goal was to determine whether the LLM-based classifications exhibited sufficient reliability comparable to human expert judgement, which would validate the methodology and the experimental results. The six additional cutting-edge and most advanced LLMs to date included in this validation were Google’s Gemini 2.0 Flash (plus the model in the Thinking Mode), Anthropic’s Claude 3.5 Sonnet, DeepSeek’s R1 model, Alibaba’s Qwen 2.5 Max, X’s Grok-3. The key metric used to measure inter-rater agreement between two raters while adjusting for chance agreement was Cohen’s Kappa [118], defined as:

$$\kappa = \frac{P_o - P_e}{1 - P_e} \quad (1)$$

where P_o is the observed agreement, and P_e is the expected agreement by chance. It ranges from -1 (complete disagreement) to 1 (perfect agreement), with standard interpretation thresholds [120] from poor ($\kappa < 0.00$) to almost perfect ($\kappa \geq 0.80$). Fleiss’ Kappa generalises Cohen’s Kappa to multiple raters, providing a single measure of agreement across all LLMs and human raters. Higher kappa values (max 0.8) indicate strong agreement, reinforcing the validity of automated classification. Low values (< 0.4) suggest significant inconsistencies, highlighting areas of marked divergence.

Agreement Between Human Raters and LLMs

To establish a baseline for IRR, we first assessed agreement between two domain-expert human raters from the author team. Prior to classification, both raters underwent training to ensure consistency in applying the classification framework (see stance-framing classification prompt in Appendix A). This entailed meetings and the provision of a classification training pack. The human raters were provided with the same prompts, criteria, and assumptions used by the LLMs, along with examples of their classifications for reference. A custom web application was developed for independent annotation (see Appendix B, Figures 11 and 10), which enabled raters to classify a random sample of 150 abstracts. For each abstract, the raters selected a classification and chose between two possible justification options for their most suitable decision. Cohen’s Kappa ($\kappa = 0.501$) indicated moderate agreement, aligning with established IRR benchmarks in literature [121–123] and prior studies in subjective classification tasks such as qualitative content analysis, medical diagnoses, and thematic coding [124]. Perfect agreement is rarely expected in complex classification tasks due to interpretive differences and variations in emphasis on textual elements.

Next, we evaluated the agreement between the final revised classifications (produced via GPT-4o-mini’s self-reflection process referred to below as ‘GPT’) and a set of alternative LLM-generated classifications (Gemini, Gemini-Thinking, Claude, DeepSeek, Qwen, Grok-3), alongside the original LLM classification and those of the two human raters. Table 1 presents Cohen’s Kappa for each pairwise comparison.

Comparison	Cohen’s Kappa
GPT vs. Original GPT Classification	0.767
GPT vs. Gemini	0.717
GPT vs. Gemini-Thinking	0.608
GPT vs. Qwen	0.600
GPT vs. Claude	0.592
GPT vs. Human Interrater 1	0.583
GPT vs. Human Interrater 2	0.508
GPT vs. DeepSeek	0.475
GPT vs. Grok-3	0.458
Human-Human Agreement	0.501

Table 1: Cohen’s Kappa agreement between GPT revised classifications after self-reflection and all other raters/classifications

Key observations emerge from validation results in Table 1:

- The strongest agreement occurred between the original and GPT’s revised classifications ($\kappa = 0.767$), indicating that the LLM self-reflection process refined classifications to a limited degree but did not fundamentally alter them. This suggests that initial classifications were internally consistent, requiring only minor adjustments. This, therefore, supports the claim that there is evidence of an underlying **stability** in the approach used, confirming both the leveraging of LLMs for this task and the suitability of the designed prompts.
- Human vs. GPT’s Revised Classification: The agreement between GPT’s revised classifications and human raters ($\kappa = 0.583$ for Interrater 1, $\kappa = 0.508$ for Interrater 2) is comparable to human-human agreement ($\kappa = 0.501$). This means that the GPT’s revised classifications are at least as consistent with human judgment as human raters are with each other, thus reinforcing the **validity** of LLM’s classifications. Additionally, on the subset of abstracts where both human raters agree, their IRR with GPT’s classification was very high at $\kappa = 0.709$. Notably, with respect to LLM model choices, we found that the highest IRR values for both human raters were with GPT and Original GPT classifications ($\kappa = 0.583$ and $\kappa = 0.538$), thus also confirming the **suitability** of the chosen type of LLM for the classification tasks.
- The highest agreement with an alternative model occurred with Gemini ($\kappa = 0.717$), followed by Qwen ($\kappa = 0.600$). This suggests that these models exhibit classification patterns closest to the GPT’s revised outputs, likely due to similarities in training data or reasoning heuristics.
- The lowest agreement is observed with Grok-3 ($\kappa = 0.458$) and DeepSeek ($\kappa = 0.475$), indicating notable divergences in their classification outputs. These discrepancies may reflect different conceptual representations of Lyme disease controversies across models and differences in the underlying datasets used for their training, suggesting that these models are not the best candidates for this task.

For completeness, to assess global agreement among all human and automated classifiers, we computed Fleiss’ Kappa ($\kappa = 0.537$). This result indicates moderate agreement, aligning with the baseline human-human agreement levels and again reinforcing the consistency of the classification framework as a whole. Fleiss’ Kappa further validates the LLM-assisted classification methodology by showing that the aggregate classification structure remains coherent despite inherent variability across all models. This aligns with prior research in natural language processing and thematic coding, where moderate agreement is a reasonable outcome in subjective classification tasks [123, 124]. Given the comparable and even higher IRR values between human raters and the GPT’s classifications concerning agreement values between the human experts, the IRR analysis confirmed the reliability and methodological rigour of our classification process and the underlying prompts. The alignment between human experts, LLM classifiers, and self-refined classifications demonstrates that LLM-assisted classification of stance or framing in the abstracts is systematic and replicable.

Finally, we also examined the validity of the justification texts generated by the LLM to provide a rationale for the classifications given to each abstract. The analysis found that the human raters achieved a Cohen’s Kappa of $\kappa = 0.61$, while both human raters scored $\kappa = 0.71$ against the LLM, again demonstrating that the human raters agreed to a higher degree with the LLM’s outputs, then with each other’s choices.

4.4 Thematic Analysis - Step 4

Identification and Classification of Overarching Themes

After all the abstracts were classified into predefined stance categories (*Supports PTLDS*, *Supports CLD*, and *Neutral*), resulting in 1,033 abstracts, a hybrid computational and a theoretically informed thematic analysis overseen by subject experts was conducted to extract deeper conceptual insights into the dominant lines of reasoning and underlying social dynamics within the Lyme disease controversy. This analysis aimed to identify recurrent patterns within the justifications for classification and the abstracts themselves, ensuring a systematic, reproducible, and theoretically grounded approach to the study of contested medical narratives consistent with established social science methodologies for discourse analysis.

Step 4 in Figure 2 depicts this multi-stage, iterative thematic identification process we employed for combining automated reasoning using multiple LLMs with structured reconciliation and expert human validation that balances computational scalability with interpretive depth. The hybrid approach enabled us to leverage the strengths of both computational pattern recognition at scale and qualitative interpretation of LLMs together with that of the critical human review of subject experts, while integrating refinement and validation in the process to ensure social science validity and conceptual richness. This approach acknowledges that while LLMs can efficiently process large volumes of text, human expertise remains crucial for interpreting social meaning and ensuring theoretical coherence within complex, contested domains like medical controversies.

Methodological Approach Description

The thematic identification process proceeded in four structured and iterative phases, aiming to ensure methodological rigour, transparency, and conceptual validity, drawing upon established qualitative thematic analysis principles [125, 126]:

1. Multiple-Thematic Identification (Step 4a):

- A textual dataset was compiled to automate the theme identification and extraction from the corpus, comprising justification texts from the abstract classification phase in Steps 2b and 2c. Due to context window limitations inherent in the LLM models, the dataset was reduced to a random sample of 800 (out of 1,033) justification texts. This sample size was the maximum feasible for effective LLM processing within the technical constraints while providing a substantial and representative corpus for thematic exploration.
- Multiple, diverse and most powerful reasoning LLMs were used for theme identification. This ensured independence in theme identification and mitigated potential biases inherent in any single LLM. Three advanced reasoning LLMs, —GPT preview-o1, Gemini 2.0 Flash, and DeepSeek R1—were each separately tasked with identifying overarching themes within the dataset.
- Each LLM was instructed to identify overarching themes without being constrained to a predefined number of themes to extract. This allowed LLMs freedom to enable thematic structures to emerge organically from the data. LLMs were merely instructed to identify “half a dozen or more overarching themes” based on clustering semantic patterns and recurrent arguments within the text, mimicking a human researcher’s inductive thematic coding process.
- This process produced three sets of independent themes identified by each LLM, requiring reconciliation.

2. Thematic Consolidation (Step 4b):

- Remarkably, despite operating independently and with different underlying architectures, all three models converged on exactly eight overarching themes (Table 2), suggesting internal consistency in the underlying thematic structures of the Lyme disease discourse and reinforcing the robustness and potential validity of the results beyond any single model’s idiosyncrasies.

GPT preview-o1	Gemini 2.0	DeepSeek R1
Persistence vs. Resolution of Infection	Persistence vs. Resolution of Infection	Etiological Mechanisms: Persistent Infection vs. Post-Infectious Immune Responses
Diagnostic Uncertainty and Misdiagnosis	Diagnostic Uncertainty and Misdiagnosis	Diagnostic Complexity and Biomarker Development
Effectiveness of Antibiotic Therapy	Effectiveness of Antibiotic Therapy	Clinical Management Controversies
Role of Immune Dysregulation	Immune Dysregulation	Autoimmune Pathways and Residual Antigenic Debris
Psychological vs. Biological Basis	Neurocognitive and Neuropsychiatric Manifestations	Long-Term Outcomes and Symptom Heterogeneity
Subjectivity vs. Objectivity of Symptoms	Patient-Centered Experiences	Advocacy and Psychosocial Burden
Sociocultural and Ethical Factors	Sociocultural and Ethical Factors	Sociocultural and Institutional Influences
Mechanisms of Pathogen Persistence	Mechanisms of Pathogen Persistence	Bacterial Pathogenesis and Host Interactions

Table 2: Summary of themes independently derived from three LLMs and their mapping.

- To establish face validity, two co-authors, possessing subject-expertise in Lyme disease controversies and social science discourse analysis, independently reviewed the initial LLM-generated themes. This review phase aimed to ensure the themes resonated with existing knowledge of the Lyme debate and relevant social science concepts. The reviewers investigated conceptual overlaps and divergences across the three models, noting areas of agreement and disagreement in thematic identification. Consideration was given to comparing emergent themes against existing scholarly literature on medical controversies, contested illnesses, and the specific dynamics of the Lyme disease debate, ensuring alignment with established knowledge. Also, an evaluation was made as to whether themes were mutually exclusive, conceptually

distinct, and analytically useful for capturing the key dimensions of the Lyme disease discourse from a social science perspective.

- Next, the reconciliation process systematically mapped all significant conceptual domains identified by the initial LLM-generated themes. The results can also be seen in Table 2, which shows the alignment of all themes across the outputs of each LLM. The cross-model thematic reconciliation was also performed via an iterative inductive-deductive hybrid approach [125, 126], blending automated pattern recognition with expert human judgment. This consisted of two stages, where we first used the advanced GPT preview-o1 LLM to generate a consolidated thematic structure across the three models. This LLM-assisted step provided an initial synthesis, highlighting commonalities and potential redundancies for human review. Next, the LLM-reconciled output was manually examined and validated by the same two subject-expert co-authors. This involved discussions and critical evaluation to ensure conceptual clarity, eliminate redundancies, and, crucially, refine the themes to align with established discourse frameworks in the social sciences, particularly STS [98].
 - The final task in Step 4b was to develop a consolidated thematic framework, presented in Table 3, which defines each theme alongside its theoretical grounding in social science. This framework aimed to integrate insights from STS, illustrating how social, cultural, and political dynamics shape scientific knowledge production and interpretation, extending beyond biomedical perspectives. By embedding established social science perspectives on Lyme disease discourse, the framework sought to reflect the layered complexity of the controversy, its epistemological tensions, and its broader societal implications, ensuring alignment with existing literature on medical controversies and contested illnesses. To ensure methodological rigour and social science validity, the thematic classification process was explicitly grounded in discourse analysis, the sociology of medical knowledge, and medical controversy research [95, 127, 128]. This process systematically integrated LLM-based pattern recognition with expert qualitative interpretation, ensuring that biomedical and sociocultural narratives were meaningfully captured while maintaining conceptual coherence. The hybrid approach preserved the interpretability of computational classifications within broader social and epistemological contexts, ensuring that automated analyses were methodologically sound and aligned with expert human judgment in the study of contested medical knowledge.
3. **Abstract Thematic Labelling (Step 4c):** With the themes identified, the last classification task was conducted, to classify all 1,033 abstracts and their classification justifications into the derived themes, which could then aid in conducting a deeper analysis. Automation was once again used for this. GPT-o1-mini was tasked with assigning each abstract/justification pair two most suitable thematic categories. The methodological choice was made recognising the multidimensional nature of Lyme disease discourse and acknowledging the potential for abstracts to address multiple thematic dimensions simultaneously.
 4. **Expert Validation (Step 4d):** Finally, we selected a random sample of 50 abstract/justification pairs comprising 100 thematic classifications from Step 4c and asked subject experts to validate the assignment of themes that rounded off the last human-in-the-loop verification of the thematic classification before we proceeded to analysis. In this step, the human validators assessed the classifications for errors. Despite the inherent subjectivity of the task and interpretative overlaps, the human evaluators agreed with 96% of thematic assignments, thus confirming the validity of this process as a whole.

