Correlation between CSF and blood neurofilament light chain protein: a systematic review and meta-analysis

Jasmini Alagaratnam, Sophia von Widekind, Davide De Francesco, Jonathan Underwood, Paul Edison, Alan Winston, Henrik Zetterberg, Sarah Fidler

ABSTRACT

Objective To assess the overall pooled correlation coefficient estimate between cerebrospinal fluid (CSF) and blood neurofilament light (NFL) protein.

Methods We searched Medline, Embase and Web of Science for published articles, from their inception to 9 July 2019, according to Preferred Reporting Items for Systematic Reviews and Meta-analyses guidelines. Studies reporting the correlation between CSF and blood NFL in humans were included. We conducted a random-effects meta-analysis to calculate the overall pooled correlation coefficient estimate, accounting for correlation technique and assay used. Heterogeneity was assessed using the I² statistic test. In sensitivity analyses, we calculated the pooled correlation coefficient estimate according to blood NFL assay: single-molecule array digital immunoassay (Simoa), electrochemiluminescence (ECL) assay or ELISA.

Results Data were extracted from 36 articles, including 3961 paired CSF and blood NFL samples. Overall, 26/36 studies measured blood NFL using Simoa, 8/36 ECL, 1/36 ELISA and 1 study reported all three assay results. The overall meta-analysis demonstrated that the pooled correlation coefficient estimate for CSF and blood NFL was r=0.72. Heterogeneity was significant: I²=83%, p<0.01. In sensitivity analyses, the pooled correlation coefficient was similar for studies measuring blood NFL using Simoa and ECL (r=0.69 and r=0.68, respectively) but weaker for ELISA (r=0.35).

Conclusion Moderate correlations are demonstrated between CSF and blood NFL, especially when blood NFL was measured using Simoa and ECL. Given its high analytical sensitivity, Simoa is the preferred assay for measuring NFL, especially at low or physiological concentrations, and this meta-analysis supports its use as the current most advanced surrogate measure of CSF NFL.

PROSPERO registration number CRD42019140469

INTRODUCTION

Cerebrospinal fluid (CSF) neurofilament light (NFL) chain protein is well recognised as a sensitive and dynamic biomarker of active central nervous system (CNS) neuro-axonal injury.1–3 The neurofilament complex is predominantly located in the neuronal cytoplasm and they provide structural stability to neurons and enable radial growth of axons.2,3 Concentrations of CSF NFL rise proportionally to the degree of neuroaxonal injury in a variety of neurological conditions including neurodegenerative, inflammatory, vascular and traumatic diseases.4–7

However, the invasive nature of CSF collection via lumbar punctures limits the widespread use of CSF NFL. The ELISA routinely used to measure that CSF NFL is not recommended for blood NFL measurement (usually 50–100 times lower than CSF NFL concentrations), due to its limited sensitivity. Electrochemiluminescence (ECL)-based NFL assays have improved analytical sensitivity,8,9 but the novel ultrasensitive single-molecule array (Simoa) digital immunoassay is 126-fold and 25-fold more sensitive than ELISA and ECL assays, respectively, for quantification of NFL.10 The manifold higher analytical sensitivity with the Simoa assay for NFL measurement enables reliable blood NFL measurement in disease and physiological conditions,11 12 while avoiding the need for CSF collection and allowing more frequent measurement given that blood is easier to obtain.

Individual studies have reported the correlation coefficients between CSF NFL and blood NFL in several, discrete neurological conditions, but the pooled overall correlation coefficient estimate has not been established. If blood NFL is to be used as a reliable surrogate marker of CSF NFL, then the overall estimated correlation between CSF and blood NFL needs to be determined.

The aim of this systematic review and meta-analysis was to determine the overall pooled correlation coefficient estimate between CSF and blood NFL in human studies. Given the lower analytical sensitivity of ELISA compared with ECL and Simoa NFL assays, which is of particular relevance when measuring blood NFL concentrations, we also assessed the
pooled correlation coefficient estimate between CSF and blood NFL, in studies that measured blood NFL using the Simoa or ECL assays only. In subanalyses, we stratified the pooled correlation coefficient estimates by blood NFL assay, and according to whether both CSF and blood NFL concentrations were measured using the Simoa assay or whether CSF NFL was measured using ELISA or ECL assays while blood NFL was measured using Simoa. Additionally, we stratified the pooled correlation coefficient estimates by plasma versus serum NFL, conditions which purely affect the CNS versus conditions that affect both the CNS and peripheral nervous system (PNS) versus control participants and statistical correlation technique used.

METHODS
Standard protocol approvals, registrations and patient consents
This systematic review and meta-analysis was conducted in accordance with Preferred Reporting Items for Systematic Reviews and Meta-analyses guidelines13–15 and is reported in compliance with the Meta-Analysis of Observational Studies in Epidemiology proposal. A protocol was registered and approved in the International Prospective Register of Systematic Reviews (PROSPERO).15 16 We searched publicly available published studies, and institutional research ethics board approval and patient consent were not required for this systematic review.

