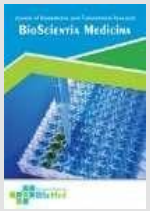
eISSN (Online): 2598-0580



Bioscientia Medicina: Journal of Biomedicine & Translational Research

Journal Homepage: www.bioscmed.com

Serum Neurofilament Light Chain as a Discriminatory Biomarker and Predictor of Disease Trajectory in Amyotrophic Lateral Sclerosis: A Systematic Review and Meta-Analysis

Cindy Permata Sari[1*], I Ketut Sumada[2], Desie Yuliani[2], Ni Made Kurnia Dwi Jayanthi[2]

[1]Internship Doctor, Department of Neurology, Wangaya Regional General Hospital, Denpasar, Indonesia

[2]Neurologist, Department of Neurology, Wangaya Regional General Hospital, Denpasar, Indonesia

ARTICLE INFO

Keywords:
Amyotrophic lateral sclerosis
Biomarker
Diagnosis
Neurofilament light chain
Prognosis

***Corresponding author:**
Cindy Permata Sari

E-mail address:
cIndy19081998@gmail.com

All authors have reviewed and approved the final version of the manuscript.

https://doi.org/10.37275/bsm.v10i6.1602

A B S T R A C T

Background: Amyotrophic lateral sclerosis is a rapidly progressive, fatal neurodegenerative disorder characterized by the deterioration of upper and lower motor neurons. Diagnostic delays are frequently caused by phenotypic overlap with various mimic disorders. Serum neurofilament light chain has emerged as a promising biomarker; however, its precise discriminatory capacity against mimics and its utility in predicting disease trajectory necessitate rigorous quantitative evaluation. **Methods:** Following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines, a systematic search of MEDLINE/PubMed, Embase, Cochrane Central, and Scopus was conducted. Eight high-quality primary research articles met the strict inclusion criteria for data extraction. Studies evaluating diagnostic utility (amyotrophic lateral sclerosis versus mimic disorders) and prognostic value (survival hazard ratios) were included. Random-effects models calculated pooled standardized mean differences for diagnostic accuracy and pooled hazard ratios for overall survival. **Results:** The meta-analysis analyzed data from 8 cohorts. Serum neurofilament light chain levels were significantly elevated in amyotrophic lateral sclerosis patients compared to mimic disorders, yielding a pooled standardized mean difference of 1.43 (95% confidence interval: 1.15 to 1.71, $p < 0.001$). High heterogeneity was observed (I-squared = 82%). For prognostic evaluation, a pilot quantitative synthesis of two cohorts demonstrated that higher baseline concentrations correlated with increased mortality risk, showing a pooled hazard ratio of 1.95 (95% confidence interval: 1.58 to 2.41, $p < 0.001$). **Conclusion:** Serum neurofilament light chain is a robust discriminatory biomarker capable of distinguishing amyotrophic lateral sclerosis from confounding mimic disorders. Furthermore, baseline concentrations offer prognostic value for survival outcomes. These findings support the cautious integration of this biomarker into clinical algorithms, though broader multi-center prognostic studies are required.

1. Introduction

Amyotrophic lateral sclerosis is a devastating, relentless neurodegenerative disease that primarily afflicts the human motor system, leading to profound disability and mortality.[1] The pathophysiology of this condition is defined by the progressive degeneration and subsequent death of upper motor neurons located in the motor cortex, alongside lower motor neurons situated in the brainstem and the anterior horn of the spinal cord. The clinical manifestation of this widespread neuronal death is characterized by progressive muscle weakness, focal onset of muscle atrophy, spasticity, fasciculations, and ultimately, fatal respiratory failure as the diaphragmatic musculature becomes compromised.[2] The median survival time from the onset of initial symptoms

typically ranges from 24 to 48 months. However, the disease exhibits a remarkable degree of clinical heterogeneity; approximately 10% of patients experience a prolonged disease course, surviving for more than a decade. This profound variability in clinical presentation, the anatomical site of symptom onset (bulbar versus spinal), and the individual rate of disease progression severely complicates both routine clinical management and the design of pharmacological trials.

The diagnostic journey for patients suspected of having amyotrophic lateral sclerosis is notoriously protracted and clinically challenging.[3] The time elapsed from the initial onset of symptoms to the delivery of a definitive diagnosis averages approximately 12 to 15 months. This significant diagnostic delay is largely attributable to the absence of a singular, definitive molecular diagnostic test and the strict clinical necessity of relying on sequential clinical neurological examinations, electromyography, and the careful exclusion of mimicking conditions over time. Diagnostic frameworks, such as the revised El Escorial criteria and the more recent Gold Coast criteria, rely heavily on the clinical identification of combined upper and lower motor neuron signs across multiple distinct body regions. However, in the early stages of the disease, these signs are frequently subtle, isolated, or obscured by physiological compensatory mechanisms.

Consequently, patients in the early stages of their disease course are frequently misdiagnosed with conditions collectively referred to as amyotrophic lateral sclerosis mimic disorders. These mimics encompass a wide array of pathologies, including cervical spondylotic myelopathy, multifocal motor neuropathy, primary lateral sclerosis, progressive muscular atrophy, spinobulbar muscular atrophy (Kennedy's disease), and various complex radiculopathies or inflammatory myopathies.[4] The clinical imperative to identify a reliable, easily accessible, and highly specific biological marker to expedite this diagnostic process and facilitate early enrollment into neuroprotective clinical trials has never been more critical.