5 Results

Figure 5 visualizes the yearly distribution of stance classifications, confirming a notable increase in publication volume over time, particularly after 2014, with perceivable surges in 2015, 2019, and 2021. This trend, viewed through an STS lens, underscores the growing societal and academic relevance of the Lyme disease controversy, signalling its transformation into a more debated area within biomedical and public health discourse. The predominance of the *Neutral* stance, consistently the largest category, suggests general epistemic caution or perhaps strategic neutrality among researchers. Given the persistent diagnostic and therapeutic uncertainties or publication prospects of the studies, this may reflect a field-wide hesitancy to endorse polarised positions. While fewer studies explicitly endorse CLD, their consistent presence indicates a sustained, albeit marginalised, counter-narrative challenging mainstream PTLDS frameworks. The increase in abstracts supporting the PTLDS perspective, especially post-2010, reveals a dynamic evolution of the scientific dialogue, with a shifting centre of gravity towards the PTLDS framework, even as CLD-aligned discourse persists as a significant dissenting voice. Figure 5 thus highlights the entrenched polarisation and the evolving and contested nature of scientific knowledge production within the Lyme debate.

Figure 6 confirms these trends by presenting the overall stance distribution across the 25-year dataset. The *Neutral* stance indeed constitutes the largest proportion (42%), followed by *Supports PTLDS* (34%) and *Supports CLD* (24%).

Final Consolidated Theme	Description	Social Science Rationale/Relevance
Active Infection vs. Post-Infectious Immune Activity	Examines whether persistent symptoms are due to an ongoing <i>Borrelia burgdorferi</i> infection or a post-infectious immune response, a central debate in Lyme disease.	Epistemological Divide & Paradigm Clash: Reflects the fundamental epistemological divide in the Lyme controversy, highlighting competing paradigms of disease causation (biological vs. immunological). Connects to STS concepts of scientific controversy, paradigm clashes, and the social construction of scientific facts.
Diagnostic Complexity and Uncertainty	Investigates challenges in Lyme disease diagnosis, including limitations of serological testing, potential misdiagnosis, and the absence of definitive biomarkers.	Social Construction of Diagnosis & Medical Uncertainty: Illustrates the social construction of diagnostic categories, the inherent uncertainty in medical knowledge, and the limitations of biomedical reductionism in complex illnesses. Relevant to medical sociology's focus on the patients' experience of diagnostic ambiguity, patient navigation of complex medical systems, and the social impact of contested diagnoses.
Therapeutic Controversies and Antibiotic Efficacy	Addresses the contentious issue of prolonged antibiotic therapy, conflicting treatment guidelines, and the clinical efficacy of alternative interventions.	Bioethics & Medical Pluralism: Highlights the social and ethical dimensions of treatment decisions in contested illnesses, including the balance between evidence-based medicine, patient autonomy, and the influence of advocacy groups on treatment choices. Relates to bioethics, the sociology of medical practice, and the study of medical pluralism and treatment-seeking behaviours.
Immune Dysregulation and Autoimmune Mechanisms	Explores the hypothesis that post-treatment Lyme disease symptoms may stem from immune dysfunction, autoimmunity, or persistent inflammation rather than active infection.	Biomedical Framing & Shifting Paradigms: Represents a biomedical framing of persistent symptoms within established immunological paradigms, potentially reflecting a shift away from infection-centric models. Connects to STS analysis of how biomedical frameworks shape research agendas and the legitimization of certain types of medical knowledge over others.
Neurocognitive and Neuropsychiatric Manifestations	Focuses on Neurological and Psychiatric Sequelae Linked to Lyme Disease, such as cognitive impairment, neuroinflammation, and psychiatric symptoms.	Psychosocial Impact of Contested Illness & Stigma: Underscores the psychosocial impact of Lyme disease, including cognitive and mental health challenges, and how these are understood and contested within the controversy. Relevant to medical sociology and medical anthropology's interest in the patient experience of chronic illness, stigma, and the social construction of mental health in contested medical conditions.
Patient-centred Experiences and Advocacy	Highlights patient narratives, diagnostic challenges, disparities in medical recognition, and the broader psychosocial context of Lyme disease.	Patient Agency & Challenging Medical Authority: Centres on patient perspectives, highlighting patient agency in navigating contested diagnoses, challenging medical authority, and advocating for recognition and alternative treatments. Connects to patient advocacy studies, sociology of patient experience, and the role of online health communities in shaping health discourses.
Sociocultural and Ethical Factors	Examines the role of advocacy groups, public discourse, legal frameworks, and media representations in shaping Lyme disease perceptions and policies.	Social Construction of Health Policy & Media Influence: Explicitly addresses the broader sociocultural and ethical dimensions of the Lyme controversy, examining the influence of advocacy groups, media framing, and legal frameworks on shaping health policy and public understanding. Directly relevant to SSM's focus on the social determinants of health, health policy analysis, and media studies of health controversies.
Mechanisms of Pathogen Persistence and Biofilm Formation	Investigates microbial survival mechanisms, including persister cells and biofilms, which may contribute to treatment resistance and chronic symptoms.	Marginalised Biomedical Research & Alternative Paradigms: Represents a more marginalised biomedical research perspective, often associated with CLD advocacy, focusing on mechanisms that challenge mainstream views of pathogen eradication and treatment failure. Connects to STS analysis of scientific marginalisation, the sociology of scientific knowledge production in contested fields, and the dynamics of alternative medical paradigms.

Table 3: Key overarching themes extracted from the dataset of abstracts, their classifications, and their definition and grounding in the Science and Technology Studies framework.

Sociologically interpreted, this aggregate distribution reveals a key structural feature: a substantial portion of the literature strategically navigates the epistemic fault lines of the controversy without fully committing to either pole. However, the combined proportion of PTLDS and CLD-supporting studies (58%) underscores that the majority of research nonetheless remains deeply polarised, reflecting the enduring scientific and clinical divides within the Lyme disease landscape.

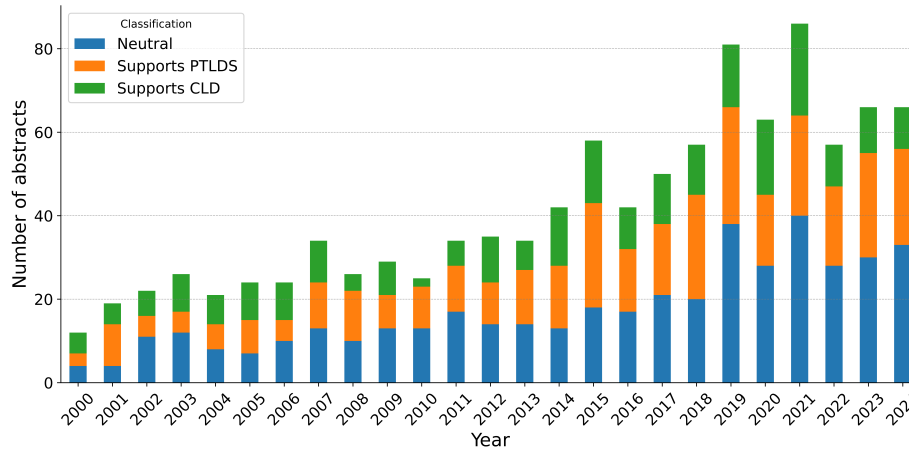


Figure 5: Number of relevant studies by year (2000-2024) and their classification. Classifications consisted of Neutral (blue), Supports PTLDS (orange), and Supports CLD (green).

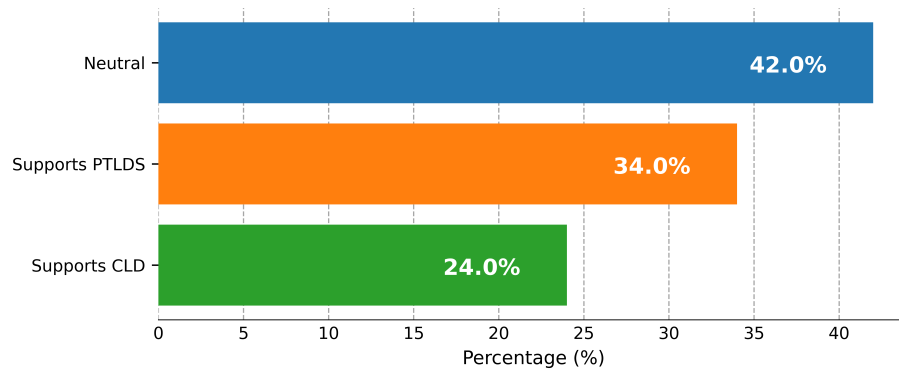


Figure 6: Percentage distribution of classifications. Neutral, Supports PTLDS, and Supports CLD are depicted in blue, orange and green, respectively.

Figure 7 provides a more granular and deconstructed view of stance distribution, explicitly showing yearly publication counts per classification and confirming earlier observations. However, Figure 7 reveals notably that Neutral and Supports PTLDS studies have driven this growth, while Supports CLD publications have remained comparatively flat, even declining recently. This differential growth, viewed through an STS lens, underscores a shifting centre of gravity in the discourse; whereas the academic conversation expands, increasingly it has centred on Neutral or PTLDS-aligned perspectives. The sustained prominence of the Neutral category across years, now visually explicit in Figure 7, further highlights strategic epistemic caution within Lyme disease research. Yearly data show Neutral studies consistently as the largest single category, often exceeding the combined count of polarised stances. This sustained neutrality, particularly amidst intense debate, likely reflects methodological conservatism where researchers may be prioritising cautious, evidence-based approaches, avoiding definitive stances given persistent diagnostic and therapeutic uncertainties.

Conversely, the relatively flat trajectory of Supports CLD publications, now clearly differentiated, underscores the persistent marginalization of this perspective. Despite overall research growth, CLD-advocating publications have not seen comparable increases, exhibiting a recent decline. This visual trend reinforces the interpretation of a field where CLD viewpoints, though consistently present, remain a minority, failing to gain broader traction in mainstream academic discourse. In contrast, Supports PTLDS publications show a pronounced upward trend with a distinct acceleration, especially post-2014, suggesting the institutionalisation of the PTLDS framework. This trend indicates a strengthening consensus around immune-mediated explanations and increasing alignment with mainstream medical guidelines.

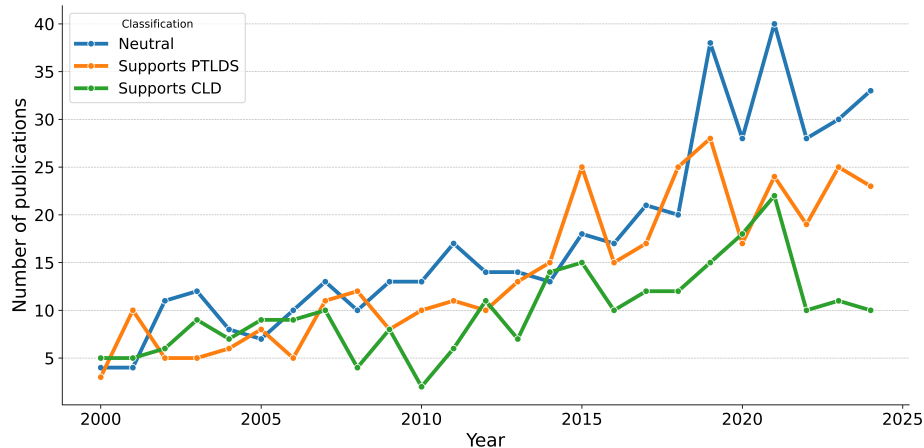


Figure 7: Study classifications on Lyme disease from 2000 to 2025. The yearly count of abstracts on Lyme disease from 2000 to 2025 was classified into three classifications: Neutral, Supports PTLDS, and Supports CLD, which are depicted as blue, orange, and green, respectively.