Data sources and search strategy
We systematically searched Ovid MEDLINE, Ovid Embase and Web of Science electronic databases for eligible published articles from their inception to 9 July 2019. The following search terms were used: [(cerebrospinal fluid or spinal fluid or CSF) and (neurofilament* light or neurofilament* light or neurofilament* light or neurofilament* light or NFL or NFLs)] and [(plasma or blood or peripheral or serum) and (neurofilament* light or neurofilament* light or NFL or NFLs)]. Grey literature sources were not accessed.

Inclusion and exclusion criteria
Studies published in English were included if the correlation coefficient between paired CSF and blood (plasma or serum) NFL in human participants was reported. Studies were excluded for the following reasons: (1) duplicate articles (where the same article was retrieved more than once during the electronic database searches), (2) CSF and blood samples retrieved at autopsy, (3) nonoriginal research (review articles, letters in response to previous articles) and (4) abstracts and conference proceedings.

Data collection
Retrieved articles were imported into Covidence, an online primary screening and data extraction tool. Two authors (JA and SvW) independently screened the titles and abstracts of all studies retrieved from the database search to identify potentially eligible studies; the two authors (JA and SvW) then independently appraised the potentially eligible full-text articles against the eligibility criteria to determine final inclusion into the systematic review and meta-analysis. Discrepancies about decisions on study inclusion were resolved by the senior author (SF). Where additional clarification was required such as correlation technique used and sample size, the corresponding author of the study was contacted and details about the specific missing information were requested.

Data extraction
Data from the full-text articles were independently extracted and checked for accuracy by two authors (JA and SvW) and imported into a database on Microsoft Excel. The following information was extracted: first author’s surname, year of publication, article title, location where study took place, study design, disease process(es) being investigated, study population, number of paired CSF and blood samples measured for NFL, CSF and blood NFL assays used, plasma or serum samples analysed for NFL, statistical correlation technique used and correlation coefficient value. Data extraction is summarised in online supplemental Table 1. Any disagreements during data extraction were resolved by the senior author (SF).

Assessment of quality
Quality assessment was performed at the study level by two authors (JA and SvW) using the National Institutes of Health: National Heart Lung and Blood Institute’s study quality assessment tools.17 Each item was scored as (✓=1 point, X or ?=0 points). Any study that scored above 6 points of the 14-point criteria was considered ‘good’ quality or ‘low risk of bias’. Studies scoring between 5 and 6 points were considered ‘fair’ quality or ‘moderate risk of bias’ and studies scoring below 5 points were considered ‘poor’ quality or ‘high risk of bias’.

Statistical analysis
Meta-analyses were performed using the statistical software R V.3.6.1 and the meta package. The pooled correlation coefficient estimate and 95% CI in the overall analysis were calculated using a random-effects meta-analysis of correlations based on Fisher’s Z-transformation, incorporating the heterogeneity between studies due to the different correlation techniques and blood NFL assays used. Heterogeneity and between-study variance were assessed using the I² statistics and τ² (Sidik-Jonkmann estimator). The rank correlation test of funnel plot asymmetry was used to assess for publication bias. A p value <0.05 was considered statistically significant throughout.

In prespecified sensitivity analyses, pooled correlation coefficient estimates were stratified according to: (1) studies measuring blood NFL using ECL and Simoa assays only, (2) blood NFL assay (Simoa, ECL or ELISA), (3) whether both CSF and blood NFL concentrations were measured using Simoa or CSF NFL was measured using ELISA or ECL and blood NFL was measured using the Simoa assay, (4) whether plasma or serum NFL was...
measured, (5) whether the neurological condition being investigated only affected the CNS or affected both the CNS and PNS and (6) correlation coefficient technique used (Spearman’s rank or Pearson’s correlation coefficient, where unspecified, Pearson’s correlation coefficient was imputed). Pooled correlation coefficient estimates and 95% CI in these subgroup analyses were calculated using a fixed-effects model.

**Data availability**

The data sets used and analysed in this study including those not published within the article can be shared with other qualified investigators on reasonable request made to the corresponding author, in accordance with International Committee of Medical Journal Editors requirements.

**RESULTS**

**Study selection**

The database searches identified 1058 unique articles (figure 1). Titles and abstracts were screened against the eligibility criteria and 826 articles were excluded (figure 1). The remaining 232 articles underwent full-text review and 36 articles \(^8\) \(^{10–12}\) \(^{18–49}\) were deemed eligible for inclusion into the meta-analysis (figure 1).

**Study characteristics**

Data were extracted from the 36 articles (online supplemental table 1) and included 3961 unique paired CSF and blood NfL measurements. Studies were from Europe (n=26), North America (n=5), China (n=1) and four studies recruited in multiple sites internationally. Online supplemental table 1 summarises the characteristics of the individual studies included.

Blood NfL was measured using the Simoa assay in 26/36 studies, ECL assay in 8/36 studies, ELISA in 1/36 study; one study reported results using all three assays (Simoa, ECL and ELISA) for each blood NfL measurement. Most studies measured CSF NfL using ELISA (25/36), while 6/36 studies measured CSF NfL using Simoa, 6/36 studies using ECL and 1 study reported CSF NfL measurements using all three assays (Simoa, ECL and ELISA). In total, 23/36 studies used serum NfL when calculating the correlation coefficient between CSF and blood NfL, 12/36 studies used plasma NfL and 1 study combined the results from serum and plasma NfL measurements. In total, 21 unique correlation coefficients between CSF and blood NfL were reported in conditions affecting the CNS only, 17 unique correlation coefficients were reported in conditions affecting the CNS and PNS and 7 correlation coefficients were reported in cohorts of control participants. Overall, 22/36 studies reported Spearman’s rank correlation, 13/36 studies reported Pearson’s correlation coefficient; correlation technique was unspecified in one study.