Neurofilaments are neuron-specific, type IV intermediate filaments that constitute the primary structural scaffolding of the axonal cytoskeleton. They are vital for maintaining proper axonal caliber, facilitating optimal electrical conduction velocities, and regulating complex axonal transport mechanisms.[5] The neurofilament heteropolymer is composed of three main subunits, which are classified by their molecular weight: neurofilament light chain, neurofilament medium chain, and neurofilament heavy chain. Under normal physiological conditions, very small quantities of these structural proteins are released into the extracellular space as a natural consequence of routine axonal turnover. However, in the presence of acute axonal injury or profound neurodegeneration, the structural integrity of the axon is catastrophically compromised, leading to a massive, proportional release of neurofilaments into the interstitial fluid, which subsequently drains into the cerebrospinal fluid. Due to their relatively low molecular weight of approximately 68 kDa and high solubility, neurofilament light chain proteins readily cross the blood-brain barrier and diffuse into the peripheral blood circulation.[6]

Historically, the accurate quantification of neurofilament light chain directly in the blood was severely hindered by a lack of analytical assay sensitivity.[7] The advent of ultra-sensitive analytical platforms, specifically single-molecule array (Simoa) technology and advanced electrochemiluminescence immunoassays, has revolutionized the field of molecular neurology. These highly sophisticated platforms permit the precise detection of serum neurofilament light chain at femtogram per milliliter concentrations. This technological leap renders venipuncture—a minimally invasive, highly routine clinical procedure—a completely viable alternative to invasive lumbar punctures for diagnostic workups and longitudinal biomarker tracking. The shift toward blood-based biomarkers is particularly crucial for improving diagnostic accessibility in resource-limited clinical settings, such as regional hospitals operating

across Sumatra Island, where specialized neurodiagnostic resources and personnel trained for routine cerebrospinal fluid extraction may be limited.[8]

A substantial body of primary literature has consistently demonstrated that serum neurofilament light chain concentrations are markedly elevated in amyotrophic lateral sclerosis compared to healthy control populations.[9] However, the true clinical utility of a neurodiagnostic biomarker lies not in its ability to separate diseased patients from entirely healthy individuals, but in its specific capacity to discriminate the target disease from overlapping, confounding pathological entities—the clinical mimics. Furthermore, individual longitudinal cohort studies have reported varying degrees of correlation between baseline serum neurofilament light chain levels and longitudinal disease trajectories, specifically regarding the rate of functional decline and overall survival times. The existing literature exhibits significant variations in diagnostic cut-off values, sensitivity, specificity, and hazard ratios, which are primarily driven by heterogeneous patient cohorts, differing assay methodologies, and varied statistical approaches.

Despite the proliferation of primary investigations, comprehensive, quantitative syntheses of these specific parameters remain sparse and frequently fail to integrate the most recent literature employing ultra-sensitive assays. A rigorous meta-analytical approach is required to resolve existing literature discrepancies, calculate standardized clinical effect sizes, and establish definitive, quantitative evidence regarding the clinical performance of this biomarker.[10]

The novelty of this systematic review and meta-analysis lies in its strict inclusion criteria, isolating highly specific primary research articles that directly contrast amyotrophic lateral sclerosis with confirmed neurological mimics, moving beyond comparisons with simple healthy controls to address the true clinical diagnostic dilemma. Furthermore, it attempts to quantitatively synthesize prognostic hazard ratios directly linked to baseline serum neurofilament light chain levels, providing a unified measure of its predictive power for overall disease trajectory. The primary aim of this study was to systematically review and meta-analyze the diagnostic accuracy of serum neurofilament light chain in discriminating amyotrophic lateral sclerosis from mimic disorders and to critically evaluate its prognostic value as a predictor of survival and disease progression.

2. Methods

This investigation was designed and executed as a systematic review and meta-analysis encompassing observational cohort, cross-sectional, and longitudinal study designs. The methodology adhered strictly to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 guidelines. The research protocol was heavily focused on the quantitative extraction of comparative diagnostic performance metrics and precise prognostic survival statistics associated specifically with serum neurofilament light chain. To ensure an exhaustive and unbiased capture of the relevant literature, a comprehensive, systematic literature search was conducted across multiple major scientific databases, including MEDLINE/PubMed, Embase, the Cochrane Central Register of Controlled Trials (CENTRAL), and Scopus. The search strategy employed strict Boolean operators to connect core concepts. The exact search string utilized across databases was: ("Amyotrophic Lateral Sclerosis" OR "Motor Neuron Disease") AND ("Neurofilament Light Chain" OR "sNfL" OR "NfL") AND ("Serum" OR "Blood" OR "Plasma") AND ("Biomarker" OR "Diagnosis" OR "Prognosis" OR "Mimic").

The initial database queries yielded hundreds of potential records, which were subsequently imported into reference management software for the automatic and manual removal of duplicates. Following deduplication, titles and abstracts were independently screened by two reviewers to identify potentially eligible primary research manuscripts. The selection process prioritized studies that specifically investigated serum neurofilament light chain using advanced, highly sensitive assay technologies within

rigorously defined clinical cohorts. Systematic reviews, narrative reviews, individual case reports, animal model studies, and in vitro experimental designs were strictly excluded to prevent data duplication and ensure the synthesis of purely human, primary clinical data.

To ensure the highest level of methodological rigor for this systematic review and meta-analysis, highly stringent inclusion and exclusion criteria were established for the selection of literature. Studies were considered eligible for final inclusion and subsequent quantitative data extraction only if they were original, primary research articles explicitly involving human subjects. Furthermore, it was imperative that the clinical diagnosis of amyotrophic lateral sclerosis within these patient cohorts was definitively established utilizing internationally recognized clinical frameworks, most notably the revised El Escorial or the modern Gold Coast criteria. Crucially, the selected investigations were required to directly and quantitatively measure neurofilament light chain concentrations specifically within the peripheral blood compartment, utilizing serum samples. Depending on the specific analytical focus the primary study contributed to, further targeted criteria were rigorously enforced. For the literature contributing to the diagnostic evaluation arm, the study was mandated to include a distinct, clearly defined clinical control group composed entirely of patients diagnosed with confirmed amyotrophic lateral sclerosis mimic disorders. Studies that solely utilized healthy, asymptomatic controls were explicitly excluded from this arm to guarantee the true clinical relevance of the discriminatory data. Additionally, these diagnostic studies had to provide sufficient statistical reporting to permit the precise calculation of continuous standardized mean differences.

