Neurofilament light chain as neuronal injury marker – what is needed to facilitate implementation in clinical laboratory practice?

Abstract: Neurobiomarkers have attracted significant attention over the last ten years. One promising biomarker is the neurofilament light chain protein (NFL). Since the introduction of ultrasensitive assays, NFL has been developed into a widely used axonal damage marker of relevance to the diagnosis, prognostication, follow-up, and treatment monitoring of a range of neurological disorders, including multiple sclerosis, amyotrophic lateral sclerosis, and Alzheimer’s disease. The marker is increasingly used clinically, as well as in clinical trials. Even if we have validated precise, sensitive, and specific assays for NFL quantification in both cerebrospinal fluid and blood, there are analytical, as well as pre- and post-analytical aspects of the total NFL testing process, including biomarker interpretation, to consider. Although the biomarker is already in use in specialised clinical laboratory settings, a more general use requires some further work. In this review, we provide brief basic information and opinions on NFL as a biomarker of axonal injury in neurological diseases and pinpoint additional work needed to facilitate biomarker implementation in clinical practice.

Keywords: analytical phase; biomarkers; blood; cerebrospinal fluid; neurodegeneration; neurofilament light chain; neurological disease; post-analytical phase; pre-analytical phase.

Introduction

Neurological disorders are the primary cause of disability and the second cause of death according to the Global Burden of Diseases, Injuries, and Risk Factors Study (GBD) [1]. Currently, clinicians and medical laboratory professionals have few tests for the diagnosis, screening, and follow-up of neurological conditions, and the availability of these tests is limited to specialised laboratories. Some clinical laboratories are using neurobiomarkers in their routine practice, for example, total and phosphorylated tau proteins, amyloid β₁₋₄₂, and the amyloid β₁₋₄₂/₁₋₄₀ ratio for patients with dementia, especially Alzheimer’s disease [2]; cerebrospinal fluid (CSF) oligoclonal bands for multiple sclerosis [3]; and central nervous system (CNS) autoantibodies (anti-AQP4 against aquaporin water channel on astrocytes and anti-MOG against myelin oligodendrocyte glycoprotein on oligodendrocyte cell/myelin sheaths) for the differential and/or diagnosis of neuroinflammatory disorders [4]. In addition to these biomarkers, we have a relatively new biomarker of potential relevance to most neurological disorders, namely neurofilament light chain protein (NFL).

Neurofilaments and NFL – basic information

Neurofilaments (Nfs) are cylindrical proteins profusely expressed in large-caliber myelinated axons [5]. They support axons for stability, enabling the radial growth of myelinated axons. Nfs are classified as intermediate filaments (IFs), which consist of three subunits classified according to their molecular weight: neurofilament light (NFL), medium (NfM), and...
heavy (NFH), as well as α-internexin (a-int) and peripherin [6]. Under normal conditions (at a normal physiological state), low concentrations of NFs are perpetually released from axons into the CSF and passed to the blood at lower concentrations across the blood-brain barrier. This release increases in an age-dependent manner [7]. The release process is accelerated if there are damaging or degenerating processes affecting neuronal axons (central or peripheral), which is the basis for the use of NFs as biomarkers [8–10]. So far, although it may seem counterintuitive, there is very little evidence of an effect of blood-brain barrier impairment on plasma or serum NF-L concentration [11]. Additionally, most psychiatric diseases without significant neurodegeneration have normal NF L levels [12–14], which has been suggested to help differentiating, e.g., affective diseases and frontotemporal dementia [15]. The most studied NF biomarker is NF-L, which is the focus of the current review.

**Neurofilaments in clinical laboratory practice – a “call to arms” for laboratory professionals**

In 2016, Kuhle et al. published an article in Clinical Chemistry and Laboratory Medicine (CCLM) and compared three analytical platforms for the quantification of NF-L in blood samples [16]. This article highlighted the need for robust laboratory validation of research-developed biomarkers. Henceforth, a substantial number of studies have been published that underline the potential role of serum/plasma NF-L as a biomarker of neuroaxonal integrity mostly in MS, Alzheimer’s disease, and amyotrophic lateral sclerosis, to assess disease activity, neurodegeneration, and treatment responses. Nevertheless, it is essential to define a clear context of use (COU) for each proposed test that is of primary interest before their laboratory implication [17]. In an editorial of CCLM [18], laboratory professionals were “called to arms” to add value to neurobiomarkers in terms of standardized (or at least harmonized) pre-analytical, analytical and post-analytical procedures. Here, we review the studies that have been published to date, addressing these issues for NF-L.