**Meta-analysis results**

Using a random-effects model, the overall pooled correlation coefficient estimate for CSF and blood NfL across all 36 eligible studies was \(r=0.723 (95\% \text{ CI} 0.540 \text{ to } 0.840)\) (figure 2A). Heterogeneity was significant with an \(I^2\) result of 83% and \(\tau^2\) (Sidik-Jonkman estimator) of 0.072, \(p<0.01\), thus stratified analyses were performed. The rank correlation test of funnel plot asymmetry showed no asymmetry (p=0.53), and, thus, no obvious publication bias (figure 2B).

In sensitivity analyses, the pooled correlation coefficient estimates for CSF and blood NfL for studies which measured blood NfL using Simoa or ECL assays only (n=3848), excluding the two studies which measured blood NfL using ELISA, was \(r=0.688 (95\% \text{ CI} 0.671 \text{ to } 0.705)\) (figure 3A). The pooled correlation coefficient was similar when blood NfL was measured using Simoa (n=3117) and ECL (n=731) assays \((r=0.689 (95\% \text{ CI} 0.670 \text{ to } 0.708)\) and \(r=0.684 (95\% \text{ CI} 0.642 \text{ to } 0.722)\), respectively) (figure 3B,C), but weaker with ELISA (n=113) \((r=0.354 (95\% \text{ CI} 0.176 \text{ to } 0.510))\) (figure 3D). The pooled correlation estimate was similar when stratified according to studies which measured both CSF and blood NfL using Simoa (n=437) and studies where blood NfL was measured...
using Simoa and CSF NfL was measured using ELISA or ECL (n=2600), r=0.712 (95% CI 0.661 to 0.756) (figure 3E) and r=0.674 (95% CI 0.652 to 0.695) (figure 3F), respectively.

With regards to studies measuring blood NfL in serum versus plasma, studies using serum NfL (n=2335) had a pooled correlation coefficient estimate of r=0.658 (95% CI 0.634 to 0.681) (figure 3G), and studies which used plasma NfL (n=1531) had a pooled estimate of r=0.706 (95% CI 0.680 to 0.731) (figure 3H).

Studies including participants with clinical disorders that affected the CNS only (n=2045) had a pooled correlation coefficient estimate of r=0.642 (95% CI 0.615 to 0.667) (figure 4A). Studies including participants with clinical disorders that affected both the CNS and PNS (n=1577) had a higher overall pooled correlation coefficient estimate of r=0.745 (95% CI 0.721, 0.766) (figure 4B) and the control participants (n=508) had the lowest pooled correlation coefficient estimate at r=0.552 (95% CI 0.466 to 0.627) (figure 4C).

When stratified according to correlation technique used, studies quoting Spearman’s rank correlation coefficient (n=3085) had a pooled correlation coefficient estimate of r=0.642 (95% CI 0.620 to 0.663) (figure 4D), while the pooled estimate was higher for studies quoting Pearson’s correlation (n=876), with r=0.794 (95% CI 0.768 to 0.818) (figure 4E).

Quality assessment of included studies
The results of the quality assessment of the included studies are summarised in online supplemental table 2. Overall, 31 studies were classified as ‘good’ quality and five as ‘fair’ quality.

DISCUSSION
Our systematic review and meta-analysis demonstrates that the overall pooled correlation coefficient estimate between CSF and blood NfL across pathologies is moderately strong, according to the Chan definition. While this is reassuring, it is important to note that blood NfL does not correlate perfectly with CSF NfL and additional factors may need to be taken into consideration when interpreting blood NfL results.