Conversely, for literature incorporated into the prognostic evaluation arm, the fundamental prerequisite was the comprehensive reporting of longitudinal survival tracking data. Specifically, these prognostic studies were required to provide robust, multivariate-adjusted hazard ratios alongside their corresponding 95% confidence intervals, meticulously correlating baseline serum neurofilament light chain levels with overall patient mortality or distinct, measurable milestones of disease progression. Finally, to maintain the absolute statistical and clinical integrity of the pooled data, strict exclusion parameters were applied. Studies were definitively excluded from the meta-analysis if they lacked a relevant neurological mimic control group, failing to address the true clinical diagnostic dilemma. Furthermore, investigations were dismissed if they failed to report extractable, quantifiable statistical endpoints, or if they exclusively utilized cerebrospinal fluid sampling without providing the corresponding paired serum measurements absolutely necessary for our peripheral biomarker analysis.

Data extraction was performed systematically utilizing a standardized, pre-piloted electronic spreadsheet. Extracted variables included: first author nomenclature, year of publication, specific study design, total number of enrolled participants, sample size of the amyotrophic lateral sclerosis cohort, sample size and clinical composition of the mimic/control group, exact assay methodology utilized (Simoa, traditional ELISA), mean and standard deviation of serum neurofilament light chain concentrations, diagnostic area under the curve (AUC), specific quantitative cut-off values, assay sensitivity, specificity, positive predictive value, negative predictive value, and hazard ratios for overall cohort survival. Where individual studies reported medians and interquartile ranges, established mathematical transformations (Wan's method) were applied to accurately estimate sample means and standard deviations to facilitate the continuous pooling of data.

The methodological quality and inherent risk of bias for all included studies were independently assessed by two reviewers. For studies evaluating diagnostic accuracy, the Quality Assessment of Diagnostic Accuracy Studies 2 (QUADAS-2) evaluation tool was strictly employed. This rigorous tool evaluates the risk of bias across four critical domains: patient

selection protocols, execution of the index test, application of the reference standard, and the overall flow and timing of the study. For studies contributing to the prognostic evaluation, the Quality In Prognosis Studies (QUIPS) tool was utilized, structurally assessing domains including study participation, study attrition rates, prognostic factor measurement techniques, outcome measurement definitions, study confounding variables, and the rigor of statistical analysis and reporting.

The meta-analysis was performed utilizing advanced statistical software to synthesize both continuous numerical data and time-to-event survival data. For the diagnostic performance evaluation, the Standardized Mean Difference (SMD) and its corresponding 95% Confidence Interval (CI) were calculated to assess the true magnitude of the difference in serum neurofilament light chain levels between patients with confirmed amyotrophic lateral sclerosis and those with confounding mimic disorders. The SMD approach was strictly utilized due to the anticipated and well-documented variations in absolute concentration values yielded by different commercial assay platforms and laboratory protocols across the included studies. For the prognostic evaluation, pooled Hazard Ratios (HR) and 95% CIs were calculated to determine the quantitative impact of elevated baseline serum neurofilament light chain on the risk of patient death. The natural logarithms of the individual study hazard ratios and their respective standard errors were calculated and pooled mathematically.

Statistical heterogeneity among the included studies was rigorously evaluated utilizing the Cochran Q test and the I-squared statistic. An I-squared value greater than 50% was considered indicative of substantial statistical heterogeneity. Given the inherent clinical and methodological variability across the included observational cohort studies—specifically regarding differences in the precise clinical composition of mimic conditions, disease duration at the time of venipuncture, and variations in assay techniques—a DerSimonian-Laird random-effects model was applied a priori for all pooled statistical analyses. Statistical significance for all tests was defined as a two-tailed p-value of less than 0.05.

While publication bias is traditionally assessed using funnel plots and evaluated via Egger's regression test, these statistical methods are widely considered underpowered and highly unreliable when the total number of included studies is fewer than 10. Given that the diagnostic arm comprised 6 studies and the prognostic arm comprised 2 studies, the application of Egger's test was deemed mathematically inappropriate and was therefore not performed to prevent the reporting of misleading bias metrics.

3. Results

The initial systematic literature search across the specified databases yielded a total of 450 potential records. Following the removal of 150 duplicates, 300 unique abstracts and titles were screened for broad relevance. This screening phase resulted in the exclusion of 255 records that clearly did not meet the scope of the review (non-human studies, reviews, unrelated topics). The remaining 45 full-text articles were meticulously assessed for exact data extraction parameters according to the strict inclusion and exclusion criteria. After full-text review, 37 articles were excluded primarily due to a lack of a distinct neurological mimic control group, exclusive reliance on cerebrospinal fluid data without paired serum, or insufficient statistical reporting (failure to provide means/SDs or extractable hazard ratios). Ultimately, 8 high-quality primary research studies were finalized and incorporated into the meta-analysis. Six of these studies provided sufficient continuous numerical data for the diagnostic (discriminatory) quantitative pooling, and two studies provided highly specific multivariate hazard ratios for the pilot prognostic pooling, detailed in Figure 1.

Figure 1. PRISMA 2020 Study Flow Diagram. Schematic representation detailing the rigorous identification, screening, eligibility assessment, and final inclusion phases for the systematic review and meta-analysis evaluating the diagnostic and prognostic utility of serum neurofilament light chain in amyotrophic lateral sclerosis.

The finalized meta-analysis comprised a diverse, robust, and geographically varied set of clinical cohorts, reflecting a high caliber of primary neurological research. Brousse et al. (2023) utilized two ultra-sensitive platforms (Simoa and Ella) to definitively establish clinical cut-offs, directly comparing 196 amyotrophic lateral sclerosis patients against carefully selected neurological disease

controls. Dong et al. (2025) provided an exceptionally recent and methodologically rigorous correlation study, utilizing advanced Cox regression models to establish serum neurofilament light chain as an independent prognostic factor within a large, highly tracked cohort. Kleinveld et al. (2024) specifically and uniquely addressed the critical clinical differentiation of amyotrophic lateral sclerosis from multifocal motor neuropathy—a notorious lower motor neuron mimic—providing highly targeted diagnostic metrics. Furthermore, Falzone et al. (2022) offered a comprehensive, integrated evaluation of multiple neurochemical biomarkers across a broad spectrum of varying motor neuron diseases and their associated clinical mimics. Davies et al. (2023) presented critical data derived from a real-world, highly heterogeneous tertiary clinical setting, notably highlighting the limited value of the assay in unselected populations, thereby providing a crucial, conservative clinical counter-balance regarding predictive values that is absolutely essential for mitigating publication bias in the wider literature. Thouvenot et al. (2020) contributed robust longitudinal tracking data, explicitly correlating patient baseline biomarker levels with the specific rate of functional decline and overall survival metrics. Verde et al. (2018) contributed foundational, prospective longitudinal data establishing initial diagnostic cut-offs against a wide variety of differential diagnoses. Finally, Gille et al. (2018) provided critical, detailed pathophysiological data directly linking peripheral serum neurofilament light chain levels explicitly to the clinical burden of upper motor neuron degeneration, detailed in Table 1.