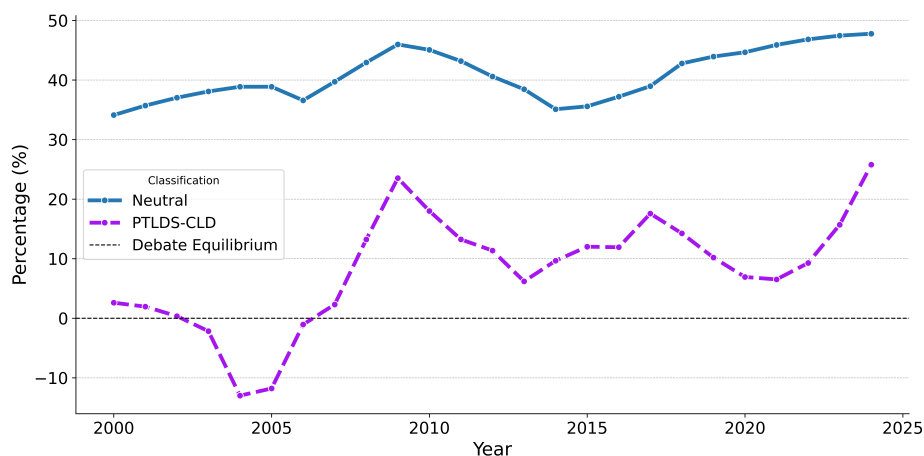


Figure 8: Smoothed trends in classification percentages by year using the Savitzky-Golay [129] filter. Lines represent smoothed trends for Neutral and PTLDS-CLD classifications over the period 2000 to 2025.

Figure 8, presenting smoothed trends⁵ in classification percentages, offers a longitudinal perspective on stance evolution. The solid line confirms the enduring prominence of neutrality as a defining feature of the discourse. Its upward trajectory, particularly in recent years, reinforces the interpretation of field-wide epistemic caution, reflecting researchers' strategic navigation of persistent uncertainties. The dashed line represents the smoothed difference in percentage points between the *Supports PTLDS* and *Supports CLD* classifications. Values above the zero threshold on the figure represent a greater predominance of PTLDS-supporting studies versus CLD and vice versa for negative values. The figure reveals a more dynamic, fluctuating pattern indicative of the shifting balance of power between competing paradigms. The negative dip in the early to mid-2000s, where CLD-supporting studies held a relative majority, suggests a historically contingent phase of alternative viewpoints gaining temporary traction, potentially fuelled by early patient activism challenging established biomedical narratives. However, the subsequent and sustained shift into positive territory, accelerating in recent years, once more empirically substantiates the institutionalisation of the PTLDS framework as the increasingly dominant paradigm. This longitudinal shift signifies a strengthening alignment with mainstream medical consensus and biomedical authority. While fluctuations persist within the PTLDS-CLD difference, the overall trend underscores the dynamic and continuously evolving nature of the Lyme disease controversy, reflected in shifts in research conclusions and publication volume. Figure 8 thus provides a macro-level view of these long-term shifts, complementing the

⁵We applied the Savitzky-Golay [129] filter kernel to reduce noise while preserving the underlying signal characteristics, using a second-order polynomial and a window size 10.

granular yearly data presented in Figure 7 and offering a broader temporal context for understanding the evolving stance landscape.

Shifting the focus to potentially latent journal-level biases, Figure 9 reveals a stratified epistemic landscape shaped by publication venue. The figure displays the difference in percentage points between *Supports PTLDS* and *Supports CLD* classifications across the top 20 journals by publication volume within our target dataset. Positive values indicate a higher proportion of PTLDS-supporting studies, while negative values suggest a more significant presence of CLD-supporting studies in the given journals' outputs. This journal-centric view, analysed through STS and Bourdieu's field theory [130] specifically, underscores how a journal's specialisation structures the Lyme controversy, channelling and differentiating the representation of competing knowledge claims. Leading infectious disease and clinical medicine journals—*Clinical Infectious Diseases*, *The Journal of Infectious Diseases*, and *The American Journal of Medicine*—predominantly publish PTLDS-supporting studies (positive values) with respect to our dataset. This sociologically interpreted proclivity reflects these field-defining venues reinforcing mainstream medical consensus and legitimising the PTLDS framework. Aligned with established clinical guidelines and institutional authorities like the IDSA, these journals can be seen as functioning as epistemic gatekeepers, preferentially disseminating research congruent with dominant biomedical narratives.

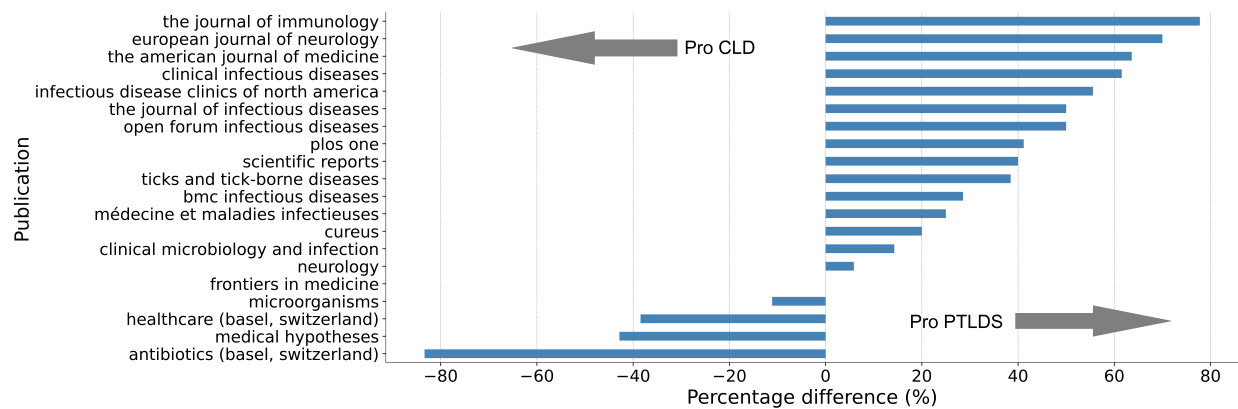


Figure 9: Top 20 journals by volume representing our dataset and depicting their potential bias. Values are the difference in percentage points between published papers classified as supporting PTLDS or CLD. Positive values indicate a higher proportion of PTLDS-supporting studies for a journal and vice-versa for CLD.

Conversely, journals with broader, hypothesis-driven scopes—*Medical Hypotheses* and *Antibiotics (Basel, Switzerland)*—exhibit a contrasting bias towards CLD-supporting studies (negative values). This suggests that CLD-aligned research, challenging mainstream paradigms, often finds outlets outside core infectious disease venues. While offering platforms for heterodox perspectives, these journals occupy a more marginalised position within the biomedical field, potentially limiting the broader impact of CLD-supporting research on clinical practice. The PTLDS preference in immunology and neurology-focused journals—*The Journal of Immunology* and *The European Journal of Neurology*—further highlights disciplinary influences on potential journal bias, reflecting a research emphasis on immune and neurological dysfunction. Meanwhile, the slightly more balanced stance in *Ticks and Tick-Borne Diseases* and *BMC Infectious Diseases* likely stems from their broader scope beyond Lyme-specific controversies. Figure 9 thus reveals a segmented epistemic landscape, where journal specialisation and editorial practices actively shape and reinforce distinct perspectives, contributing to the enduring polarisation of the Lyme disease controversy.

To further probe the epistemic influence of publications in Figure 9, we complement it with the analysis of citation impact data. Our analysis shows that the top 20 most-cited abstracts—a mere 2% of the dataset—account for 45% of all citations, thus highlighting a skewed distribution of epistemic power towards a small, highly influential subset of publications. This concentration, viewed through STS, underscores how a limited number of studies disproportionately shape the discourse's trajectory and impact. Analysing stance distribution within these top 20 papers reveals a clear bias: Supports PTLDS studies dominate (13 of 20), eclipsing Neutral (4) and Supports CLD (3) papers. Sociologically interpreted, this citation bias suggests a reinforcement of the PTLDS framework as the most impactful paradigm. The concentration of citations within PTLDS-supporting studies amplifies their visibility, credibility, and perceived scientific authority. This pattern also extends to the overall citation share across the whole dataset. While PTLDS-supporting papers constitute 34% of studies in our dataset, they accrue a disproportionate 52% of citations. Conversely, Neutral studies, the largest abstract proportion (42%), garner a smaller citation share (26%), and CLD-supporting papers, representing 24% of abstracts, receive the smallest share (22%). This discrepancy between publication volume and

citation impact, analysed through the prism of Bourdieu [130], reveals an epistemic hierarchy. While Neutral studies represent the largest research output, and CLD-supporting studies maintain a consistent presence, the PTLDS framework commands greater epistemic influence, attracting disproportionate scholarly attention and citations. Citation analysis thus further substantiates the institutionalisation of the PTLDS framework as the dominant paradigm, shaping not only publication trends but also the perceived impact and influence of Lyme disease research.

5.1 Thematic Analysis

To systematically assess the dominant conceptual dimensions in the Lyme disease debate, we categorised each abstract (see Step 4 in Figure 2) according to overarching themes (defined in Table 3) that capture distinct aspects of the controversy. Table 4 presents the thematic distribution of 1,033 classified papers, illustrating their absolute counts and proportional representation. Additionally, it highlights how each theme aligns with the broader debate by indicating the proportion of papers classified as *Neutral*, *Supports PTLDS*, or *Supports CLD*. This analysis provides a structured view of how different scientific perspectives are distributed across key areas of contention.

Theme	Papers	Percent	Neutral %	Supports PTLDS (%)	Supports CLD (%)
Active Infection vs. Post-Infectious Immune Activity	579	56.1	11.7	49.4	38.9
Diagnostic Complexity and Uncertainty	530	51.3	77.0	16.2	6.8
Therapeutic Controversies and Antibiotic Efficacy	365	35.3	20.8	40.5	38.6
Neurocognitive and Neuropsychiatric Manifestations	196	19.0	61.7	21.4	16.8
Immune Dysregulation and Autoimmune Mechanisms	192	18.6	28.1	62.5	9.4
Patient-Centered Experiences and Advocacy	149	14.4	82.6	8.1	9.4
Mechanisms of Pathogen Persistence and Biofilm Formation	30	2.9	10.0	0.0	90.0
Sociocultural and Ethical Factors	25	2.4	76.0	24.0	0.0

Table 4: Distribution of studies by theme and classifications supporting PTLDS/CLD or neutrality.

Among the themes in Table 4, *Active Infection vs. Post-Infectious Immune Activity* emerged as the dominant topic associated with more than half of the studies. This theme reflects the fundamental divide in Lyme disease research, where one faction attributes persistent symptoms to immune dysfunction (Supports PTLDS, 49.4%). At the same time, the other endorses the hypothesis of bacterial persistence (Supports CLD, 38.9%). The relatively low proportion of neutral papers (11.7%) compared to the overall percentage (42% refer to Figure 6) underscores the polarising nature of this theme, as most studies align explicitly with one of the two competing explanatory models.

Closely related to this core controversy is the *Diagnostic Complexity and Uncertainty* theme, which also appears in half of the target studies. In contrast to the previous theme, here there is an overwhelming neutrality of studies in this category (77%), highlighting the ongoing challenges in establishing clear diagnostic criteria over the last quarter of a century and is a problem that has been exacerbated by the limitations of serological tests and the absence of universally accepted biomarkers [57, 75]. Despite the overwhelming neutrality that has sustained the controversy amidst diagnostic ambiguity, a larger proportion of studies lean towards PTLDS (16%), while only 6.8% align with CLD, indicating that, while uncertainty dominates, the prevailing inclination, although modest, favours an immune-mediated rather than infection-driven explanation for persistent symptoms.

Discussions surrounding treatment strategies are equally contentious, as reflected in the *Therapeutic Controversies and Antibiotic Efficacy* theme, which accounts for around a third of the studies. The classification breakdown reveals a near-even split between PTLDS-supporting (40.5%) and CLD-supporting (38.6%) papers, with a significantly smaller than average distribution to a neutral stance, illustrating the ongoing debate over whether extended antibiotic regimens provide a therapeutic benefit or pose unnecessary risks. This division again mirrors the long-standing disagreements between major health organisations regarding treatment guidelines, further reinforcing the unresolved nature of this issue.

Beyond these core debates, several secondary themes provide insights into the broader implications of Lyme disease research and discourse. The theme of *Immune Dysregulation and Autoimmune Mechanisms*, which is identified in a fifth of studies, is predominantly associated with PTLDS (62.5%), suggesting a growing recognition of immune dysfunction as a plausible driver of persistent symptoms. Conversely, bacterial persistence receives significantly less support within this thematic category (9.4%), confirming that research into immune dysregulation tends to align more closely with the PTLDS framework than the CLD perspective.