Only two studies measured blood NfL using ELISA, which has a low analytical sensitivity for measuring NfL concentration. When these two studies were excluded, the pooled correlation coefficient estimate between CSF and blood NfL across pathologies remained moderately strong, which may reflect the small number of unique paired CSF and blood NfL measurements contributed by these two studies to the overall pooled correlation coefficient estimate in this meta-analysis. When stratified according to blood NfL assay used, moderately strong correlations were demonstrated between CSF and blood NfL in studies using Simoa and ECL blood NfL assays, but the correlations in studies using ELISA to measure blood NfL.
Figure 3  
(A) Forest plot of the pooled correlation coefficients for CSF and blood neurofilament light protein, in studies that used the Simoa and ECL assays only to measure blood NfL concentration. Forest plot of the summary correlation coefficients with corresponding 95% CIs for the correlation between CSF and blood neurofilament light protein from studies that used the Simoa digital immunoassay to measure both CSF and blood NfL concentrations. (B) Forest plot of the pooled correlation coefficients for studies using Simoa to measure blood NfL. Forest plot of the summary correlation coefficients with corresponding 95% CIs for the correlation between CSF and blood neurofilament light protein from studies that used the Simoa digital immunoassay to measure blood NfL. (C) Forest plot of the pooled correlation coefficients for studies using the ECL assay to measure blood NfL. Forest plot of the summary correlation coefficients with corresponding 95% CIs for the correlation between CSF and blood neurofilament light protein from studies using the ECL assay to measure blood NfL. (D) Forest plot of the pooled correlation coefficients for studies using the ELISA assay to measure blood NfL. Forest plot of the summary correlation coefficients with corresponding 95% CIs for the correlation between CSF and blood neurofilament light protein from studies using the ELISA assay to measure blood NfL. (E) Forest plot of the pooled correlation coefficients for studies using the Simoa assay to measure both CSF and blood NfL concentrations. Forest plot of the summary correlation coefficients with corresponding 95% CIs for the correlation between CSF and blood neurofilament light protein from studies using the Simoa digital immunoassay to measure blood NfL. (F) Forest plot of the pooled correlation coefficients for studies using the ELISA or ECL assays to measure CSF NfL and Simoa assay to measure blood NfL. Forest plot of the summary correlation coefficients with corresponding 95% CIs for the correlation between CSF and blood neurofilament light protein from studies using the Simoa digital immunoassay to measure blood NfL. (G) Forest plot of the pooled correlation coefficients for studies that measured serum NfL. Forest plot of the summary correlation coefficients with corresponding 95% CIs for the correlation between CSF and blood neurofilament light protein from studies which measured serum NfL. (H) Forest plot of the pooled correlation coefficients for studies that measured plasma NfL. Forest plot of the summary correlation coefficients with corresponding 95% CIs for the correlation between CSF and blood neurofilament light protein from studies which measured plasma NfL. CSF, cerebrospinal fluid; NfL, neurofilament light chain protein; ECL, electrochemiluminescence.
were much lower. The pooled correlation estimate was higher in studies that measured both CSF and blood NfL using Simoa compared with studies that measured blood NfL using Simoa and CSF NfL using ELISA or ECL, suggesting that in the setting of these studies, measuring CSF and blood NfL using Simoa improved the association in NfL between the two compartments. Simoa is the current preferred blood NfL assay, especially at low or physiological concentrations due to its high analytical sensitivity (low detection limit), and this meta-analysis supports the use of blood NfL measured using Simoa as the current most advanced surrogate measure of CSF NfL.

The correlation coefficients were similarly moderately strong when stratified according to studies which measured plasma versus serum NfL, in keeping with published literature suggesting the lack of difference in NfL concentration when measured in these two matrices.

When stratified according to underlying condition being studied (purely CNS conditions, conditions with CNS and PNS components and control participants), the pooled correlation coefficient estimates were highest in participants with CNS and PNS disease and lowest in the control participants. The NfL concentration range was lower in control participants and closer to the analytical sensitivity of the NfL assays employed, thus, more variable, resulting in a lower correlation between CSF and blood NfL. This suggests that blood NfL is a better surrogate marker of CSF NfL at higher CSF NfL concentration ranges. Additionally, participants with disorders affecting the CNS may have a more disrupted blood–brain barrier, and, thus, leak more NfL from the CSF into the blood, compared with control participants, who are more likely to have intact blood–brain barriers, and, thus, leak less CSF NfL into the bloodstream. The pooled correlation coefficient between CSF and blood NfL is higher in studies reporting Pearson’s correlation compared with those reporting Spearman’s rank correlation, possibly due to the presence of outliers or to non-normality of NfL concentrations.

Strengths of our review are the high methodological standards used to conduct the systematic review, and the inclusion of potential confounders in sensitivity analyses. Limitations include publication bias, which may cause an overestimation of the pooled correlation coefficient estimates. Most studies enrolled participants in Western Europe and North America, and it is unknown whether our results can be extrapolated to individuals globally. We included publications in English language only as part of our search strategy.

**Figure 4** (A) Forest plot of the pooled correlation coefficients for studies involving CNS-only disorders. Forest plot of the summary correlation coefficients with corresponding 95% CIs for the correlation between CSF and blood neurofilament light protein from studies including participants with neurological disorders that only affect the central nervous system only. (B) Forest plot of the pooled correlation coefficients for studies involving CNS and PNS disorders. Forest plot of the summary correlation coefficients with corresponding 95% CIs for the correlation between CSF and blood neurofilament light protein from studies including participants with neurological disorders that affect both the CNS and PNS. (C) Forest plot of the pooled correlation coefficients from control participants. Forest plot of the summary correlation coefficients with corresponding 95% CIs for the correlation between CSF and blood neurofilament light protein from control participants. (D) Forest plot of the pooled correlation coefficients from studies reporting Spearman’s rank correlation coefficient. Forest plot of the summary correlation coefficients with corresponding 95% CIs for the correlation between CSF and blood neurofilament light protein from studies reporting Spearman’s rank correlation coefficient. (E) Forest plot of the pooled correlation coefficients from studies reporting Pearson’s correlation coefficient. Forest plot of the summary correlation coefficients with corresponding 95% CIs for the correlation between CSF and blood neurofilament light protein from studies reporting Pearson’s correlation coefficient.