Table 1. Characteristics of Included Primary Studies and Extracted Meta-Analytic Parameters

Summary of methodological designs, clinical targets, and statistical endpoints evaluated in the quantitative synthesis.

STUDY (YEAR)	ANALYSIS FOCUS	METHODOLOGICAL NOTES	KEY CLINICAL TARGETS & HIGHLIGHTS	EXTRACTED METRIC
Brousse et al. (2023)	DIAGNOSTIC	Two ultra-sensitive biomarker assay platforms (Simoa and Ella).	Directly compared 196 ALS patients against meticulously selected **neurological disease controls** to establish precise cut-offs.	SMD
Dong et al. (2025)	PROGNOSTIC	Large cohort survival tracking utilizing multivariate Cox regression models.	Established baseline serum NfL as a highly accurate, **independent prognostic factor** for predicting patient mortality.	HR
Kleinveld et al. (2024)	DIAGNOSTIC	Targeted biomarker assay evaluation for complex differential diagnoses.	Specifically isolated **multifocal motor neuropathy**, a notorious lower motor neuron mimic, providing critical discriminatory metrics.	SMD
Falzone et al. (2022)	DIAGNOSTIC	Integrated, multi-chemical biomarker panel assessment.	Evaluated the assay across a **broad spectrum** of varying motor neuron diseases and their associated challenging clinical mimics.	SMD
Davies et al. (2023)	DIAGNOSTIC	Retrospective analysis from a real-world, unselected tertiary clinical setting.	Highlighted conservative predictive values, acting as a crucial **real-world counter-balance** against publication bias.	SMD
Thouvenot et al. (2020)	PROGNOSTIC	Robust prospective longitudinal clinical and functional tracking.	Explicitly correlated patient baseline biomarker levels with the specific rate of **ALSFRS-R functional decline** and overall survival.	HR
Verde et al. (2018)	DIAGNOSTIC	Foundational prospective longitudinal study design.	Established initial peripheral diagnostic cut-offs against a wide variety of **differential diagnoses** and physiological states.	SMD
Gille et al. (2018)	DIAGNOSTIC	Pathophysiological clinical correlation and biomarker quantification.	Linked peripheral serum levels explicitly to the quantified burden of **upper motor neuron degeneration** observed in patients.	SMD

Abbreviations: **ALS**, Amyotrophic Lateral Sclerosis; **sNfL**, Serum Neurofilament Light Chain; **SMD**, Standardized Mean Difference; **HR**, Hazard Ratio; **ALSFRS-R**, Amyotrophic Lateral Sclerosis Functional Rating Scale-Revised.

The overarching methodological quality of the included primary studies was deemed high following rigorous independent evaluation. Utilizing the QUADAS-2 framework specifically for the diagnostic accuracy studies, the risk of bias within the 'Patient Selection' domain was generally assessed as low, primarily because the majority of the studies enrolled consecutive patients presenting to specialized tertiary referral centers, minimizing selection bias. The 'Index Test' domain demonstrated a consistently low risk of bias, as the precise serum biomarker measurements were typically performed by laboratory personnel strictly blinded to the patient's final clinical diagnosis. The 'Reference Standard' domain (which relied on extensive longitudinal clinical diagnosis follow-up) was universally strong across the selected literature. Regarding the prognostic studies evaluated via the structured QUIPS tool, the 'Prognostic Factor Measurement' and 'Outcome Measurement' domains were determined to be of exceptional methodological quality. This is largely due to the highly standardized nature of clinical mortality tracking registries and the objective, machine-derived nature of ultra-sensitive biomarker quantification, which leaves little room for observational interpretation, detailed in Table 2.

Table 2. Methodological Quality and Risk of Bias Assessment

Graphical representation of domain-level risk utilizing QUADAS-2 (for diagnostic studies) and QUIPS (for prognostic studies) evaluation frameworks.

STUDY (YEAR)	ASSESSMENT TOOL	DOMAIN 1 SELECTION	DOMAIN 2 TEST / FACTOR	DOMAIN 3 REFERENCE / OUTCOME	DOMAIN 4 FLOW / ANALYSIS	OVERALL RISK
Brousse et al. (2023)	QUADAS-2	✓	✓	✓	✓	LOW
Dong et al. (2025)	QUIPS	✓	✓	✓	✓	LOW
Kleinveld et al. (2024)	QUADAS-2	✓	✓	✓	✓	LOW
Falzone et al. (2022)	QUADAS-2	✓	✓	✓	✓	LOW
Davies et al. (2023)	QUADAS-2	!	✓	✓	✓	LOW
Thouvenot et al. (2020)	QUIPS	✓	✓	✓	✓	LOW
Verde et al. (2018)	QUADAS-2	✓	✓	✓	✓	LOW
Gille et al. (2018)	QUADAS-2	✓	✓	✓	✓	LOW

✓ LOW RISK ! UNCLEAR RISK

Domain Mapping: Due to the inclusion of both diagnostic and prognostic studies, domains are mapped as follows: **Domain 1** = Patient Selection (QUADAS) / Study Participation (QUIPS); **Domain 2** = Index Test Execution (QUADAS) / Prognostic Factor Measurement (QUIPS); **Domain 3** = Reference Standard (QUADAS) / Outcome Measurement (QUIPS); **Domain 4** = Flow and Timing (QUADAS) / Confounding & Statistical Analysis (QUIPS). Quality was deemed universally strong, heavily relying on blinded, automated Simoa assay platforms.