A related but distinct theme, *Neurocognitive and Neuropsychiatric Manifestations*, also emerges in a fifth of the studies. A significant majority of these, however, remain neutral (61.7%), reflecting the ambiguity surrounding whether neurocognitive symptoms arise from residual infection, immune-mediated mechanisms, or other secondary effects. While PTLDS-supporting papers (21.4%) slightly outnumber CLD-supporting papers (16.8%), the relatively balanced distribution suggests that both explanatory models remain viable considerations in this domain in the presence of predominant uncertainty. Outside of the biomedical discourse, the role of patient experiences and advocacy remains a notable but underexplored dimension. The theme of *Patient-Centered Experiences and Advocacy* represents approximately an eighth of all studies characterised by overwhelming neutrality (82.6%). This high neutrality suggests that, while patient narratives are acknowledged, few studies explicitly frame them within either the PTLDS or CLD paradigm. Nonetheless, small proportions of studies equally support PTLDS (8.1%) or CLD (9.4%), reflecting the residual tensions between scientific discourse and patient advocacy efforts.

Although the theme of *Mechanisms of Pathogen Persistence and Biofilm Formation* is among the least represented (2.9% of studies), it stands out for its strong association with CLD (90.0%). This overwhelming alignment suggests that research on bacterial persistence may primarily be conducted within the CLD framework, reinforcing its position as the central scientific rationale for the chronic infection hypothesis. However, the limited number of papers in this category indicates that, despite its prominence in CLD discourse, empirical investigations into biofilms and persister cells remain relatively scarce. Lastly, *Sociocultural and Ethical Factors* constitute the least explored theme, appearing in only 2.4% of the studies. The high proportion of neutral papers (76.0%) suggests that while sociopolitical influences are acknowledged, they are seldom the primary focus of scientific inquiry in this controversy. However, within this small subset of studies, PTLDS receives more explicit support (24.0%) than CLD (0%), potentially reflecting how institutional and ethical discussions more frequently align with the mainstream medical consensus rather than the alternative chronic Lyme paradigm.

Table 5 recasts the above analysis into a temporal frame. The table, therefore, reveals temporal shifts in thematic focus across the past three decades. The thematic analysis from this table identifies shifts in the scientific focus of Lyme disease research and discourse over time, normalising the percentages to account for the incompleteness of the current decade. The most studied theme, *Active Infection vs. Post-Infectious Immune Activity*, has declined from 33% of papers in the 2000s to 24% in the 2020s, reflecting a broader transition from infection-centric to immune-mediated explanations. Meanwhile, *Diagnostic Complexity and Uncertainty* have grown (21% → 25% → 29%), pointing toward continued and unresolved challenges in establishing reliable biomarkers and diagnostic criteria. Similarly, *Therapeutic Controversies and Antibiotic Efficacy*, though declining in prevalence (20% → 20% → 14%), remains one of the most contested areas as discussed previously.

Theme	2000s (%)	2010s (%)	2020s (%)
Active Infection vs. Post-Infectious Immune Activity	33	28	24
Diagnostic Complexity and Uncertainty	21	25	29
Therapeutic Controversies and Antibiotic Efficacy	20	20	14
Neurocognitive and Neuropsychiatric Manifestations	11	8	11
Immune Dysregulation and Autoimmune Mechanisms	10	8	10
Patient-Centered Experiences and Advocacy	3	8	10
Sociocultural and Ethical Factors	1	1	2
Mechanisms of Pathogen Persistence and Biofilm Formation	1	2	1

Table 5: Shifts in the percentage of studies by thematic focus across decades.

These findings highlight distinct thematic patterns in Lyme disease research, reinforcing the entrenched divisions that define the controversy. The central debate over persistent infection versus immune-mediated pathology remains a dominant axis of discourse. Yet, its proportional representation has declined as research has increasingly shifted toward diagnostic complexity and uncertainty. Studies focusing on diagnostic complexity have steadily increased over the decades, now representing the most expanding area of research, underscoring the persistent difficulty in establishing definitive diagnostic criteria. Treatment strategies remain a significant point of contention, though research on antibiotic efficacy has proportionally declined, suggesting a decreased emphasis on therapeutic debates over time. While immune dysregulation and neurocognitive symptoms continue to be actively explored within the PTLDS framework, research on bacterial persistence and biofilm formation remains confined to a niche subset of CLD-focused studies, reflecting its limited expansion within mainstream biomedical discourse. The high neutrality observed in patient-centred and sociocultural discussions underscores the marginal integration of experiential perspectives into scientific debates despite a modest increase in advocacy-related research in recent years. These thematic distributions illustrate the scientific dimensions of the Lyme disease debate and its broader structural divisions where competing frameworks persist

across clinical, immunological, and sociopolitical contexts, shaping the discourse, research priorities and institutional positioning.

6 Discussion

Prior research has offered crucial insights into the contentious character of Lyme disease by examining patient advocacy, the influence of alternative medical viewpoints, and the effect of scientific ambiguity on public confidence. Nevertheless, these investigations have predominantly been qualitative, depending on discourse analysis, patient trajectory studies, and theoretical frameworks. This research used a hybrid and novel AI-driven technology combined with the human-in-the-middle oversight to identify attitudes and analyse themes in a large dataset spanning 25 years. This methodology enabled a structured examination of the discussion on Lyme disease over time and at scale, offering a quantifiable view of persistent debates that have exceeded prior research endeavours.

6.1 Evolution of the Academic Discourse on CLD and PTLDS

The evolution of Lyme disease research revealed a significant shift in stance distributions over time (**RQ1**). In the early 2000s, CLD-supporting studies were relatively more frequent, reflecting ongoing debates about the persistence of *Borrelia burgdorferi* post-antibiotic treatment. Yet, after 2010, PTLDS-supporting studies increasingly outnumbered CLD-supporting ones, a trend that has accelerated in recent years. This shift coincided with the consolidation of the PTLDS framework within major medical organisations and an increasing emphasis on immune-mediated rather than infection-driven explanations for persistent symptoms. Alongside this shift, thematic trends revealed a transition in research priorities. This trend underscored the prolonged difficulty in establishing reliable biomarkers and diagnostic criteria, which remains a key barrier to consensus in the field. Similarly, research on therapeutic controversies, including antibiotic efficacy, has proportionally declined (20% → 14%), suggesting a decreasing emphasis on long-term antibiotic treatment trials. This shift may also reflect clinical trial results discouraging prolonged antibiotic use and institutional funding trends prioritising alternative research directions. These findings indicate that while stance polarisation persists, the scientific focus has reoriented from direct treatment interventions toward diagnostic and mechanistic questions. This pattern aligns with more significant changes in discussions about Lyme disease.

This study's primary contribution in quantifying discourse shifts throughout time on this controversy using a large-scale dataset exceeds prior research efforts. Although, investigations by Pascal et al. [49], Puppo et al. [84], Hinds and Sutcliffe [85] examined the polarisation and impact of conflicting medical viewpoints on Lyme disease discourse, frequently highlighting the significance of social movements, patient advocacy, and media influence, these studies have not incorporated a systematic chronological analysis. Our confirmation that since 2010, there has been a significant rise in research endorsing PTLDS perspectives, especially in high-impact journals accompanied by disproportionate citation volumes and impact, whilst studies supporting CLD have diminished, has provided empirical evidence of a shift in scientific consensus, which was formerly assumed from limited qualitative evaluations.

6.2 Influence of Journal Specialisation and Editorial Focus

Our analysis also supports the notion that journal specialisations shape the discourse on Lyme disease (**RQ2**), where we find that leading infectious disease and clinical medicine journals, predominantly publish PTLDS-supporting studies. This reflects alignment with mainstream clinical guidelines, particularly those endorsed by the IDSA, which maintains that prolonged antibiotic therapy for CLD lacks sufficient evidence. Uzzell et al. [83] and Bloor et al. [90] illustrated that journalistic framing frequently favours established medical institutions, reinforcing mainstream opinions and marginalising alternative viewpoints. In contrast, CLD-supporting studies are more frequently published in journals with broader or hypothesis-driven scopes, underscoring the claim that research favouring bacterial persistence and extended antibiotic treatment often finds publication outlets outside high-impact infectious disease journals, potentially limiting its influence on clinical practice. The strong preference for PTLDS-supporting studies in neurology and immunology-focused journals further indicates that PTLDS is increasingly framed as an immunological or neurological disorder rather than an active bacterial infection. These journal-level biases may contribute to a self-reinforcing effect, wherein PTLDS-supporting research is more likely to appear in high-impact venues, increasing its credibility and citation impact. At the same time, CLD-supporting studies remain within a more restricted epistemic niche. Future investigations could explore whether this distribution reflects methodological quality differences or broader editorial and peer-review preferences that exclude CLD perspectives from mainstream discourse.

Despite the increasing acknowledgement of patient-centred perspectives in Lyme disease discourse, their representation in high-impact journals remains limited. Studies by Rebman et al. [13], Baarsma et al. [86] and Singer [88] highlight the medical scepticism and social stigma experienced by patients diagnosed with persistent Lyme disease symptoms,

leading many to seek alternative healthcare solutions. While prior research has documented these qualitative aspects, our study quantifies the growing scholarly emphasis on patient experiences, particularly since the late 2010s. However, this research remains descriptive mainly, with limited integration into clinical or mechanistic frameworks.

6.3 Thematic Structures and Their Scientific and Institutional Implications

Our thematic analysis provided novel and deeper insights into framing Lyme disease debates (**RQ3**) across time. While the infection persistence vs. immune-mediated pathology controversy has remained central, its proportional representation has declined (33% → 24%) over the decades, suggesting a shift in research emphasis. In contrast, studies on diagnostic complexity and uncertainty have increased (21% → 29%), highlighting ongoing methodological challenges in standardising Lyme disease diagnostics. Therapeutic debates remain polarised but have decreased in prominence (20% → 14%), reflecting a redirection in focus toward immune and neurological complications rather than antimicrobial persistence. Research on bacterial persistence and biofilm formation remains a niche field, with over 90% of studies in this category aligning with CLD perspectives. However, this area has not expanded significantly within mainstream Lyme disease research.

Beyond biomedical discourse, research on patient-centred experiences and advocacy has grown (3% → 10%), reflecting increasing acknowledgment of patient narratives. However, high neutrality within this category (82.6%) suggests that patient experiences are often studied descriptively rather than integrated into clinical or scientific frameworks. Similarly, sociocultural and ethical discussions have remained a small but expanding domain, underscoring the broader societal implications of the Lyme disease controversy. Our findings align with prior studies by Pascal et al. [49] and Dumes [87], which examined the role of media in shaping public and scientific perceptions of Lyme disease. While these studies analysed these trends qualitatively, our research quantifies the increasing prominence of patient-centred research relative to infection-focused studies, providing empirical validation for the reciprocal influence between scientific discourse and public narratives.

6.4 Implications of the Hybrid AI-Computational and Human-in-the-loop Text Analysis Methodology

This study's hybrid computational-interpretive methodology offers a scalable yet theoretically rigorous approach to systematically analysing large-scale textual datasets in social science research and beyond. This study demonstrated its value for studying complex and contested medical conditions. Combined with structured human oversight and validation grounded in established social science frameworks, it has demonstrated how AI-assisted classification and content analysis can be leveraged without displacing critical human interpretation or compromising reliability. The proposed methodological approach ensured the robustness and validity of replicable results, with the approach transferable to similar domains.

6.5 Broader Implications and Future Directions

Beyond specific stance distributions, this research highlights structural and epistemic challenges in Lyme disease. Key considerations include:

- The continued diagnostic uncertainty signals a need for better biomarkers and molecular tools to bridge discordant perspectives.
- Reduced emphasis on infection persistence and therapeutic controversies suggests shifting funding and research priorities toward immune-mediated explanations and diagnostic improvements.
- The incremental attention to patient advocacy underscores the importance of inclusive research integrating qualitative and patient-reported data.
- The limited expansion of bacterial persistence research indicates remaining contention over empirical evidence, funding, or shifting clinical frameworks.
- Institutional factors and clinical guidelines shape research agendas. Additional examination of policy, medical guidelines, and funding mechanisms will elucidate how shifts in public health priorities affect the landscape of Lyme disease research.

Our study aligns with agenda-setting [131], diffusion of innovations [132], and framing theory [100] to illustrate how institutional gatekeepers, clinical norms, and framing tactics shape the discourse, and how scientific disagreements are influenced by communication and institutional factors, offering a new viewpoint on the underlying reasons for ongoing Lyme disease challenges by applying these theories. This research also underscores the interplay between scientific authority, patient autonomy, and public health communication for disputed medical conditions [89, 90].

Overall, our large-scale, AI-assisted analysis deepens and extends prior qualitative findings, emphasising the sociopolitical factors that render Lyme disease a complex, evolving controversy. These insights reinforce the need for policy interventions supporting equitable research funding, patient-centred care, and improved diagnostic clarity. Future research should bridge biomedical consensus with authentic patient experiences, employing AI-driven methodologies to detect emerging patterns and improve health communication strategies. The persistence of unresolved debates around Lyme disease calls for a balanced approach integrating scientific rigour with empathetic, patient-focused practice.