**Pre-analytical phase – where we are and future directions**

Up to 70% of laboratory errors arise from pre-analytical variation, most of which are related to patient preparation, sample collection, transportation to laboratories, preparation period for analysis, and storage [19]. Before clinical implementation, it is important to determine how sensitive a biomarker is to pre-analytical variation and define standard operating procedures (SOPs) that minimize variation and bias due to relevant pre-analytical factors. In Figure 1, we summarize what factors to test for new biomarker candidates.

**Usual suspects of assay interference – what is relevant to NF-L?**

Haemolysis, lipemia, icterus, and biotin are ‘usual suspects’ of assay interference in clinical chemistry, and they may significantly affect test results. Before implementing NF-L in clinical laboratory practice, the possible effects of these interfering substances should also be elucidated. Thus far, some efforts have already been made for this goal. Recently, Lee et al. developed a highly sensitive NF-L assay for an automated immunoassay platform [20]. In this study, the researchers defined assay-specific cut points of each endogenous interfering substance. Besides, Midde et al. investigated the effects of interference from haemolysed and lipemic samples in recovering NF-L on the Simoa platform [21]. They detected no interference from haemolysis and lipemia in their study. However, it should be kept in mind that some of these interferences are assay-dependent – this is a topic of validation that should be re-visited for each assay developed.

**Effect of matrix choice on NF-L levels**

To date, it has been shown that serum and plasma result in slightly different absolute concentrations of NF-L [22, 23]: EDTA plasma shows around 10% lower NF-L concentration than serum, although plasma and serum NF-L concentrations correlate strongly [24]. Recently, van Lierop et al. investigated delayed centrifugation, centrifugation temperature, and delayed storage, as well as tube types in patients with multiple sclerosis (pwMS) and healthy controls (HC) [25]. They observed that sodium citrate samples had substantially lower NF-L concentrations than serum and EDTA plasma. Furthermore, they found that delayed centrifugation (delay before centrifugation between 6 and 24 h) resulted in increased variation. Centrifugation temperature and delayed storage did not lead to relevant changes in NF-L concentration. They observed similar results in samples from pwMS and HC for all experiments. Several studies have investigated the stability of NF-L in serum and plasma using
multiple freeze-thaw steps and prolonged exposure to room temperature. There is agreement that serum and plasma NFL remains stable in both freeze-thaw and prolonged exposure experiments [24, 26–28].

**Sampling time**

Blood sampling time is crucial for interpreting the result of many clinical chemistry tests. However, there is no clear consensus about fasting status or blood sampling time in regards to NFL although some publications provide recommendations for other markers of relevance to CNS diseases [29, 30]. Diurnal variation needs to be resolved, especially in an NFL-specific manner, even if longitudinal CSF and blood sampling are challenging, especially for paired CSF and blood samples. Benedict et al. have shown that morning plasma NFL concentrations are more than 10% higher than evening concentrations, although their results are preliminary [31].

Additionally, Hviid et al. investigated the biological variation of NFL in 33 healthy individuals in a total of 184 blood samples for three days in a row [32]. They found a minute within-subject variation with neglectable day-to-day variation without semidiurnal changes. Even if they did not have paired CSF-serum samples from the same individual, they concluded that NFL is linked to tight homeostatic regulation with no or negligible semidiurnal and day-to-day variation.

We conclude that NFL is relatively robust to preanalytical variation, that the biomarker can be measured in CSF, serum, and EDTA plasma but with different absolute concentrations, necessitating sample matrix-specific reference limits, that the biomarker results are not sensitive to repeated freeze-thawing or storage temperature, but that delayed centrifugation of samples may result in higher

---

**Figure 1:** A basic checklist for what pre-analytical factors to examine for novel biomarkers.
variation. More studies are needed on the effects of fasting, and if there is any clinically meaningful diurnal variation; in the meantime, we recommend that plasma or serum NfL is measured in samples collected in the morning fasting. A suggested SOP for pre-analytical sample handling for NfL quantification is presented in Figure 2.