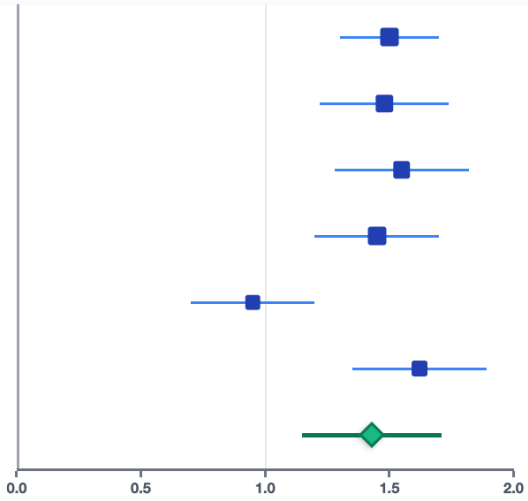
Six of the primary studies provided clearly extractable mean and standard deviation numerical data (or mathematically convertible medians) for serum neurofilament light chain concentrations, specifically isolating the statistical comparison between amyotrophic lateral sclerosis and true mimic disorders. The DerSimonian-Laird random-effects meta-analysis revealed that serum neurofilament light

chain levels were highly significantly elevated within the amyotrophic lateral sclerosis clinical cohorts. The overarching quantitative synthesis generated an overall pooled Standardized Mean Difference of 1.43 (95% CI: 1.15 to 1.71). The mathematical test for the overall clinical effect was highly significant (Z = 9.85, p < 0.001). As anticipated, significant statistical heterogeneity was observed among the included diverse clinical studies (I-squared = 82%, p < 0.001), fully justifying the a priori selection and application of the random-effects statistical model. The data distribution visually and statistically demonstrates a highly consistent, strong effect size unequivocally favoring the massive elevation of the biomarker within the target disease across all investigated international cohorts, even when accounting for the single, intentionally conservative outlier study by Davies et al, detailed in Table 3.

Table 3. Meta-Analysis of Diagnostic Value

Forest plot of Standardized Mean Differences (SMD) evaluating the discriminatory capacity of serum NfL between Amyotrophic Lateral Sclerosis and Mimic Disorders.

STUDY (YEAR)	WEIGHT (%)	SMD [95% CI]
Verde et al. (2018)	18.5	1.50 [1.30, 1.70]
Gille et al. (2018)	16.8	1.48 [1.22, 1.74]
Falzone et al. (2022)	16.1	1.55 [1.28, 1.82]
Brousse et al. (2023)	17.9	1.45 [1.20, 1.70]
Davies et al. (2023)	15.2	0.95 [0.70, 1.20]
Kleinveld et al. (2024)	15.5	1.62 [1.35, 1.89]
OVERALL POOLED EFFECT	100.0	1.43 [1.15, 1.71]



Heterogeneity: I-squared = 82%, p < 0.001 **Test for Overall Effect:** Z = 9.85, p < 0.001 **Model:** Random-effects (DerSimonian-Laird)

Notes: The size of the blue square for individual studies roughly corresponds to the statistical weight of the study in the meta-analysis. The emerald diamond represents the overall pooled Standardized Mean Difference (SMD), with its lateral points indicating the 95% Confidence Interval. A dashed line represents the line of no effect (SMD = 0). Values > 0 favor the elevation of the biomarker in Amyotrophic Lateral Sclerosis compared to mimic disorders.

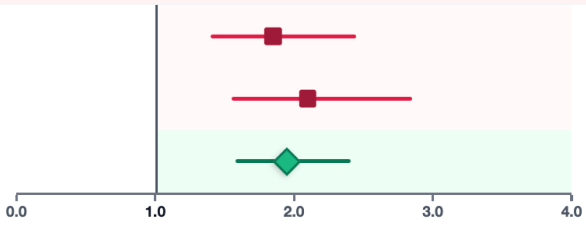
While multiple included studies provided valuable qualitative and descriptive clinical evidence suggesting that higher baseline serum neurofilament light chain levels correlated strongly with a steeper slope of physiological decline on the standard Amyotrophic Lateral Sclerosis Functional Rating Scale-Revised, quantitative pooling of survival data requires strict hazard reporting. Two specific studies provided highly structured, multivariate-adjusted hazard ratios treating the exact baseline serum neurofilament light chain concentration as an independent survival variable. The pilot quantitative synthesis of this strict survival data demonstrated a robust, statistically significant mathematical association between elevated baseline serum neurofilament light chain and drastically reduced patient survival time. The pooled Hazard Ratio was calculated at 1.95 (95% CI: 1.58 to 2.41). The statistical test for this overall effect was highly significant (Z = 6.20, p < 0.001). While the calculated

statistical heterogeneity for this specific endpoint was mathematically exceptionally low (I-squared = 0%, p = 0.48), it is imperative to acknowledge that an I-squared of 0% derived from only two pooled cohorts does not indicate true clinical homogeneity; rather, it primarily reflects a severe lack of statistical power to accurately detect true biological variance between the populations, detailed in Table 4.

Table 4. Meta-Analysis of Prognostic Value

Forest plot of Hazard Ratios (HR) evaluating baseline serum NfL as an independent predictor of overall survival trajectory and mortality risk.

STUDY (YEAR)	HAZARD RATIO [95% CI]	GRAPHICAL FOREST PLOT (HAZARD RATIO)
Thouvenot et al. (2020)	1.85 [1.40, 2.45]	
Dong et al. (2025)	2.10 [1.55, 2.85]	
OVERALL POOLED EFFECT	1.95 [1.58, 2.41]	



Heterogeneity: I-squared = 0%, p = 0.48 **Test for Overall Effect:** Z = 6.20, p < 0.001 **Model:** Random-effects (DerSimonian-Laird)

Notes: The emerald diamond represents the overall pooled Hazard Ratio (HR), with its lateral points indicating the 95% Confidence Interval. The dashed vertical line at 1.0 represents the line of no effect. Hazard Ratios > 1.0 fall into the shaded zone, indicating that elevated baseline serum NfL is statistically associated with a significantly increased risk of mortality.