6.6 Limitations

Several limitations must be acknowledged. First, the paper classification relied on abstracts, which may not fully capture nuanced positions or methodological details that may have been more clearly expressed in the full texts of the studies. The dataset was limited to indexed, peer-reviewed literature, excluding unpublished and non-English studies, potentially biasing stance distributions. The IRR findings highlighted some degree of variability in reasoning and outputs between the LLMs and human experts. It is possible that ensembling models' outputs and aggregating their classifications instead of relying on only one LLM could have enhanced classification reliability further. Journal-level biases may suggest editorial influence on discourse, but this study does not account for factors such as funding sources or peer review dynamics that may shape publication trends. Thematic categorisation, while informative, may oversimplify overlapping topics and emerging interdisciplinary perspectives - some abstracts may have spanned more than two themes. Additionally, stance shifts over 25 years may reflect evolving clinical definitions rather than changes in underlying scientific evidence. This study has mapped the structure of Lyme disease research but does not assess the empirical validity of competing claims.

7 Conclusion

Over the past quarter-century, the academic discourse surrounding Chronic Lyme Disease (CLD) and Post-Treatment Lyme Disease Syndrome (PTLDS) has become a deeply entrenched and complex controversy, fuelled by scientific disputes, patient activism, and pervasive media narratives. This study offers an unprecedented, large-scale quantification of this epistemic landscape through a rigorous and innovative hybrid AI-driven methodology. Our research transcended the limitations of traditional qualitative approaches for content analysis by analysing thousands of scholarly studies spanning this period, providing a temporally sensitive and granular understanding of the evolving dynamics of this enduring medical debate. The core novelty lies in our synergistic combination of cutting-edge large language models with structured human validation, enabling a level of systematic and scalable discourse analysis previously unattainable.

With our novel hybrid approach, our analysis revealed a significant and consequential shift in the academic conversation on this controversy, indicating that a clear transition from infection-centric models of Lyme disease towards immune-mediated explanations for persistent symptoms has occurred. This trend is particularly pronounced within high-impact clinical and immunology journals, signalling a growing institutional endorsement of the PTLDS perspective. Conversely, research supporting CLD, while consistently present, has remained largely relegated to hypothesis-driven publications, highlighting a persistent epistemic asymmetry within mainstream biomedical discourse. This divergence powerfully underscores the impact of journal specialisation and potential editorial biases in shaping competing scientific perspectives' dissemination and perceived validity. Furthermore, our AI-driven thematic analysis showed the enduring centrality of diagnostic complexity and therapeutic controversies, even as the dominant explanatory models and research priorities undergo significant transformation, demonstrating the robustness of our methodology in capturing nuanced thematic evolutions.

The key contribution of this study has been twofold: first, the provision of empirically grounded, large-scale insights into the structural and epistemic dynamics of the Lyme disease controversy, and second, the establishment of a robust and replicable hybrid AI-driven methodology for analysing complex medical debates. This innovative approach, demonstrably effective in dissecting the Lyme disease discourse, offers a powerful new tool for social science research in other fields. Our findings have direct implications for policymakers, clinicians, and communication strategists. For policymakers, they highlight the urgent need for equitable research funding policies that actively promote diverse scientific inquiry and challenge epistemic marginalisation. The persistent diagnostic ambiguity, revealed with unprecedented clarity through our AI-driven analysis, necessitates a renewed commitment to establishing standardised diagnostic criteria. For clinicians, our study reinforces the critical importance of adopting patient-centred care models that validate lived experiences, irrespective of diagnostic uncertainties. For communication strategists, we underscore nuanced and empathetic communication strategies are pivotal in bridging the persistent divide between scientific consensus and patient narratives.

References

- [1] Kiersten J Kugeler, Amy M Schwartz, Mark J Delorey, Paul S Mead, and Alison F Hinckley. Estimating the frequency of lyme disease diagnoses, united states, 2010–2018. *Emerging Infectious Diseases*, 27(2):616, 2021.
- [2] Robert A Sykes and Phoebe Makiello. An estimate of lyme borreliosis incidence in western europe. *Journal of public health*, 39(1):74–81, 2017.
- [3] Nancy A Shadick, Charlotte B Phillips, Eric L Logigian, Allen C Steere, Richard F Kaplan, Victor P Berardi, Paul H Duray, Martin G Larson, Elizabeth A Wright, Katherine S Ginsburg*, et al. The long-term clinical outcomes of lyme disease: a population-based retrospective cohort study. *Annals of Internal Medicine*, 121(8): 560–567, 1994.
- [4] Elizabeth L Maloney. Controversies in persistent (chronic) lyme disease. *Journal of infusion nursing*, 39(6): 369–375, 2016.
- [5] Alison Tonks. Lyme wars. *BMJ*, 335(7626):910–912, 2007.
- [6] Leonard H Sigal. Lyme disease controversy: Use and misuse of language. *Annals of Internal Medicine*, 137(9): 776–777, 2002.
- [7] BJ Kullberg, A Berende, and JW Van der Meer. The challenge of lyme disease: tired of the lyme wars. *Neth J Med*, 69(3):98–100, 2011.
- [8] Patrick Peretti-Watel, Jeremy Ward, Romain Lutaud, and Valérie Seror. Lyme disease: Insight from social sciences. *Médecine et Maladies Infectieuses*, 49(2):133–139, 2019.
- [9] Guillaume Coiffier and Pierre Tattevin. Lyme disease: “end of the debate?”. *Joint Bone Spine*, 88(4):105181, 2021.
- [10] Gary P Wormser, Raymond J Dattwyler, Eugene D Shapiro, John J Halperin, Allen C Steere, Mark S Klempner, Peter J Krause, Johan S Bakken, Franc Strle, Gerold Stanek, et al. The clinical assessment, treatment, and prevention of lyme disease, human granulocytic anaplasmosis, and babesiosis: clinical practice guidelines by the infectious diseases society of america. *Clinical Infectious Diseases*, 43(9):1089–1134, 2006.
- [11] Adriana Marques. Chronic lyme disease: a review. *Infectious Disease Clinics of North America*, 22(2):341–360, June 2008. doi:10.1016/j.idc.2007.12.011. PMID: 18452806; PMCID: PMC2430045.
- [12] John N. Aucott, Alison W. Rebman, Lisa A. Crowder, and Kathleen B. Kortte. Post-treatment lyme disease syndrome symptomatology and the impact on life functioning: is there something here? *Quality of Life Research*, 22(1):75–84, February 2013. doi:10.1007/s11136-012-0126-6. PMID: 22294245; PMCID: PMC3548099.
- [13] Alison W. Rebman, Katie T. Bechtold, Ting Yang, Emily A. Mihm, Mark J. Soloski, Charles B. Novak, and John N. Aucott. The clinical, symptom, and quality-of-life characterization of a well-defined group of patients with posttreatment lyme disease syndrome. *Frontiers in Medicine*, 4:224, December 2017. doi:10.3389/fmed.2017.00224. PMID: 29312942; PMCID: PMC5735370.
- [14] Jop Ursinus, Hendrik D. Vrijmoeth, Mark G. Harms, Joke A. D. Tulen, Hans Knoop, Sander A. Gauw, Tim P. Zomer, Audrey Wong, Ingrid H. M. Friesema, Yvonne M. Vermeeren, Leo A. B. Joosten, Joppe W. Hovius, Bart Jan Kullberg, and Carolien C. van den Wijngaard. Prevalence of persistent symptoms after treatment for lyme borreliosis: A prospective observational cohort study. *The Lancet Regional Health – Europe*, 6:100142, May 2021. doi:10.1016/j.lanep.2021.100142. Erratum in: *The Lancet Regional Health – Europe*. 2023 Mar 15;27:100622. doi: 10.1016/j.lanep.2023.100622. PMID: 34557833; PMCID: PMC8454881.
- [15] Alison W Rebman, John N Aucott, Eric R Weinstein, Kathleen T Bechtold, Katherine C Smith, and Lori Leonard. Living in limbo: contested narratives of patients with chronic symptoms following lyme disease. *Qualitative Health Research*, 27(4):534–546, 2017.
- [16] Stephen Mac, Simran Bahia, Frances Simbulan, Eleanor M Pullenayegum, Gerald A Evans, Samir N Patel, and Beate Sander. Long-term sequelae and health-related quality of life associated with lyme disease: a systematic review. *Clinical Infectious Diseases*, 71(2):440–452, 2020.
- [17] Robert F. Kaplan, Robert P. Trevino, Gary M. Johnson, Leonard Levy, Robert Dornbush, Linden T. Hu, John Evans, Alan Weinstein, Christopher H. Schmid, and Mark S. Klempner. Cognitive function in post-treatment lyme disease: do additional antibiotics help? *Neurology*, 60(12):1916–1922, June 2003. doi:10.1212/01.wnl.0000068030.26992.25. PMID: 12821733.
- [18] Alison W. Rebman, Ting Yang, and John N. Aucott. Symptom heterogeneity and patient subgroup classification among us patients with post-treatment lyme disease: an observational study. *BMJ Open*, 11(1):e040399, January 2021. doi:10.1136/bmjopen-2020-040399. PMID: 33441353; PMCID: PMC7812114.

- [19] Steven Shor, Christine Green, Beatrice Szantyr, Steven Phillips, Kenneth Liegner, Joseph J. Jr Burrascano, Robert Bransfield, and Elizabeth L. Maloney. Chronic lyme disease: An evidence-based definition by the ilads working group. *Antibiotics*, 8(4):269, December 2019. doi:10.3390/antibiotics8040269. PMID: 31888310; PMCID: PMC6963229.
- [20] Paul M Lantos. Chronic lyme disease: the controversies and the science. *Expert review of anti-infective therapy*, 9(7):787–797, 2011.
- [21] Hanna Schmid and Ulrich Heininger. Posttreatment lyme disease syndrome—what it might be and what it is not. *The Pediatric Infectious Disease Journal*, 40(5S):S31–S34, 2021.
- [22] Allen C Steere, Robert E Levin, Philip J Molloy, Robert A Kalish, James H Abraham Iii, Nancy Y Liu, and Christopher H Schmid. Treatment of lyme arthritis. *Arthritis & Rheumatism: Official Journal of the American College of Rheumatology*, 37(6):878–888, 1994.
- [23] Peter Conrad and Kristin K Barker. The social construction of illness: Key insights and policy implications. *Journal of health and social behavior*, 51(1_suppl):S67–S79, 2010.
- [24] Robert A Aronowitz. When do symptoms become a disease? *Annals of internal medicine*, 134(9_Part_2): 803–808, 2001.
- [25] Joseph Dumit. Illnesses you have to fight to get: Facts as forces in uncertain, emergent illnesses. *Social science & medicine*, 62(3):577–590, 2006.
- [26] Kevin C Elliott. The value-ladenness of transparency in science: lessons from lyme disease. *Studies in History and Philosophy of Science*, 88:1–9, 2021.
- [27] Paul G. Auwaerter, Johan S. Bakken, Raymond J. Dattwyler, J. Stephen Dumler, John J. Halperin, Edward McSweegan, Robert B. Nadelman, Susan O’Connell, Eugene D. Shapiro, Sunil K. Sood, Allen C. Steere, and Alan Weinstein. Antiscience and ethical concerns associated with advocacy of lyme disease. *The Lancet Infectious Diseases*, 11(9):713–719, September 2011. doi:10.1016/S1473-3099(11)70132-2. PMID: 21867956.
- [28] Henry M Feder Jr, Barbara JB Johnson, Susan O’Connell, Eugene D Shapiro, Allen C Steere, Gary P Wormser, and Ad Hoc International Lyme Disease Group. A critical appraisal of “chronic lyme disease”. *New England Journal of Medicine*, 357(14):1422–1430, 2007.
- [29] Lorraine Johnson, Spencer Wilcox, Jennifer Mankoff, and Raphael B Stricker. Severity of chronic lyme disease compared to other chronic conditions: a quality of life survey. *PeerJ*, 2:e322, 2014.
- [30] Paul G Auwaerter, Johan S Bakken, Raymond J Dattwyler, J Stephen Dumler, John J Halperin, Edward McSweegan, Robert B Nadelman, Susan O’Connell, Sunil K Sood, Arthur Weinstein, et al. Scientific evidence and best patient care practices should guide the ethics of lyme disease activism. *Journal of medical ethics*, 37(2): 68–73, 2011.
- [31] Daniel J Cameron, Lorraine B Johnson, and Elizabeth L Maloney. Evidence assessments and guideline recommendations in lyme disease: the clinical management of known tick bites, erythema migrans rashes and persistent disease. *Expert review of anti-infective therapy*, 12(9):1103–1135, 2014.
- [32] N Yiannakoulias, R Tooby, and SL Sturrock. Celebrity over science? an analysis of lyme disease video content on youtube. *Social Science & Medicine*, 191:57–60, 2017.
- [33] Jennifer Arney. Chronic illness, contested diagnoses, and the medical merry-go-round: A gendered analysis of the experiences of female patients with chronic lyme disease. *Social Science & Medicine*, 284:114240, 2021. doi:10.1016/j.socscimed.2021.114240.
- [34] Sunil K. Sood. Controversies in lyme disease: Perspectives from the clinic and the media. *Expert Review of Anti-infective Therapy*, 9(2):155–158, 2011. doi:10.1586/eri.10.159.
- [35] Annemarie Jutel. *Putting a name to it: Diagnosis in contemporary society*. JHU Press, 2024.
- [36] Adele E Clarke and Janet Shim. Medicalization and biomedicalization revisited: technoscience and transformations of health, illness and american medicine. In *Handbook of the sociology of health, illness, and healing: A blueprint for the 21st century*, pages 173–199. Springer, 2010.
- [37] Steven Epstein. *Impure science: AIDS, activism, and the politics of knowledge*, volume 7. Univ of California Press, 1996.
- [38] Vololona Rabeharisoa and Michel Callon. Patients and scientists in french muscular dystrophy research. In *States of knowledge*, pages 142–160. Routledge, 2004.
- [39] Eugene D Shapiro, Phillip J Baker, and Gary P Wormser. False and misleading information about lyme disease. *The American Journal of Medicine*, 130(7):771–772, 2017.