**Analytical phase – general information, future directions**

**Analytical methods**

The blood concentration of NfL is approximately 50 times lower than in CSF, owing to the proximity of the CSF to neuronal tissues, as well as dilution of NfL in the large blood volume [8]. Therefore, it is critical to choose the right method taking into account the sample matrix and the analytical sensitivity of the method (its lower limit of detection [LLoD] and quantification [LLoQ]). Until now, researchers have investigated NfL in different matrices using various generations of immunoassays. These are categorized as first (western blot [WB] [33]), second (enzyme-linked immunosorbent assay [7, 34]), third (electrochemiluminescence [ECL] [16]), and fourth (Single molecule array [Simoa]) generation immunoassays [35]. There are also emerging mass spectrometry-based methods for NfL quantification, but they are not ready for clinical use yet, although they may have value as candidate reference methods for the protein [36].

**Advantages-disadvantages of the analytical methods used for NfL measurement**

First-generation immunoassays are no longer used to quantify NfL because of their limited sensitivity, as well as...
their semi-quantitative nature. However, WB may be used in research settings to characterise NfL. Sandwich ELISA for NfL quantification has been commercially available since 2003. ELISA has been used for NfL measurements in both the CSF and blood; nevertheless, the original ELISA was restricted to CSF, due to its limited sensitivity for measuring low concentrations of NfL in blood [16]. Although several validation studies have been performed using CSF ELISA [37], there are no certified reference methods or materials available; hence, further standardization work is required. Currently, several companies have improved ELISAs that may work for serum and plasma samples and not only CSF; their asserted sensitivity should be considered when blood NfL measurement is planned. Another method, ECL, depends on the measurement of the electrochemiluminescent signal formed by the binding of specific monoclonal antibodies conjugated with electrochemiluminescent labels that emit light when electricity is applied to the measurement wells. ECL was presented in 2013 for the measurement of NfL in the blood with high sensitivity [38]. Despite improved sensitivity, some healthy control samples were still not measurable owing to their low concentration in the blood [39]. The fourth-generation immunoassay, Simoa, is based on anti-NfL antibody-conjugated magnetic beads that can be pulled down together with the target analyte and enzyme-labeled detector antibody into microwells to allow for the counting of individual NfL molecules [40]. Owing to this compartmentalisation, it is possible to reliably quantify NfL in blood at very low concentrations. Additionally, several companies have used the NfL antibodies produced by Uman Diagnostics to develop new assays on other platforms, such as Ella™ (a microfluidic platform) [41]. Recently, a highly sensitive prototype NfL assay has been developed and evaluated on ADVIA Centaur® XP immunoassay system by Siemens Healthineers [20]. Similar NfL assays on fully automated platforms such as Cobas (Roche Diagnostics) and Lumipulse (Fujirebio) are currently underway. Once available, further method comparison work is needed.

The choice of calibrators

Researchers have used different calibrators when measuring NfL in blood (serum or plasma) and CSF. Early studies on Simoa were primarily performed using calibrators made from NfL purified from bovine brain [16, 40]. Nevertheless, current studies have mainly used the Quanterix NF-light™ assay kit, in which recombinant human NfL calibrators are used. Since the bovine and recombinant calibrators produce notably different signal outputs in the assays, the final NfL concentrations from samples can be different depending on the which calibrators used. Hendricks et al. indicated that results obtained from the two calibrators (bovine and recombinant human calibrators) correlated well [42]. They also recommended a conversion factor of 5 when using bovine vs. recombinant human NfL calibrator, i.e., concentrations obtained using recombinant human NfL calibrator should be multiplied by 5 to make them comparable to results obtained with the bovine calibrator [42]. Laboratory researchers and clinicians should consider which assay has been used because of the probability of positive bias (5:1) of the home-brew assay relative to the commercial NF-light™ assay [42, 43]. Therefore, this discrepancy highlights the need for assay standardization to ensure consistent reproducibility and standardization protocols between clinical laboratories. The same should also be considered for different platforms, and meticulous work is needed to confirm inter-platform equivalence.