CNS, central nervous system; CSF, cerebrospinal fluid; NfL, neurofilament light chain protein; PNS, peripheral nervous system.
and may have excluded studies reporting the correlation coefficient between CSF and plasma NfL that were not in English language, which may affect the pooled correlation coefficient estimates. The number of blood samples measured for NfL using ELISA was much smaller (n=113) compared with Simoa (n=3117) and ECL (n=731), which may contribute to the much wider 95% CI for the pooled correlation coefficient estimate between CSF and blood NfL in samples using ELISA technique. However, the most likely explanation for the variable results is that blood NfL concentration measured by ELISA simply reflects noise, as the analytical sensitivity of the assay is insufficient to quantify NfL in blood reliably.

The impact of heterogeneous factors that may influence NfL measurement and interpretation such as unicentric versus multicentric studies, cross-sectional versus longitudinal samples and the duration between CSF and blood sampling were not explored in this review, due to the data not being readily available from the publications. Data on preanalytical factors that may affect NfL measurements were also not consistently available, thus, it could not be systematically assessed between the studies. Preanalytical factors to consider include different sampling methods, duration of NfL stability at room temperature and number of freeze-thaw cycles prior to NfL measurement.

Data were not routinely accessible for the following factors that may increase plasma NfL independently of CSF NfL and affect the correlation between CSF and blood NfL. There is evidence to suggest an association between increased blood–brain barrier permeability and increased blood NfL concentration, but other studies have not demonstrated this relationship. Other factors that may be associated with increased blood NfL include lower body mass index (possibly due to decreased blood volume), pregnancy (possibly due to the developing fetal brain), peripheral nerve injury, concomitant use of neurotoxic drugs and lower estimated glomerular filtration rate.

Using rigorous systematic review and meta-analysis, we report an overall moderately strong correlation between CSF and blood NfL. Until now, the strength of the correlation between CSF and blood NfL has been questionable due to the uncertainty of agreement between the studies. Our findings support the use of blood NfL measurement as a promising surrogate marker of CSF NfL. Additional studies are warranted to validate the blood NfL assay and to assess how blood NfL performs in clinical and research settings.

Author affiliations
1Department of Infectious Disease, Imperial College London, London, UK
2Department of Genitourinary Medicine & HIV, Imperial College Healthcare NHS Trust, London, UK
3Faculty of Medicine, Imperial College London, London, UK
4Institute for Global Health, University College London, London, UK
5Division of Infection and Immunity, Cardiff University, Cardiff, UK
6Department of Brain Sciences, Imperial College London, London, UK
7Department of Neurodegenerative Disease, UCL Institute of Neurology, London, UK
8Department of Psychiatry and Neurochemistry, University of Gothenburg Sahlgrenska Academy, Goteborg, Sweden

Contributors JA, AW and SF conceptualised the idea for the study, JA, SW, AW, HZ and SF designed the study, JA, SW, JU, PE, AW, HZ and DDF formulated the data analysis plan; JA and SW performed the data collection, SF resolved conflicts during data collection, DDF performed the statistical analysis, JA, SW and DDF interpreted the results, JA drafted the first version of the manuscript, all authors reviewed and edited the manuscript for intellectual content.

Funding Imperial College National Institute of Health Research (NIHR) Biomedical Research Centre (BRC) provided infrastructural support.

Competing interests JA has received financial support to attend scientific conferences from MSD and Gilead Sciences. JU has received honoraria for preparation of educational materials and has served on an advisory board for Gilead Sciences. PE has received grants from the Medical Research Council, Alzheimer’s Research UK, Alzheimer’s Drug Discovery Foundation, Alzheimer’s Association, Alzheimer’s Society, Novo Nordisk®, Life Molecular Imaging, GE Healthcare, Eli Lilly and Company, Novartis International AG and NIHR Imperial Biomedical Research Centre (BRC), PE was also a consultant to Pfizer, is now a consultant to Novo Nordisk, and serves on the advisory board for Novo Nordisk. PE has received consultancy and speaker fees from Piramal Life Sciences, Pfizer and Novo Nordisk. AW has received honoraria or research grants on behalf of Imperial College London or been a consultant or investigator in clinical trials sponsored by Bristol-Myers Squibb, Gilead Sciences, GlaxoSmithKline, Janssen-Cilag, Roche and ViVi Healthcare. HZ has served at scientific advisory boards for Denali, Roche Diagnostics, Wave, Samumed and CogRx, has given lectures in symposia sponsored by Fujirebio, Alzecure and Biogen, and is a co-founder of Brain Biomarker Solutions in Gothenburg AB, a GI Ventures-based platform company at the University of Gothenburg (outside submitted work). SF received funding and research grants from the National Institutes of Health (NIH), Bill & Melinda Gates Foundation (BMGF) and Medical Research Council (MRC).

Patient consent for publication Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/.