- [40] Sue Archer and Kelly Edwards. Online patient activism and the transformation of public engagement in contested illnesses: A case study of chronic lyme disease. *Information, Communication & Society*, 21(12):1706–1720, 2018. doi:10.1080/1369118X.2017.1348535.
- [41] Chaoqun Zeng, Yi Tang, Dena Rahmani, Stephen M. Croucher, and Leah Gilbert. The sculpture of tick-borne disease media coverage in the united states and china. *Journal of Intercultural Communication*, (48), November 2018. URL <https://immi.se/intercultural/nr48/zeng.html>.
- [42] Eugene D. Shapiro. Lyme disease in 2014: Social media changing how we think, talk, and act. *Clinical Infectious Diseases*, 58(8):1105–1106, 2014. doi:10.1093/cid/ciu062.
- [43] Kristin K Barker. Listening to lyrical: contested illnesses and pharmaceutical determinism. *Social science & medicine*, 73(6):833–842, 2011.
- [44] Stuart Blume. *Immunization: how vaccines became controversial*. Reaktion Books, 2017.
- [45] DJ Hess. *Undone science: Social movements, mobilized publics, and industrial transitions*. MIT Press, 2016.
- [46] Michel Callon, Pierre Lascoumes, and Yannick Barthe. *Acting in an uncertain world: An essay on technical democracy*. MIT press, 2011.
- [47] Adam Briggles and Carl Mitcham. *Ethics and science: An introduction*. Cambridge University Press, 2012.
- [48] Tommaso Venturini. Building on faults: How to represent controversies with digital methods. *Public understanding of science*, 21(7):796–812, 2012.
- [49] Clélia Pascal, Jocelyne Arquembourg, Philippe Vorilhon, and Olivier Lesens. Emergence of lyme disease as a social problem: analysis of discourse using the media content. *European Journal of Public Health*, 30(3): 475–481, 2020.
- [50] Raphael B Stricker and Lorraine Johnson. Chronic lyme disease and the ‘axis of evil’. *Future Microbiology*, 3(6):621–624, 2008.
- [51] Mark David Walker. The portrayal of lyme disease by a public service broadcaster. *Journal of Communication in Healthcare*, 14(4):303–311, 2021.
- [52] Allen C Steere, Stephen E Malawista, David R Snyderman, Robert E Shope, Warren A Andiman, Martin R Ross, and Francis M Steele. An epidemic of oligoarticular arthritis in children and adults in three connecticut communities. *Arthritis & Rheumatism: Official Journal of the American College of Rheumatology*, 20(1):7–17, 1977.
- [53] Allen C Steere, Jenifer Coburn, Lisa Glickstein, et al. The emergence of lyme disease. *The Journal of clinical investigation*, 113(8):1093–1101, 2004.
- [54] Justin D Radolf, Klemen Strle, Jacob E Lemieux, and Franc Strle. Lyme disease in humans. *Current issues in molecular biology*, 42(1):333–384, 2021.
- [55] John N Aucott, Lauren A Crowder, and Kathleen B Kortte. Development of a foundation for a case definition of post-treatment lyme disease syndrome. *International Journal of Infectious Diseases*, 17(6):e443–e449, 2013.
- [56] G Stanek, V Fingerle, K-P Hunfeld, B Jaulhac, R Kaiser, A Krause, W Kristoferitsch, S O’connell, Katharina Ornstein, F Strle, et al. Lyme borreliosis: clinical case definitions for diagnosis and management in europe. *Clinical Microbiology and Infection*, 17(1):69–79, 2011.
- [57] Ginny Brunton, Katy Sutcliffe, Kate Hinds, Meena Khatwa, Helen Burchett, Kelly Dickson, Rebecca Rees, Antonio Rojas-Garcia, Gillian Stokes, Melissa Harden, et al. Stakeholder experiences of the diagnosis of lyme disease: A systematic review. 2017.
- [58] Jason R Bobe, Brandon L Jutras, Elizabeth J Horn, Monica E Embers, Allison Bailey, Robert L Moritz, Ying Zhang, Mark J Soloski, Richard S Ostfeld, Richard T Marconi, et al. Recent progress in lyme disease and remaining challenges. *Frontiers in medicine*, 8:666554, 2021.
- [59] Marie Claire Van Hout. The controversies, challenges and complexities of lyme disease: implications for medical education, clinical practice and research. *Journal of Pharmacy and Pharmaceutical Sciences*, 21(1):429–436, 2018.
- [60] Arnstein Finset, Hayden Bosworth, Phyllis Butow, Pål Gulbrandsen, Robert L Hulsman, Arwen H Pieterse, Richard Street, Robin Tschoetschel, and Julia van Weert. Effective health communication—a key factor in fighting the covid-19 pandemic. *Patient education and counseling*, 103(5):873, 2020.
- [61] Adele E Clarke, Janet K Shim, Laura Mamo, Jennifer Ruth Fosket, and Jennifer R Fishman. Biomedicalization: Technoscientific transformations of health, illness, and us biomedicine. *American sociological review*, 68(2): 161–194, 2003.

- [62] Stefan Timmermans and Marc Berg. *The gold standard: the challenge of evidence-based medicine*. Temple University Press, 2010.
- [63] Dominique Brossard and Dietram A Scheufele. Science, new media, and the public. *science*, 339(6115):40–41, 2013.
- [64] Dale E Brashers. Communication and uncertainty management. *Journal of communication*, 51(3):477–497, 2001.
- [65] Eugene D Shapiro and Gary P Wormser. Lyme disease in 2018: what is new (and what is not). *Jama*, 320(7):635–636, 2018.
- [66] Paul M Lantos. Chronic lyme disease. *Infectious Disease Clinics*, 29(2):325–340, 2015.
- [67] Alison W Rebman and John N Aucott. Post-treatment lyme disease as a model for persistent symptoms in lyme disease. *Frontiers in medicine*, 7:524338, 2020.
- [68] Rick Dersch, Gabriel Torbahn, and Sebastian Rauer. Treatment of post-treatment lyme disease symptoms—a systematic review. *European Journal of Neurology*, page e16293, 2024.
- [69] Prat Sébastien, Dalbin Jacques, Plotton Catherine, and Gocko Xavier. Diagnosis and treatment of “chronic lyme”: primum non nocere. *BMC Infectious Diseases*, 23(1):642, 2023.
- [70] T Joseph Mattingly and Kalpana Shere-Wolfe. Clinical and economic outcomes evaluated in lyme disease: a systematic review. *Parasites & Vectors*, 13:1–8, 2020.
- [71] Marie Claire Van Hout. The controversies, challenges and complexities of lyme disease: A narrative review. *Journal of Pharmacy and Pharmaceutical Sciences*, 21(1):429–436, November 2018. doi:10.18433/jpps30254. URL <https://journals.library.ualberta.ca/jpps/index.php/JPPS/article/view/30254>.
- [72] Liesbeth Borgermans, Geert Goderis, Jan Vandevoorde, and Dirk Devroey. Relevance of chronic lyme disease to family medicine as a complex multidimensional chronic disease construct: a systematic review. *International Journal of Family Medicine*, 2014(1):138016, 2014.
- [73] Vladimir V Bamm, Jordan T Ko, Iain L Mainprize, Victoria P Sanderson, and Melanie KB Wills. Lyme disease frontiers: reconciling borrelia biology and clinical conundrums. *Pathogens*, 8(4):299, 2019.
- [74] Marissa Brackett, Jacklyn Potts, Allison Mehofer, Yumna Indorewala, Alina Ali, Sarah Lutes, Emma Putnam, Sophie Schuelke, Aisha Abdool, Emma Woldenberg, et al. Neuropsychiatric manifestations and cognitive decline in patients with long-standing lyme disease: A scoping review. *Cureus*, 16(4), 2024.
- [75] Maria E Aguero-Rosenfeld and Gary P Wormser. Lyme disease: diagnostic issues and controversies. *Expert Review of Molecular Diagnostics*, 15(1):1–4, 2015.
- [76] Adriana R Marques. Laboratory diagnosis of lyme disease: advances and challenges. *Infectious Disease Clinics of North America*, 29(2):295–307, 2015.
- [77] MH Beaman. Lyme disease: why the controversy? *Internal Medicine Journal*, 46(12):1370–1375, 2016.
- [78] Robert T Schoen. Challenges in the diagnosis and treatment of lyme disease. *Current Rheumatology Reports*, 22(1):3, 2020.
- [79] Brandee L Stone, Yvonne Tourand, and Catherine A Brissette. Brave new worlds: the expanding universe of lyme disease. *Vector-borne and zoonotic diseases*, 17(9):619–629, 2017.
- [80] Yan Dong, Guozhong Zhou, Wenjing Cao, Xin Xu, Yu Zhang, Zhenhua Ji, Jiaru Yang, Jingjing Chen, Meixiao Liu, Yuxin Fan, et al. Global seroprevalence and sociodemographic characteristics of borrelia burgdorferi sensu lato in human populations: a systematic review and meta-analysis. *BMJ Global Health*, 7(6):e007744, 2022.
- [81] Stephen Mac, Sara R da Silva, and Beate Sander. The economic burden of lyme disease and the cost-effectiveness of lyme disease interventions: A scoping review. *PLoS One*, 14(1):e0210280, 2019.
- [82] Casey Olechnowicz, Jessica Leahy, Allison Gardner, and Carly C Sponarski. Perceived vulnerability for lyme disease questionnaire: A social science tool for understanding tick-borne disease attitudes. *Ticks and Tick-borne Diseases*, 14(2):102120, 2023.
- [83] David Uzzell, Konstantina Vasileiou, Afrodita Marcu, and Julie Barnett. Whose lyme is it anyway? subject positions and the construction of responsibility for managing the health risks from lyme disease. *Health & Place*, 18(5):1101–1109, 2012.
- [84] C Puppo, Y Hansmann, L Moinot, X Duval, C Chirouze, and M Préau. The social representations of diagnosing lyme disease. *Plos one*, 18(2):e0276800, 2023.