IFCC – WG-CSF proteins

Commutability is defined as the equivalence of the mathematical association amongst the results of divergent measurement procedures for a reference material and for representative samples of the type intended to be measured according to the CLSI EP30-A document [44]. Principally, a regression protocol with 95% prediction interval is used to assess commutability of reference materials to show inter-assay properties. Basically, we can define as property of a reference material in such a manner that the same amount of measurand in the reference material and in clinical samples yields the same measurement response in different measurement procedures. Non-commutable calibrators can cause unreliable and inconsistent clinical sample results. Besides, a non-commutable calibrator breaks the traceability chain. If the traceability chain is broken, different working calibrators cause different results from different end-user in vitro diagnostics (IVD) medical devices. Matrix-based secondary calibrators are required to be commutable with clinical samples to achieve metrological traceability of results from a clinical laboratory measurement procedure to higher order references [45, 46] (Figure 3). In order to achieve commutable NfL results, several studies are ongoing now.

Under the auspices of IFCC, the working group on CSF proteins (WG-CSF) is working on NfL standardisation as part of the ‘Development of reference method procedures and reference material for neurofilament light polypeptides in CSF and blood’ [47]. For this purpose, the first commutability study for the measurement of NfL in plasma and serum was accomplished by measuring NfL in patient samples using
five assay platforms in seven laboratories, with high correlations across the methods but slightly different absolute concentrations [48]. The WG-CSF proteins are to undergo a second commutability study for different purposes, for example, to test the stability of candidate certified reference materials (CRM) and the performance of pooled samples for a CRM.

Quality controls and recommended directions

The basis of the “risk-based statistical quality control” is defined as follows by Westgard et al.: Defining the quality specifications for the test; selecting appropriate control materials and levels; determining the stable (in control) performance of the measurement procedure; identifying candidate quality control strategies; specifying desirable goals for the QC performance characteristics; selecting a quality control strategy (control rules, number of control measurements), whose predicted performance meets or exceeds the quality control performance goals [49]. There are no pre-defined quality control checklists for NfL measurement in any guidelines; therefore, to support measurement validity, researchers and laboratory professionals need to consider some checkpoints while measuring NfL. Some of the essential checkpoints that should be scrutinized before and during the analysis are the measurement of samples, quality controls, and calibrators (preferably in duplicate) [50]. A flowchart for NfL quality control is presented in Figure 4.

Post-analytical phase

The post-analytical phase is as crucial as the pre-analytical and analytical phases. In clinical laboratories, we provide numbers, add comments to the results, and think of them in relation to other laboratory parameters and clinical data. The biomarker result is related to reference limits or cut-points for optimal disease detection.
Age-related reference values for NfL

For NfL, the most important parameter that influences the normal reference limits is age. This can be dealt with through the development of age-related normal reference limits established in individuals free from neurological diseases that may affect NfL concentration. Simrén et al. [7] recently published age-related reference values for plasma NfL. In this paper, the authors analyzed plasma samples from a large number of neurologically healthy participants across the lifespan (aged 5–90 years) to provide reliable age-partitioned reference limits for clinical use. Specific cut-off values in five different age categories were defined to reflect the effect of age on concentrations of plasma NfL [7]. Another and more novel approach is the use of NfL percentiles and Z scores. Benkert et al. [51] generated age- and body mass index (BMI)-adjusted serum NfL concentrations by obtaining 10,133 serum samples using percentiles and Z scores to correct for confounding factors to differentiate pathological from physiological levels of serum NfL. Using Z-scores may be an appropriate technique for coping with age-related increases in serum NfL [51].

Essential to both approaches is in-house data that verifies stability of the measurements in relation to the material in which the reference limits or cut-points were established. In several papers, the authors established reference intervals using samples from different cohorts [7, 24, 52, 53]. Although relative changes in the concentrations of NfL between relevant groups were similar...
between studies, we may not directly compare absolute NfL concentrations from one study to others owing to several factors. Firstly, there may be potential concentration differences between different assays used in the studies (there is no available certified reference material for assay standardization yet). Secondly, since all available NfL assays are research-grade, there may be lot-to-lot variations in the absolute concentrations also when the same assay is used. The matrix chosen for the NfL measurement should also be considered. For example, concentrations of NfL in serum have been shown to be ∼10% higher than those determined in plasma samples [24]. Furthermore, differences in interval age ranges used for stratifying cohorts may hinder direct comparison and yield variations between studies. Most data suggest that CSF NfL is a bit more sensitive than plasma/serum NfL to detect neuroaxonal injury or degeneration [54]; hence, a normal plasma or serum NfL concentration does not exclude neurodegenerative disease. In contrast, an abnormal test result suggests that there is axonal injury in need of further examination.