ORCID iDs
Jasmini Alagaratnam http://orcid.org/0000-0003-2256-423X
Paul Edison http://orcid.org/0000-0002-6551-2002
Henrik Zetterberg http://orcid.org/0000-0003-3930-4354

REFERENCES
4 Rosengren LE, Karlsson JE, Karlsson JO, et al. Patients with amyotrophic lateral sclerosis and other neurodegenerative diseases...
have increased levels of neurofilament protein in CSF. J Neurochem 1996;67
5 Hall S, Ohrlert A, Constantinescu R, et al. Accuracy of a panel of
5 cerebrospinal fluid biomarkers in the differential diagnosis of
5 patients with dementia and/or parkinsonian disorders. Arch Neurol
5 2012;69:1445.
5 Magdalino NK, Paterson RW, Schott JM, et al. A panel of
5 nine cerebrospinal fluid biomarkers may identify patients with
5 atypical parkinsonian syndromes. J Neurol Neurosurg Psychiatry
5 Teunissen CE, Khalil M. Neurofilaments as biomarkers in multiple
5 Gaiottino J, Norgren N, Dobson R, et al. Increased neurofilament light
5 chain blood levels in neurodegenerative neurological diseases. PLoS
5 One 2012;8:e76931.
5 Kuhle J, Nourbakhsh B, Grant D, et al. Serum neurofilament is
5 associated with progression of brain atrophy and disability in early
5 Kuhle J, Barro C, Andreasson U, et al. Comparison of three analytical
5 platforms for quantification of the neurofilament light chain in blood
5 samples: ELISA, electrochemiluminescence immunoassay and
5 of the Neurofilament Light Protein (NFL) is a Biomarker of CNS
5 biomarker of neuronal damage in multiple sclerosis. Ann Neurol
5 2017;81:857–70.
5 Moher D, Liberati A, Tetzlaff J, et al. Preferred reporting items for
5 systematic reviews and meta-analyses: the PRISMA statement. PLoS
5 Med 2009;6:e1000097. doi:10.1371/journal.pmed.1000097
5 Moher D, Shamseer L, Clarke M. Preferred reporting items for
5 systematic review and meta-analysis protocols (PRISMA-P) 2015
5 431-4-5.
5 Shamseer L, Moher D, Clarke M, et al. Preferred reporting items for
5 systematic review and meta-analysis protocols (PRISMA-P) 2015:
5 elaboration and explanation. BMJ 2015;349:g7647.
5 Booth A, Chialvo M, Dooley G, et al. The nuts and bolts of Prosos:
5 an international prospective register of systematic reviews. Syst
5 NHLBI. Study quality assessment tools | National heart, lung, and
5 health-topics/study-quality-assessment-tools [Accessed 6 Mar
5 2020].
5 de Flon P, Laurell K, Sundström P, et al. Comparison of plasma and
5 cerebrospinal fluid neurofilament light in a multiple sclerosis trial. 
5 biomarkers for the diagnosis of Alzheimer’s disease in adults with
5 Down syndrome: a cross-sectional study. Lancet Neurol 2018;
5 17:960–9.
5 the cerebrospinal fluid and serum of patients with amyotrophic lateral
5 Häkansson I, Tissell A, Cassel P, et al. Neurofilament levels, disease
5 activity and brain volume during follow-up in multiple sclerosis. J
5 Neuroinflammation 2019;16:239.
5 Halaas NB, Blenkov N, Iland AV. Neurofilament light in serum and
5 cerebrospinal fluid of hip fracture patients with delirium. Dement 
5 Geriatr Cogn Disord 2019;46. doi:10.1159/000494754
5 for differential diagnosis of parkinsonian disorder. Neurology
5 Kalm M, Bostrom M, Sandelius Asa, et al. Serum concentrations of
5 the axonal injury marker neurofilament light protein are not influenced by
5 in early relapsing remitting MS is increased and correlates with
5 CSF levels and with MRI measures of disease severity. Mult Scler
5 a prognostic biomarker in amyotrophic lateral sclerosis. Neurology
5 2015;84:2247–57.
5 Mariotti S, Farinazzo A, Magliozzi R, et al. Serum and cerebrospinal
5 neurofilament light chain levels in patients with acquired peripheral
5 Mariotti S, Gajofato A, Zuliani L, et al. Serum and CSF
5 neurofilament light chain levels in antibody-mediated encephalitis. J
5 Neurol Neurosurg Psychiatry 2019;266:1643–8.