- [85] Kate Hinds and Katy Sutcliffe. Heterodox and orthodox discourses in the case of lyme disease: a synthesis of arguments. *Qualitative Health Research*, 29(11):1661–1673, 2019.
- [86] ME Baarsma, SA Claassen, HE Van Der Horst, JW Hovius, and JM Sanders. Knowing the entire story—a focus group study on patient experiences with chronic lyme-associated symptoms (chronic lyme disease). *BMC Primary Care*, 23(1):139, 2022.
- [87] Abigail A Dumes. *Divided bodies: Lyme disease, contested illness, and evidence-based medicine*. Duke University Press, 2020.
- [88] Sarah Ann Singer. *The Empowerment Paradox: Rhetorics of Lyme Disease and the Future of Chronic Illness*. PhD thesis, The University of North Carolina at Chapel Hill, 2019.
- [89] Sarah Ciotti. Three papers on patient experiences with symptom persistent lyme disease in canada. 2023.
- [90] Kate Bloor, Vahsti Hale, and Alex Faulkner. Knowledge and uncertainty in lyme disease detection: an evidence-based activism research study in the uk. *Critical Public Health*, 31(4):416–428, 2021.
- [91] Mark B Brown. *Gil eyal, the crisis of expertise*: Polity press, cambridge, uk, 2019, 190 pp, isbn: 978-0745665788 (pbk), 2020.
- [92] Alan Petersen, Claire Tanner, and Megan Munsie. Navigating the cartographies of trust: how patients and carers establish the credibility of online treatment claims. *Sociology of Health & Illness*, 41:50–64, 2019.
- [93] Anna Kata. A postmodern pandora’s box: anti-vaccination misinformation on the internet. *Vaccine*, 28(7):1709–1716, 2010.
- [94] Bruno Latour and Steve Woolgar. *Laboratory life: The construction of scientific facts*. Princeton university press, 2013.
- [95] Sheila Jasanoff, Gerald E Markle, James C Peterson, and Trevor Pinch. *Handbook of science and technology studies*. Sage publications, 2001.
- [96] Bruno Latour. Science in action: How to follow scientists and engineers through society. *Harvard UP*, 1987.
- [97] Karin Knorr Cetina. *Epistemic cultures: How the sciences make knowledge*. harvard university press, 1999.
- [98] Thomas S. Kuhn. *The Structure of Scientific Revolutions*. University of Chicago Press, 1962.
- [99] William C Cockerham. *Medical sociology*. Routledge, 2017.
- [100] Robert M Entman. Framing: Toward clarification of a fractured paradigm. *Journal of communication*, 43(4):51–58, 1993.
- [101] Tom Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared D Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, et al. Language models are few-shot learners. *Advances in neural information processing systems*, 33:1877–1901, 2020.
- [102] Jason Wei, Yi Tay, Rishi Bommasani, Colin Raffel, Barret Zoph, Sebastian Borgeaud, Dani Yogatama, Maarten Bosma, Denny Zhou, Donald Metzler, et al. Emergent abilities of large language models. *arXiv preprint arXiv:2206.07682*, 2022.
- [103] Fabio Dennstädt, Johannes Zink, Paul Martin Putora, Janna Hastings, and Nikola Cihoric. Title and abstract screening for literature reviews using large language models: an exploratory study in the biomedical domain. *Systematic reviews*, 13(1):158, 2024.
- [104] Emily M Bender, Timnit Gebru, Angelina McMillan-Major, and Shmargaret Shmitchell. On the dangers of stochastic parrots: Can language models be too big? In *Proceedings of the 2021 ACM conference on fairness, accountability, and transparency*, pages 610–623, 2021.
- [105] Danah Boyd and Kate Crawford. Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon. *Information, communication & society*, 15(5):662–679, 2012.
- [106] Anne-Wil Harzing. Publish or perish, 2007. Available from <https://harzing.com/resources/publish-or-perish/>.
- [107] Teo Susnjak. Applying bert and chatgpt for sentiment analysis of lyme disease in scientific literature. In *Borrelia burgdorferi: Methods and Protocols*, pages 173–183. Springer, 2024.
- [108] Aaron M Cohen, William R Hersh, Kim Peterson, and Po-Yin Yen. Reducing workload in systematic review preparation using automated citation classification. *Journal of the American Medical Informatics Association*, 13(2):206–219, 2006.

- [109] Nodira Nazyrova, Salma Chahed, Thierry Chausalet, and Miriam Dwek. Leveraging large language models for medical text classification: a hospital readmission prediction case. In *2024 14th International Conference on Pattern Recognition Systems (ICPRS)*, pages 1–7. IEEE, 2024.
- [110] Qusai Khraisha, Sophie Put, Johanna Kappenberg, Azza Warraitch, and Kristin Hadfield. Can large language models replace humans in the systematic review process? evaluating gpt-4’s efficacy in screening and extracting data from peer-reviewed and grey literature in multiple languages. *arXiv preprint arXiv:2310.17526*, 2023.
- [111] Eugene Syriani, Istvan David, and Gauransh Kumar. Screening articles for systematic reviews with chatgpt. *Journal of Computer Languages*, page 101287, 2024.
- [112] Fabrizio Gilardi, Meysam Alizadeh, and Maël Kubli. Chatgpt outperforms crowd workers for text-annotation tasks. *Proceedings of the National Academy of Sciences*, 120(30):e2305016120, 2023.
- [113] Mayur Wankhade, Annavarapu Chandra Sekhara Rao, and Chaitanya Kulkarni. A survey on sentiment analysis methods, applications, and challenges. *Artificial Intelligence Review*, 55(7):5731–5780, 2022.
- [114] Dilek Küçük and Fazli Can. Stance detection: A survey. *ACM Computing Surveys (CSUR)*, 53(1):1–37, 2020.
- [115] Xiaochong Lan, Chen Gao, Depeng Jin, and Yong Li. Stance detection with collaborative role-infused llm-based agents. In *Proceedings of the International AAAI Conference on Web and Social Media*, volume 18, pages 891–903, 2024.
- [116] Norman Fairclough. *Discourse and Social Change*. Polity Press, 1992.
- [117] Matthew Renze and Erhan Guven. Self-reflection in llm agents: Effects on problem-solving performance. *arXiv preprint arXiv:2405.06682*, 2024.
- [118] Jacob Cohen. A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20(1):37–46, 1960. doi:10.1177/001316446002000104.
- [119] Joseph L. Fleiss. Measuring nominal scale agreement among many raters. *Psychological Bulletin*, 76(5):378–382, 1971. doi:10.1037/h0031619.
- [120] J. Richard Landis and Gary G. Koch. The measurement of observer agreement for categorical data. *Biometrics*, 33(1):159–174, 1977. doi:10.2307/2529310.
- [121] Mary L. McHugh. Interrater reliability: The kappa statistic. *Biochemia Medica*, 22(3):276–282, 2012. URL <https://pubmed.ncbi.nlm.nih.gov/23092060/>.
- [122] Anthony J. Viera and Joanne M. Garrett. Understanding interobserver agreement: The kappa statistic. *Family Medicine*, 37(5):360–363, 2005. URL <https://pubmed.ncbi.nlm.nih.gov/15883903/>.
- [123] Julius Sim and Chris C. Wright. The kappa statistic in reliability studies: Use, interpretation, and sample size requirements. *Physical Therapy*, 85(3):257–268, 2005. URL <https://academic.oup.com/ptj/article/85/3/257/2805022>.
- [124] Kevin A. Hallgren. Computing inter-rater reliability for observational data: An overview and tutorial. *Tutorials in Quantitative Methods for Psychology*, 8(1):23–34, 2012. URL <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3402032/>.
- [125] Virginia Braun and Victoria Clarke. Using thematic analysis in psychology. *Qualitative research in psychology*, 3(2):77–101, 2006.
- [126] M Schreier. Qualitative content analysis in practice. *Jacobs University Bremen*, 2012.
- [127] Adele E Clarke and Carrie Friese. Grounded theorizing using situational analysis. *The Sage handbook of grounded theory*, pages 363–397, 2007.
- [128] Phil Brown, Stephen Zavestoski, Sabrina McCormick, Brian Mayer, Rachel Morello-Frosch, and Rebecca Gasior Altman. Embodied health movements: new approaches to social movements in health. *Sociology of health & illness*, 26(1):50–80, 2004.
- [129] Abraham Savitzky and Marcel J. E. Golay. Smoothing and differentiation of data by simplified least squares procedures. *Analytical Chemistry*, 36(8):1627–1639, 1964. doi:10.1021/ac60214a047. URL <https://doi.org/10.1021/ac60214a047>.
- [130] Pierre Bourdieu. The specificity of the scientific field and the social conditions of the progress of reason. *Social Science Information*, 14(6):19–47, 1975. doi:10.1177/053901847501400602.
- [131] Maxwell E McCombs and Donald L Shaw. The agenda-setting function of mass media. *Public opinion quarterly*, 36(2):176–187, 1972.
- [132] Everett M Rogers. *Diffusion of innovations (5th Ed.)*. Free Press, 2003.

A Classification Prompts and Example Outputs

Below is the prompt used for pre-screening classification GPT-4o-mini to execute the Step 2a of the methodology depicted in Figure 2.

```
You are a world-leading expert in medical literature, specialising in Lyme disease and the controversies surrounding chronic Lyme disease (CLD) and post-treatment Lyme disease syndrome (PTLDS).

Your task is to evaluate a single scientific paper abstract at a time to determine if it has relevance to either CLD or PTLDS for further processing.

Definitions:
- Chronic Lyme Disease (CLD): A contested condition wherein some believe symptoms result from an ongoing Borrelia burgdorferi infection is often cited as requiring extended antibiotic therapy.
- Post-Treatment Lyme Disease Syndrome (PTLDS): Refers to persistent symptoms (e.g., fatigue, pain, cognitive issues) following standard Lyme disease treatment, generally understood not to involve active infection.

Categories:
Each abstract should be categorised into one of the following three categories:

1. Potentially Related to CLD/PTLDS:
- Select this category if the paper involves human studies and contains explicit or implicit references, findings or terminology potentially related to CLD or PTLDS. This includes abstracts where the connection to CLD or PTLDS is indirect, uncertain, or requires additional processing to determine relevance.

2. Definitely Unrelated:
- Select this category if the study has no relevance to CLD or PTLDS in human subjects. Abstracts discussing other medical conditions and unrelated tick-borne diseases.

3. Animal Study:
- Select this category if the study is entirely focused on animal subjects with no direct relevance to human cases of CLD or PTLDS.

Input:
For each abstract, you will receive an index number, a paper title, and the abstract text.

JSON Output Structure:
Your output must be a JSON object containing only the evaluation of the current abstract. Each response should include the following fields:
- 'index': The index number of the abstract, which MUST match the input.
- 'classification': One of "Potentially Related to CLD/PTLDS," "Definitely Unrelated," or "Animal Study."
- 'confidence': Confidence level in the classification ("High," "Medium," or "Low").

Example Output:
{
  "index": 700,
  "classification": "Potentially Related to CLD/PTLDS",
  "confidence": "High"
}

ABSTRACT:
< input >
```

Following the above classification, stance-framing detection was then executed as outlined in Step 2b of the methodology in Figure 2 with the abridged prompt below, where the three abstracts provided to the LLM as part of few-shot learning are omitted for brevity, but can be found on the Github repository of the project:

```
You are a world-leading expert in medical literature analysis and particularly in medical literature, specialising in Lyme disease and the debates surrounding chronic Lyme disease (CLD) and post-treatment Lyme disease syndrome (PTLDS). Your task is to evaluate a series of scientific paper abstracts related
```

to Lyme disease and determine each abstract's stance on PTLDS and CLD. This task requires scientific impartiality to minimize any bias in classifying abstracts concerning disputed subjects such as CLD or PTLDS. The current literature on this topic is polarised and can be expressed as follows:

"Some medical experts argue that post-treatment symptoms, experienced by a subset of patients after completing standard antibiotic therapy, can be attributed to what they define as Post-Treatment Lyme Disease Syndrome (PTLDS). According to some health organisations, these symptoms-ranging from fatigue to cognitive impairment-are likely caused by immune responses or tissue damage, rather than persistent infection. PTLDS concerns patients who experience persistent symptoms for at least six months after the completion of recommended treatment for Lyme disease.

Conversely, other organisations advocate for the recognition of chronic Lyme disease (CLD), contending that ongoing infection or immune dysfunction may be responsible for these symptoms and recommend extended antibiotic regimens, pointing to contested evidence of patient improvement. CLD is a broader term that encompasses a range of unexplained symptoms that patients attribute to a persistent Lyme disease infection that survives recommended treatment. Both PTLDS and CLD involve persistent symptoms, such as fatigue, pain, and neurological issues, and both face with diagnostic difficulties due to nonspecific symptoms and lack of clear biomarkers, contributing to ongoing debates in the medical community."

Based on the above context and the debate, your task is to classify each abstract and determine each abstract's explicit or implicit stance regarding CLD or PTLDS based on the following classifications:

- Supports PTLDS
- Supports CLD
- Neutral
- Unrelated
- Animal Study

Definitions for Classification:

1. Supports PTLDS:

- Attributes persistent symptoms after successful Lyme disease treatment to other factors like immune dysfunction, chronic fatigue syndrome, or depression.
- Opposes the use of prolonged or repeated antibiotic treatments for PTLDS.
- Suggests or implies that CLD lacks scientific support.

2. Supports CLD:

- A contested condition wherein some believe symptoms result from an ongoing or active *Borrelia burgdorferi* infection after standard treatment, often cited as requiring extended antibiotic therapy.
- Suggests that persistent symptoms may require prolonged or repeated antibiotic treatment.

3. Neutral:

- Presents balanced arguments without clearly supporting or refuting either PTLDS or CLD.
- Discusses aspects of Lyme disease that relate to both sides of the PTLDS and CLD debate without taking a clear stance.
- No clear stance or support or refutation is provided regarding PTLDS or CLD, either explicitly or implicitly.

4. Unrelated:

- The abstract does not mention PTLDS or CLD, or the surrounding debate at all.
- Focuses on other aspects of Lyme disease, such as acute Lyme disease, vector control, ecological studies, epidemiology, or diagnostic methods without mentioning chronic conditions.
- References CLD without any explicit or implied previous treatment.

5. Animal Study:

- The abstract pertains exclusively to animal models or non-human subjects without direct implications for PTLDS or CLD in humans.
- Does not involve human clinical data or direct conclusions about PTLDS or CLD in humans.
- If the animal study provides evidence directly relevant to PTLDS or CLD in humans, classify it accordingly under "Supports PTLDS", "Supports CLD", or "Neutral".