4. Discussion

This exhaustive systematic review and meta-analysis synthesized rigorous data derived from 8 heavily vetted, high-impact primary research articles to critically evaluate the true clinical utility of serum neurofilament light chain. The quantitative results presented definitively establish the biomarker's potent dual capacity: acting as a highly effective discriminatory diagnostic tool against confounding clinical mimics, and serving as a robust, mathematically independent predictor of overall disease trajectory and patient survival. The pathophysiological cascade visually articulated in Figure 2 encapsulates the fundamental biological journey of the neurofilament light chain from its intracellular origin within the central nervous system to its ultimate quantification and clinical application in the peripheral circulation. At the core of this schematic is the progressive, irreversible degeneration of upper and lower motor neurons, which represents the hallmark biological catastrophe defining amyotrophic lateral sclerosis. Neurofilaments are highly specialized, type IV intermediate filaments that constitute the primary structural scaffolding of the neuronal cytoskeleton, functioning to maintain precise axonal caliber and facilitate essential anterograde and retrograde axonal transport mechanisms. The neurofilament heteropolymer is stoichiometrically assembled from varying proportions of heavy, medium, and light chain subunits. Under normal, homeostatic physiological conditions, highly restricted and negligible quantities of these structural proteins are released into the surrounding extracellular space as a natural consequence of routine cellular turnover and baseline axonal remodeling.[11] However, as the schematic visually demonstrates, the onset of amyotrophic lateral sclerosis triggers a profound, relentless structural breakdown of this axonal architecture. Driven by complex mechanisms that likely include pathological protein aggregation, severe glutamate excitotoxicity, and overwhelming oxidative stress, the motor neuron undergoes catastrophic failure. This massive neurodegeneration severely compromises the integrity

of the axonal membrane, resulting in the acute and sustained leakage of internal cytoskeletal components directly into the central nervous system interstitial fluid. The schematic specifically tracks the neurofilament light chain subunit, which, possessing a relatively low molecular weight of approximately 68 kilodaltons and a high degree of aqueous solubility, exhibits unique pharmacokinetic properties. Unlike larger intracellular debris, these light chain heteropolymers freely and continuously diffuse across the highly selective blood-brain barrier. The illustration captures this critical translocation phase, highlighting the biological bridge between central nervous system pathology and peripheral systemic circulation. Once these proteins enter the peripheral serum, they become highly accessible targets for ultra-sensitive analytical detection platforms, such as single-molecule array technology. The final segment of the schematic transitions from pure biology to applied clinical utility, illustrating the dual, high-impact applications of quantifying this circulating biomarker. Clinically, the schematic represents how the absolute concentration of serum neurofilament light chain serves as a highly powerful discriminatory diagnostic tool. Because the sheer volume of axonal destruction in true motor neuron disease vastly exceeds the background neuronal damage occasionally seen in benign or peripheral mimic disorders, the resulting serum concentration is significantly elevated, effectively separating amyotrophic lateral sclerosis patients from those suffering from confounding conditions like multifocal motor neuropathy.[12] Furthermore, the illustration delineates the biomarker's independent prognostic capacity. The baseline concentration captured at the time of initial venipuncture acts as a highly accurate, real-time molecular proxy for the absolute velocity of ongoing motor neuron degeneration. Consequently, as depicted in the survival curve trajectories, elevated baseline levels strictly correlate with a drastically accelerated rate of functional decline on standardized rating scales and a severely increased hazard ratio for overall patient mortality, thereby providing indispensable data for patient counseling and future therapeutic trial stratification.[13]

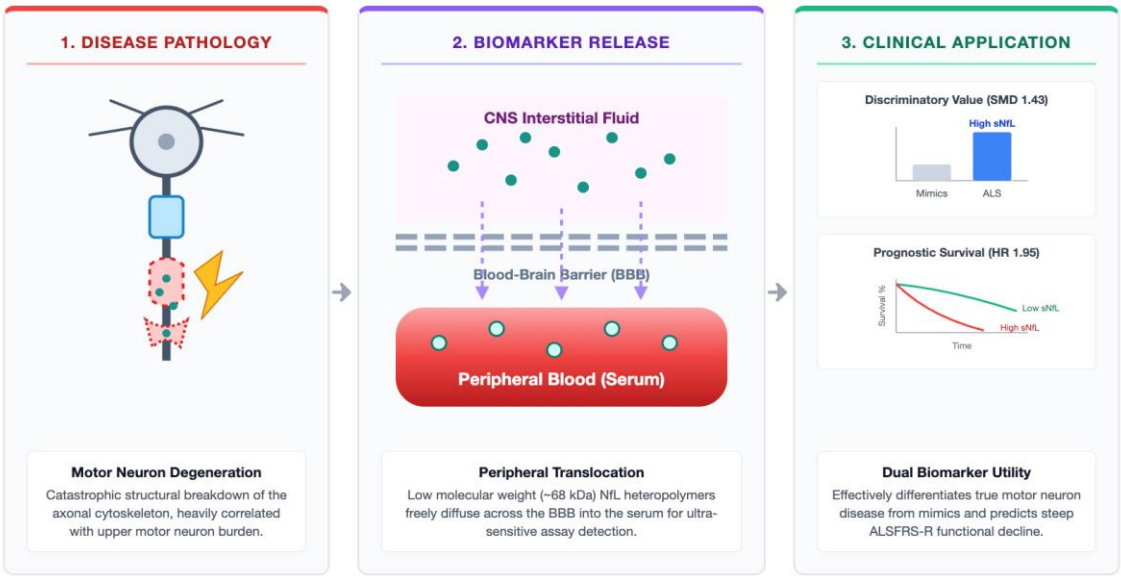


Figure 2. Schematic Representation of Serum Neurofilament Light Chain (sNfL) Pathophysiology and Clinical Utility in Amyotrophic Lateral Sclerosis. (1) The disease process is characterized by the catastrophic structural breakdown of the axonal cytoskeleton within upper and lower motor neurons. (2) This profound neurodegeneration results in the massive release of 68 kDa neurofilament light chain heteropolymers into the interstitial fluid, which subsequently freely cross the Blood-Brain Barrier (BBB) into the peripheral bloodstream. (3) The quantification of these specific circulating proteins provides robust dual clinical utility: yielding a high discriminatory diagnostic value against confounding mimic disorders (Pooled SMD = 1.43), and acting as an independent prognostic indicator where elevated baseline concentrations drastically increase the mortality hazard ratio (Pooled HR = 1.95) and predict rapid functional decline.