**Supplemental Table 1**: Characteristics of the studies included in the meta-analysis

<table>
<thead>
<tr>
<th>Study</th>
<th>Source and number of paired CSF and blood NfL samples</th>
<th>Disease affecting CNS only, CNS+PNS, or PNS only</th>
<th>Plasma or serum NfL measured</th>
<th>Blood NfL assay used</th>
<th>Correlation technique used</th>
<th>Correlation coefficient value</th>
<th>Quality score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abdellhak et al[1], 2018</td>
<td>n=80 MS, n=20 controls</td>
<td>CNS only</td>
<td>Serum</td>
<td>Simoa</td>
<td>Spearman's</td>
<td>MS: r=0.7 controls: r= 0.6</td>
<td>7</td>
</tr>
<tr>
<td>Abdellhak et al[2], 2019</td>
<td>n=71 PPMS</td>
<td>CNS only</td>
<td>Serum</td>
<td>Simoa</td>
<td>Spearman's</td>
<td>r=0.6</td>
<td>7</td>
</tr>
<tr>
<td>Benatar et al[3], 2017</td>
<td>n=154 (n=18 controls, n=110 at risk ALS, n=16 ALS, n=10 phenoconverters)</td>
<td>CNS+PNS</td>
<td>Serum</td>
<td>ECL</td>
<td>Spearman's</td>
<td>r=0.7</td>
<td>8</td>
</tr>
<tr>
<td>Bergman et al[4], 2016</td>
<td>n=60 PMS</td>
<td>CNS only</td>
<td>Serum</td>
<td>Simoa</td>
<td>Pearson's</td>
<td>r=0.775</td>
<td>7</td>
</tr>
<tr>
<td>Byrne et al[5], 2017</td>
<td>n=29 (n=12 controls, n=2 premanifest HD, n=15 manifest HD)</td>
<td>CNS+PNS</td>
<td>Plasma</td>
<td>Simoa</td>
<td>Pearson's</td>
<td>r=0.868</td>
<td>7</td>
</tr>
<tr>
<td>Byrne et al[6], 2018</td>
<td>n=80 (n=20 controls, n=20 manifest HD, n=40 pre-HD)</td>
<td>CNS+PNS</td>
<td>Plasma</td>
<td>Simoa</td>
<td>Pearson's</td>
<td>r=0.914</td>
<td>8</td>
</tr>
<tr>
<td>deFlon et al[7], 2019</td>
<td>n=281 RRMS</td>
<td>CNS only</td>
<td>Plasma</td>
<td>Simoa</td>
<td>Spearman's</td>
<td>r=0.445</td>
<td>7</td>
</tr>
<tr>
<td>Disanto et al[8], 2017</td>
<td>n=132 MS</td>
<td>CNS only</td>
<td>Serum</td>
<td>Simoa</td>
<td>Pearson's</td>
<td>r=0.77</td>
<td>7</td>
</tr>
<tr>
<td>Fortea et al[9], 2018</td>
<td>n=94 (n=54 asymptomatic, n=18 prodromal AD, n=22 AD)</td>
<td>CNS only</td>
<td>Plasma</td>
<td>Simoa</td>
<td>not reported</td>
<td>r=0.80</td>
<td>8</td>
</tr>
<tr>
<td>Gaiottino et al[10], 2013</td>
<td>n=20 AD, n=19 GBS, n=45 ALS, n=66 controls</td>
<td>AD: CNS only, GBS: PNS only, ALS: CNS+PNS</td>
<td>Serum</td>
<td>ECL</td>
<td>Spearman's</td>
<td>AD: r=0.48 GBS: r=0.79 ALS: r=0.70 controls r=0.11</td>
<td>6</td>
</tr>
<tr>
<td>Gisslen et al[11], 2016</td>
<td>n=140 (n=121 with HIV, n=19 controls)</td>
<td>CNS+PNS</td>
<td>Plasma</td>
<td>Simoa</td>
<td>Pearson's</td>
<td>r=0.89</td>
<td>7</td>
</tr>
<tr>
<td>Gong et al[12], 2018</td>
<td>N=40 ALS, n=40 controls</td>
<td>CNS+PNS</td>
<td>Serum</td>
<td>ELISA</td>
<td>Spearman's</td>
<td>ALS r=0.5997 controls r=0.0226</td>
<td>7</td>
</tr>
<tr>
<td>Study</td>
<td>n (MS, controls)</td>
<td>Region</td>
<td>Media</td>
<td>Assay</td>
<td>Correlation</td>
<td>r</td>
<td>Study 4</td>
</tr>
<tr>
<td>-------------------------------</td>
<td>------------------</td>
<td>--------</td>
<td>-------</td>
<td>-------</td>
<td>---------------</td>
<td>-----</td>
<td>------------</td>
</tr>
<tr>
<td>Hakansson et al [13], 2018</td>
<td>n=63 (n=41 MS, n=22 controls)</td>
<td>CNS only</td>
<td>Serum</td>
<td>Simoa</td>
<td>Pearson's</td>
<td>0.74</td>
<td>Study 1</td>
</tr>
<tr>
<td>Halaas et al [14], 2019</td>
<td>n=130 with hip fracture and delirium, n=123 with hip fracture and no delirium</td>
<td>CNS+PNS</td>
<td>Serum</td>
<td>Simoa assay</td>
<td>Spearman's</td>
<td>Cases: r= 0.56 controls r= 0.69</td>
<td></td>
</tr>
<tr>
<td>Hansson et al [15], 2017</td>