Additionally, assign a Confidence Level to each classification you make: High, Medium, or Low, based on how confident you are in the given classification.

- High: The chosen classification is indisputably true and cannot possibly be any of the other options. The abstract contains explicit wording relating to its stance.
- Medium: The chosen classification requires a mix explicit and implicit evidence in the abstract. There is some degree of certainty, but the classification is inferred rather than explicitly stated. Typically arises from indirect language, ambiguous terms, or implied support for a stance.
- Low: The classification is uncertain due to lack of clear evidence or conflicting information in the abstract, relying more on inference and contextual understanding. Typically occurs when the abstract is vague, uses ambiguous language, or includes evidence that could be interpreted in multiple ways. Reflects significant uncertainty in the classification, but that the chosen classification is more likely than the others.

Also, provide a reason comprising 2-3 sentences for your classification above. Clearly state your justification and if the abstract implicitly or explicitly supports PTLDS or CLD in humans. When determining the abstract's stance, consider both explicit statements and implications of the abstracts given context and chosen topic. Look for language that suggests support or opposition even if not directly stated.

Input:

For each abstract, you will receive an index number of the abstract, a paper title, and the abstract text.

JSON Output Structure:

Your output must be a JSON object containing the following fields:

- 'index': The index number of the abstract, which MUST match the input.
- 'classification': One of "Supports PTLDS", "Supports CLD", "Neutral", "Unrelated", or "Animal Study"
- 'confidence': Confidence level in the classification ("High", "Medium", or "Low").
- 'reason': Provide a 2-3 sentence justification for your classification.

< few-shot learning examples comprising 3 abstracts and their classification >

ABSTRACT:

< input >

A.1 Example classification results post self-reflection

A selection of abstracts whose classification changed from one category to the next post self-reflection prompting together with the justification texts, are shown in Table 6.

B Inter-rater Reliability Web Application

B.1 Web App with Example Texts and Outputs

Figure 10 illustrates our custom-built web application created specifically for the human validation process (Step 3 in Figure 2). In this interface, a user (in this case, a subject-expert rater) logs in to evaluate a given abstract. The application displays the original abstract text and prompts the user to choose one of several classification labels (e.g., Supports PTLDS, Supports CLD, Neutral, or Animal Study). Next, the rater selects a confidence level in their choice (e.g., Low, Medium, or High) and selects one of the two options that best describes the justification for the classification.

Figure 11 depicts another view of the same web application but for a different abstract. This example emphasizes the classification process itself, highlighting how a user can read through the abstract, reflect on the core content (e.g., whether it focuses on CLD, PTLDS, or otherwise), and then decide upon the appropriate classification label. The interface displays the choices that a user has selected in green. Together, both figures demonstrate the workflow that expert human raters followed in the IRR analysis process to validate the AI-based stance detection and justification texts provided for the rationale behind the LLM's decisions.

Table 6: Selected examples of abstracts with original and revised classifications and justifying text.

Abstract	Original Classification	Revised Classification	Original Justification	Revised Justification
<p><i>Background: IL-26 has demonstrated antimicrobial properties, as well as in the degradation of DNA from the Lyme disease spirochete Borrelia burgdorferi (Bb). Additionally, IL-26 can promote macrophage activation and enhance Bb phagocytotic activity. It is unclear if cell-mediated immune responses are modulated through TLR9 signaling when exposed to IL-26 Bb DNA complexes in post-treatment Lyme disease syndrome (PTLDS). Objective: We here aim to explore the effect of IL-26 in human Toll-like receptor (TLR)-9's activation upon the recognition of Bb DNA. Methods: We utilized a single-receptor cell system, HEK-Dual™ hTLR9 cells, which harbors two reporter plasmids for the NF-κB and IL-8 signaling pathways. Bb DNA was exposed to increasing concentrations of IL-26 in monomeric or dimeric form and then used to stimulate the cells for 4 h. The TLR-9 ligand CpG was used as a control. Results: We observed that NF-κB and IL-8 activation was maximal when the cells were stimulated with Bb DNA that had been treated with 5 μM of IL-26 monomer and 1 μM of IL-26 dimer. The same was observed for IL-8 activation upon CpG stimulation. We observed, however, a decrease in NF-κB activation when treated with either form of IL-26. An NF-κB activation increase did not occur with IL-26-treated TLR9 ligand CpG. Conclusions: Our study shows an enhancement in NF-κB and IL-8 activation upon the recognition of IL-26-treated Bb DNA by TLR9, which suggests an increase in sensing by the TLR9 of Bb DNA when it is in the form of an IL-26-Bb DNA complex. These findings will prompt further studies on the interaction between IL-26 and Bb DNA. 145</i></p>	Neutral	Supports PTLDS	The abstract discusses the role of IL-26 in the immune response to Borrelia burgdorferi DNA and its potential implications for post-treatment Lyme disease syndrome (PTLDS). However, it does not take a clear stance on whether persistent symptoms are due to ongoing infection or other factors, nor does it advocate for or against the concepts of PTLDS or chronic Lyme disease (CLD). The focus is primarily on the biochemical interactions and immune mechanisms without explicit support for either side of the debate.	Upon reevaluation, the abstract explicitly mentions 'post-treatment Lyme disease syndrome (PTLDS)' and explores immune responses related to this condition. The focus on IL-26's role in enhancing immune sensing of Bb DNA suggests a mechanism that aligns with the PTLDS perspective, which attributes persistent symptoms to immune dysfunction rather than ongoing infection. This indicates a clear inclination towards supporting PTLDS, justifying a revision to 'Supports PTLDS' with a High confidence level.
<p><i>Arthritis is the most common late manifestation of Borrelia burgdorferi infection in the United States, usually beginning months after the tick bite. In most patients with Lyme arthritis (LA) today, arthritis is the presenting manifestation of the disease. Patients have swelling and pain in one or a few large joints, especially the knee. Serologic testing is the mainstay of diagnosis. Responses to antibiotic treatment are generally excellent, although a small percentage of patients have persistent, postinfectious synovitis after 2 to 3 months of oral and IV antibiotics, which respond to anti-inflammatory therapies. Herein we review the clinical presentation, diagnosis, and management of LA. 7787</i></p>	Unrelated	Neutral	The abstract focuses on Lyme arthritis as a manifestation of Borrelia burgdorferi infection, discussing its clinical presentation, diagnosis, and management. While it mentions persistent postinfectious synovitis, it does not engage with the debates surrounding PTLDS or CLD, nor does it imply a stance on these conditions. The content is primarily concerned with the acute and post-treatment aspects of Lyme arthritis rather than chronic conditions.	Upon reevaluation, the abstract does mention persistent postinfectious synovitis, which could be interpreted as a connection to ongoing symptoms after treatment, potentially linking it to the PTLDS debate. However, it does not explicitly support or refute either PTLDS or CLD, making it more appropriate to classify it as Neutral rather than Unrelated. The confidence level is adjusted to Medium, as the abstract does touch on relevant themes but lacks a clear stance.
<p><i>The immuno-pathogenetic mechanisms underlying chronic Lyme neuroborreliosis are mainly unknown. Human Borrelia burgdorferi (Bb) infection is associated with Bb-specific secretion of interferon-γ (IFN-γ), which may be important for the elimination of Bb, but this may also cause tissue injury. In order to increase the understanding of the pathogenic mechanisms in chronic neuroborreliosis, we investigated which cell types that secrete IFN-γ. Blood mononuclear cells from 13 patients with neuroborreliosis and/or acrodermatitis chronica atrophicans were stimulated with Bb antigen and the phenotypes of the induced IFN-γ-secreting cells were analyzed with three different approaches. Cells expressing CD8 or TCRγδ, which both have cytolytic properties, were the main phenotypes of IFN-γ-secreting cells, indicating that tissue injury in chronic neuroborreliosis may be mediated by cytotoxic cells.</i></p>	Supports CLD	Neutral	The abstract discusses the immunological response in patients with chronic Lyme neuroborreliosis, suggesting that the secretion of interferon-γ by specific cell types may contribute to tissue injury. This aligns with the CLD perspective, which posits that ongoing immune responses and potential persistent infection could lead to chronic symptoms. However, the abstract does not explicitly advocate for prolonged antibiotic treatment or directly state that ongoing infection is the cause of symptoms, leading to a medium confidence classification.	Upon reevaluation, the abstract primarily focuses on the immunological mechanisms involved in chronic Lyme neuroborreliosis without explicitly supporting the notion of ongoing infection or the need for prolonged antibiotic treatment. While it discusses the role of immune responses in tissue injury, it does not provide clear evidence that aligns with the CLD perspective. Therefore, it is more accurately classified as Neutral, as it does not take a definitive stance on either PTLDS or CLD.

Lyme Classifier Classifier Hello user1@lyme.nz! Log off

Classifier

Article 104 out of 181 [See full List](#)

Abstract
We previously showed that numerous polymorphonuclear leukocyte (PMN) granule components efficiently kill *Borrelia burgdorferi*, the agent of Lyme disease. In addition, motile, granule-poor cytoplasts (U-Cyt) from human blood PMN can exert anti-*Borrelia* activity against opsonized *B. burgdorferi* independently of oxidative mechanisms. Here we show that lysates of U-Cyt also possess anti-*Borrelia* activity, a portion of which comes from the abundant cytosolic protein calprotectin. The anti-*Borrelia* activity of U-Cyt lysates and recombinant calprotectin was partially or completely reversed by specific antibody to calprotectin and by Zn(2+), a cation essential for the growth of *B. burgdorferi* and known to inhibit the antimicrobial activity of calprotectin. Quantitative microscopic and regrowth assays revealed that calprotectin acted in a bacteriostatic fashion against *B. burgdorferi*. We conclude that calprotectin, a potent bacteriostatic agent from a cell primarily recognized for its oxidative and granular antibacterial mechanisms, may play a modulatory role in infection by the Lyme spirochete, particularly at sites of acute inflammation.

Classification

Neutral	Supports PTLDS	Supports CLD	Unrelated	Animal Study
---------	----------------	--------------	-----------	--------------

Confidence

Low	Medium	High
-----	--------	------

Justification

Reasoning1	Reasoning2
Upon reevaluation, while the abstract focuses on healthcare providers' knowledge and experiences, this could be interpreted as indirectly relevant to the PTLDS/CLD debate by highlighting the challenges in communicating with patients about chronic Lyme.	Upon reevaluation, the abstract does not specifically address PTLDS or CLD, nor does it focus on animal studies. Instead, it discusses the bacteriostatic properties of calprotectin against <i>Borrelia burgdorferi</i> , which is relevant to the understanding of Lyme disease but does not engage with the chronic aspects of the disease or the debates surrounding PTLDS and CLD. Therefore, it is more accurately classified as 'Unrelated' to the PTLDS and CLD discussions, with a Medium confidence level due to the abstract's focus on a specific protein's role in infection rather than chronic disease outcomes.

Option 1 Option 2

Save

Figure 10: IRR web app showing the user abstract classification interface

Lyme Classifier Classifier Hello user1@lyme.nz! Log off

Classifier

Article 106 out of 181 [See full List](#)

Abstract
Lyme disease (LD) is most often diagnosed clinically, and the differential diagnosis of erythema migrans may be challenging. Recent advances have raised questions about the efficacy of traditional diagnostic modalities, but may soon facilitate consistent identification of patients with Lyme borreliosis. Therapeutic recommendations vary with the stage of disease, and treatment usually leads to complete resolution. The management of patients with "chronic Lyme disease" is controversial. A number of preventative measures have been evaluated; those involving the avoidance of tick bites with protective clothing and insect repellents remain the simplest and most effective. © 2010 by the American Academy of Dermatology, Inc.

Classification

Neutral

Supports PTLDS

Supports CLD

Unrelated

Animal Study

Confidence

Low

Medium

High

Justification

Reasoning1	Reasoning2
Upon reevaluation, the abstract maintains a neutral stance as it discusses the management and prevention of Lyme disease without endorsing either PTLDS or CLD. It acknowledges the controversy surrounding 'chronic Lyme disease' but does not provide evidence or language that would suggest support for either side. The mention of treatment leading to complete resolution further reinforces the neutral position, as it does not imply ongoing infection or persistent symptoms. Therefore, the classification remains appropriate, but the confidence level is adjusted to High due to the clarity of the abstract's neutral position.	Upon reevaluation, the abstract does not maintain a neutral stance; instead, it implicitly leans toward supporting CLD by acknowledging the controversy surrounding 'chronic Lyme disease' while not explicitly dismissing the possibility of ongoing infection or persistent symptoms after treatment. The mention of treatment leading to complete resolution could be interpreted as a dismissal of the experiences of patients who do not achieve complete resolution, subtly undermining the PTLDS perspective. Therefore, the classification is inappropriate and the confidence level should be adjusted to Low due to the potential for misinterpretation of the abstract's stance.

Option 1

Option 2

Save

2025 Lyme Classifier

Figure 11: IRR web app demonstrating the classification process