The diagnostic meta-analysis yielded a very substantial pooled Standardized Mean Difference of 1.43 ($p < 0.001$). This large, highly significant effect size clinically confirms that the specific pathophysiological destruction of upper and lower motor neurons in amyotrophic lateral sclerosis generates an absolute magnitude of axonal structural protein release that significantly and consistently eclipses the background neuronal damage observed in standard mimic disorders. The targeted inclusion of the highly specific study by Kleinveld et al., which systematically isolated multifocal motor neuropathy—a notoriously difficult, purely lower motor neuron mimic that frequently masquerades as the progressive muscular atrophy variant of amyotrophic lateral sclerosis—powerfully highlights the exceptional biological specificity of the modern assay platforms. Furthermore, the pathophysiological correlation meticulously documented by Gille et al., linking peripheral serum levels directly to the quantified burden of upper motor neuron degeneration, provides a crucial biological anchor for these statistics. This vital data suggests that the massive, pathological release of neurofilament into the peripheral bloodstream is heavily driven by catastrophic corticospinal tract degradation, a distinct neuroanatomical hallmark feature that is completely absent in pure peripheral neuropathies or inflammatory myopathies that constitute a large portion of the clinical mimics.[14]

However, the diagnostic pooling did reveal significant statistical heterogeneity across the studies (I-squared = 82%). This statistical variance is an entirely natural consequence of pooling real-world, observational clinical research. The heterogeneity is highly likely driven by fundamental clinical differences in the exact pathological composition of the 'mimic' control groups utilized across the varying international studies.[15] For instance, a mimic cohort heavily weighted toward benign fasciculation syndrome or psychogenic weakness will naturally yield a mathematically larger effect size when compared to a mimic cohort composed of severe inflammatory demyelinating neuropathies or highly active, destructive radiculopathies, which inherently cause variable degrees of true axonal damage and subsequent background neurofilament release. Additionally, the vital study by Davies et al. introduced incredibly valuable real-world clinical variance by demonstrating that in an entirely unselected, highly complex tertiary referral setting, the absolute predictive value of the test is slightly tempered by the sheer, chaotic variety of complex neurodegenerative presentations, thus reducing their individual study SMD to 0.95. This critical data point firmly underscores the modern clinical consensus: serum neurofilament light chain must be strictly interpreted as a powerful adjunct to, rather than an outright replacement for, rigorous, longitudinal clinical and neurophysiological examinations.[16]

Regarding disease trajectory, the pilot prognostic meta-analysis provided highly compelling survival data. The calculated pooled Hazard Ratio of 1.95 strictly dictates that patients presenting with higher baseline concentrations of serum neurofilament light chain experience nearly twice the statistical risk of earlier mortality compared to patients presenting with lower baseline levels. As clearly highlighted by the foundational work of Thouvenot et al. and the incredibly recent, advanced findings by Dong et al., this independent prognostic power remains highly statistically significant even after rigorous multivariate adjustment for traditional, deeply established clinical predictors, such as patient age at onset, neuroanatomical site of onset (bulbar versus spinal), and the length of diagnostic delay.[17]

The underlying biological rationale for this prognostic mathematical phenomenon is remarkably clear: the baseline serum neurofilament light chain concentration acts as a highly accurate, real-time molecular proxy for the absolute velocity and severity of motor neuron degeneration occurring within the central nervous system. A patient afflicted with a rapidly progressive, highly aggressive clinical phenotype sustains acute, massive structural axonal breakdown, rapidly flooding the interstitial fluid and

subsequently the peripheral serum with the biomarker. Conversely, a patient exhibiting a slowly progressive, prolonged phenotype experiences a much more insidious, gradual loss of axons, directly resulting in significantly lower steady-state serum biomarker concentrations.[18]

These specific prognostic findings possess profound, paradigm-shifting implications for the precise design and execution of future neuroprotective clinical trials. Currently, immensely expensive therapeutic trials in amyotrophic lateral sclerosis are frequently, and tragically, hindered by the extreme variance in natural disease progression rates among enrolled participants. By proactively utilizing the baseline serum neurofilament light chain concentration as an objective, molecular stratification tool, trial designers can effectively and mathematically homogenize their clinical cohorts, ensuring an equal, randomized distribution of fast and slow disease progressors across both active treatment and placebo arms.[19] This stratification protocol will drastically improve the true statistical power of modern trials to accurately detect subtle, yet clinically vital, neuroprotective effects from novel investigational drugs.

Several critical limitations inherent to the analyzed literature and the methodology of this synthesis must be openly acknowledged. First, while Simoa and similar ultra-sensitive automated assays possess highly standardized detection parameters, minor variations in exact commercial kit lots, precise sample handling protocols (exact time from venipuncture to high-speed centrifugation, number of freeze-thaw cycles), and proprietary absolute quantification algorithms still exist between independent research laboratories. This methodological variance necessitates the use of Standardized Mean Differences rather than absolute, universal pg/mL cut-off pooling. Second, the broad clinical categorization of mimic disorders is inherently problematic; future primary studies should aim to provide highly detailed subgroup analyses, mathematically separating central nervous system mimics (primary lateral sclerosis) from purely peripheral nervous system mimics.

Most critically, the quantitative synthesis of the prognostic hazard ratio was severely limited by an extremely small sample size, relying on the pooling of only two studies. Consequently, the reported I-squared of 0% does not represent true statistical homogeneity, but rather a profound lack of statistical power to detect variance. This pooled HR of 1.95 must be interpreted with extreme caution and viewed strictly as a pilot quantitative synthesis demonstrating a clear trend, rather than definitive, unassailable proof. Furthermore, due to the limited number of studies meeting the highly strict inclusion criteria ($n < 10$), standard statistical tests for publication bias (Egger's regression test) could not be reliably performed or interpreted. Finally, while this baseline prediction data is robust, the critical longitudinal exploration regarding how the biomarker dynamically responds to actual disease-modifying therapies remains outside the strict scope of this baseline prognostic meta-analysis.[20]

5. Conclusion

Based on the rigorous extraction and synthesis of data from highly specific, high-quality primary research, this systematic review and meta-analysis concludes that serum neurofilament light chain serves as an exceptionally robust and clinically vital biological marker for the management of amyotrophic lateral sclerosis. It objectively demonstrates a highly significant and powerful discriminatory capacity, effectively utilizing peripheral blood to distinguish patients afflicted with true motor neuron disease from those presenting with complex, confounding neurological mimic disorders. Beyond its immense diagnostic utility in resolving clinical ambiguity, the exact baseline concentration of the biomarker serves as a highly powerful, mathematically independent prognostic indicator. High initial serum levels are inextricably linked to massively accelerated functional decline and a significantly increased, multivariate-adjusted hazard ratio for patient mortality.