Other possible confounders of NfL concentration

Factors in addition to age that may influence plasma NfL concentration are kidney disease (higher plasma NfL concentration in patients with kidney disease [55, 56]) and body composition (lower plasma NfL in patients with high BMI [56]). An additional approach in regards to NfL interpretation is to examine longitudinal change of the biomarker; however, the delta that constitutes a clinically meaningful change in longitudinal/serial measurements needs to be defined in different age groups and disease contexts. This is also important in regards to the use of NfL to monitor the response to disease-modifying treatments against, e.g., multiple sclerosis [43, 50, 57]. In the future, it may be possible to design an auto-verification system for release of NfL results, which may reduce the burden on clinical laboratories.

Conclusions

CSF and plasma NfL are clearly useful as general markers of neuroaxonal injury across CNS and peripheral nervous system (PNS) diseases. The availability of the biomarker as a simple blood test speaks for its more widespread adoption in clinical laboratory practice. During recent years, we have learnt a lot about pre-analytical factors that may confound the results and can conclude that NfL is a relatively robust biomarker. Its age-related increase makes it less useful in older individuals, and the fact that it is not specific to any particular disease but rather a general marker of neuroaxonal injury makes it important to find the underlying cause of any abnormal result. Finally, a normal plasma or serum NfL result does not exclude neurodegenerative diseases or other diseases that may result in low-grade neuroaxonal injury. From our point of view, the marker is ready for more widespread implementation in clinical laboratory practice. However, we must also remember that we do not fully understand what regulates the concentration of this biomarker in biofluids – passive release from injured axons, increased expression and release of the protein, and/or impaired clearance of the protein, e.g., by microglia, or a combination thereof. Whilst examining these details, clinicians and medical laboratory professionals should work with each other by exchanging their ideas, and laboratory professionals should devote more time to this work in translational research projects [58, 59].

Research funding: BA was supported by grant from IUBMB–FEBS–PABMB Young Scientists’ Forum (YSF), IUBMB–FEBS–PABMB Congress 2022, 9th – 14th July 2022, Lisbon, Portugal. HZ is a Wallenberg Scholar supported by grants from the Swedish Research Council (#2018-02532), the European Union’s Horizon Europe research and innovation programme under grant agreement No 101053962, Swedish State Support for Clinical Research (#ALFGBG-71320), the Alzheimer Drug Discovery Foundation (ADDF), USA (#201809-2016862), the AD Strategic Fund and the Alzheimer’s Association (#ADSF-21-831376-C, #ADSF-21-831381-C, and #ADSF-21-831377-C), the Bluefield Project, the Olav Thon Foundation, the Erling-Persson Family Foundation, Stiftelsen för Gamla Tjänarinnor, Hjärnfonden, Sweden (#FO2022-0270), the European Union’s Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 860197 (MIRIADE), the European Union Joint Programme – Neurodegenerative Disease Research (JPND2021-00694), and the UK Dementia Research Institute at UCL (UKDRI-1003).

Author contributions: All authors have accepted responsibility for the entire content of this manuscript and approved its submission.

Competing interests: HZ has served at scientific advisory boards and/or as a consultant for Abbvie, Acumen, Alector, ALZPath, Annexon, Apellis, Artery Therapeutics, AZTherapies, CogRx, Denali, Eisai, Nervgen, Novo Nordisk, Passage Bio, Pinteon Therapeutics, Red Abbey Labs, reMYND, Roche, Samumed, Siemens Healthineers, Triplet Therapeutics, and Wave, has given lectures in symposia
sponsored by Cellectron, Fujirebio, Alzecure, Biogen, and Roche, and is a co-founder of Brain Biomarker Solutions in Gothenburg AB (BBS), which is a part of the GU Ventures Incubator Program (outside submitted work). BA reports no conflicts of interest.

**Informed consent:** Not applicable.

**Ethical approval:** Not applicable.

## References