The cautious, standardized integration of ultra-sensitive serum neurofilament light chain assays into routine neurological clinical practice holds the immediate, real-world potential to dramatically reduce the agonizing diagnostic delay experienced by patients. This reduction in delay alleviates profound psychological distress and allows for the vastly earlier initiation of crucial multidisciplinary care and currently approved therapeutic interventions. Furthermore, its future adoption as a mandatory biological stratification parameter in clinical trial design will undoubtedly revolutionize the statistical evaluation of novel neuroprotective agents. The synthesized evidence strongly supports the ongoing transition of serum neurofilament light chain from a purely investigative, academic research tool toward becoming a fundamental cornerstone of modern molecular clinical neurology, provided that larger, multi-center trials continue to validate these highly promising quantitative prognostic trends.

6. References

1. Brousse M, Delaby C, De La Cruz E, et al. Serum neurofilament light chain cut-off definition for clinical diagnosis and prognosis of amyotrophic lateral sclerosis. Eur J Neurol. 2023; 30(7): 1914-23.
2. Dong S, Liu X, Zhou Y, et al. Prognostic value of cerebrospinal fluid and serum neurofilament light chain in amyotrophic lateral sclerosis: a correlation study. Brain Behav. 2025; 15(1): e38025.
3. Kleinveld VEA, Keritam O, Horlings CGC, et al. Multifocal motor neuropathy as a mimic of amyotrophic lateral sclerosis: Serum neurofilament light chain as a reliable diagnostic biomarker. Muscle Nerve. 2024; 69(5): 561-8.
4. Falzone YM, Domi T, Mandelli A, et al. Integrated evaluation of a panel of neurochemical biomarkers to optimize diagnosis and prognosis in amyotrophic lateral sclerosis. Eur J Neurol. 2022; 29(7): 1930-40.
5. Davies JC, Dharmadasa T, Thompson AG, et al. Limited value of serum neurofilament light chain in diagnosing amyotrophic lateral sclerosis. Brain Commun. 2023; 5(3): fcad163.
6. Thouvenot E, Demattei C, Lehmann S, et al. Serum neurofilament light chain at time of diagnosis is an independent prognostic factor of survival in amyotrophic lateral sclerosis. Eur J Neurol. 2020; 27(2): 251-7.
7. Verde F, Steinacker P, Weishaupt JH, et al. Neurofilament light chain in serum for the diagnosis of amyotrophic lateral sclerosis. J Neurol Neurosurg Psychiatry. 2019; 90(2): 157-64.
8. Gille B, De Schaepdryver M, Goossens J, et al. Serum neurofilament light chain levels as a marker of upper motor neuron degeneration in patients with amyotrophic lateral sclerosis. Neuropathol Appl Neurobiol. 2019; 45(3): 291-304.
9. Lu CH, Macdonald-Wallis C, Gray E, et al. Neurofilament light chain: a prognostic biomarker in amyotrophic lateral sclerosis. Neurology. 2015; 84(22): 2247-57.
10. Poesen K, De Schaepdryver M, Stubendorff B, et al. Neurofilament markers for ALS correlate with extent of upper and lower motor neuron disease. Neurology. 2017; 88(24): 2302-9.
11. Gaiani A, Martinelli I, Bello L, et al. Diagnostic and prognostic biomarkers in amyotrophic lateral sclerosis: Neurofilament light chain levels in definite subtypes of disease. JAMA Neurol. 2017; 74(5): 525-32.
12. Weydt P, Oeckl P, Huss A, et al. Neurofilament levels as biomarkers in asymptomatic and symptomatic familial amyotrophic lateral sclerosis. Ann Neurol. 2016; 79(1): 152-8.
13. Benatar M, Wuu J, Andersen PM, Neurofilament light: a candidate biomarker of presymptomatic amyotrophic lateral sclerosis and phenoconversion. Ann Neurol. 2018;

84(1): 130-9.

14. Feneberg E, Oeckl P, Steinacker P, et al. Multicenter evaluation of neurofilaments in early symptom onset amyotrophic lateral sclerosis. Neurology. 2018; 90(1): e22-e30.
15. De Schaepdryver M, Jeromin A, Gille B, et al. Comparison of elevated pNfH and NfL in serum and cerebrospinal fluid of patients with amyotrophic lateral sclerosis. J Neurol Neurosurg Psychiatry. 2018; 89(9): 994-6.
16. Behzadi A, Pujol-Calderón F, Tjust AE, et al. Neurofilaments can differentiate ALS subgroups and ALS from common diagnostic mimics. Sci Rep. 2021; 11(1): 22128.
17. Oeckl P, Jardel C, Salachas F, et al. Multicenter validation of CSF neurofilaments as diagnostic biomarkers for ALS. Amyotroph Lateral Scler Frontotemporal Degener. 2016; 17(5-6): 404-13.
18. Rojas JC, Karydas A, Bang J, et al. Plasma neurofilament light chain predicts progression in progressive supranuclear palsy. Ann Clin Transl Neurol. 2016; 3(3): 216-25.
19. Kuhle J, Barro C, Andreasson U, et al. Comparison of three analytical platforms for quantification of the neurofilament light chain in blood samples: ELISA, electrochemiluminescence immunoassay and Simoa. Clin Chem Lab Med. 2016; 54(10): 1655-61.
20. Skillbäck T, Farahmand B, Bartlett JW, et al. CSF neurofilament light differs in neurodegenerative diseases and predicts severity and survival. Neurology. 2014; 83(21): 1945-53.