<td>Lund: n=245 (n=147 PD, n=28 MSA, n=15 PSP, n=5 CBS, n=50 controls) London: n=97 (n=5 PD, n=29 MSA, n=26 PSP, n=11 CBS, n=26 controls)</td>
<td>CNS+PNS</td>
<td>Plasma</td>
<td>Simoa assay</td>
<td>Spearman's</td>
<td>Lund cohort: r=0.730 London cohort: r=0.840</td>
<td></td>
</tr>
<tr>
<td>Kalm et al [16], 2017</td>
<td>n=43 (n=20 CSF:serum albumin &lt;10, n=15 CSF:serum albumin 10–20, n=8 CSF:serum albumin &gt; 20)</td>
<td>CNS+PNS</td>
<td>Serum</td>
<td>Simoa</td>
<td>Pearson's</td>
<td>0.6303</td>
<td></td>
</tr>
<tr>
<td>Kuhle et al [17], 2016</td>
<td>n=33 (a variety of neurological disorders)</td>
<td>CNS+PNS</td>
<td>Serum</td>
<td>ELISA, ECL, Simoa</td>
<td>Spearman's</td>
<td>Simoa: r = 0.88 ECL: r = 0.78 ELISA: r = 0.38</td>
<td></td>
</tr>
<tr>
<td>Kuhle et al [18], 2016</td>
<td>n=31 RRMS</td>
<td>CNS only</td>
<td>Serum</td>
<td>ECL</td>
<td>Pearson's</td>
<td>0.62</td>
<td>Study 4</td>
</tr>
<tr>
<td>Lu et al [19], 2015</td>
<td>n=38 ALS, n = 20 controls</td>
<td>CNS+PNS</td>
<td>Serum</td>
<td>ECL</td>
<td>Spearman's</td>
<td>0.78</td>
<td>Study 4</td>
</tr>
<tr>
<td>Mariotto et al [20], 2018</td>
<td>n=12 acquired peripheral neuropathy</td>
<td>PNS only</td>
<td>Serum</td>
<td>Simoa</td>
<td>Pearson's</td>
<td>0.71</td>
<td>Study 4</td>
</tr>
<tr>
<td>Mariotto et al [21], 2019</td>
<td>n = 23 CNS syndromes+autoantibodies to neuronal cell surface antigens</td>
<td>CNS only</td>
<td>Serum</td>
<td>Simoa</td>
<td>Spearman's</td>
<td>0.461</td>
<td>Study 4</td>
</tr>
<tr>
<td>Marques et al [22], 2019</td>
<td>n=137 (n=55 PD, n=29 APD, n=53 controls)</td>
<td>CNS+PNS</td>
<td>Serum</td>
<td>Simoa</td>
<td>Spearman's</td>
<td>0.72</td>
<td>Study 4</td>
</tr>
<tr>
<td>Mattsson et al [23], 2017</td>
<td>n=391 (n=112 controls, n=189 MCI, n=90 AD)</td>
<td>CNS only</td>
<td>Plasma</td>
<td>Simoa</td>
<td>Spearman's</td>
<td>0.59</td>
<td>Study 4</td>
</tr>
<tr>
<td>Meeter et al [24], 2016</td>
<td>n=55 genetic FTD</td>
<td>CNS only</td>
<td>Serum</td>
<td>ECL</td>
<td>Spearman's</td>
<td>0.87</td>
<td>Study 4</td>
</tr>
<tr>
<td>Menke et al [25], 2015</td>
<td>n=25 ALS, n=17 controls</td>
<td>CNS+PNS</td>
<td>Serum</td>
<td>ECL</td>
<td>Spearman's</td>
<td>0.642</td>
<td>Study 4</td>
</tr>
<tr>
<td>Mielke et al [26], 2019</td>
<td>n=79 without dementia (n=64 cognitively unimpaired, n=15 MCI)</td>
<td>CNS only</td>
<td>Plasma</td>
<td>ELISA</td>
<td>Spearman's</td>
<td>0.568</td>
<td>Study 4</td>
</tr>
<tr>
<td>Study</td>
<td>Sample Size</td>
<td>Region</td>
<td>Fluid</td>
<td>Platform</td>
<td>Correlation Type</td>
<td>Correlation</td>
<td>Comments</td>
</tr>
<tr>
<td>-------------------------------</td>
<td>-------------</td>
<td>------------</td>
<td>---------</td>
<td>----------</td>
<td>------------------</td>
<td>-------------</td>
<td>----------</td>
</tr>
<tr>
<td>Novakova et al[27], 2017</td>
<td>n=434 MS, n=45 OND, n=42 controls</td>
<td>CNS only</td>
<td>Serum</td>
<td>Simoa</td>
<td>Spearman's</td>
<td>r=0.620</td>
<td>controls: r=0.385 OND: r=0.740</td>
</tr>
<tr>
<td>Piehl et al[28], 2018</td>
<td>n=66 MS</td>
<td>CNS only</td>
<td>Serum</td>
<td>Simoa</td>
<td>Spearman's</td>
<td>r=0.672</td>
<td></td>
</tr>
<tr>
<td>Rojas et al[29], 2018</td>
<td>n=22 PSP</td>
<td>CNS only</td>
<td>Plasma</td>
<td>Simoa</td>
<td>Pearson's</td>
<td>r=0.64</td>
<td></td>
</tr>
<tr>
<td>Rojas et al[30], 2016</td>
<td>n=15 PSP</td>
<td>CNS only</td>
<td>Plasma</td>
<td>Simoa</td>
<td>Pearson's</td>
<td>r=0.74</td>
<td></td>
</tr>
<tr>
<td>Sanchez-Valle et al[31], 2018</td>
<td>n=35 AD</td>
<td>CNS only</td>
<td>Serum</td>
<td>Simoa</td>
<td>Spearman's</td>
<td>r=0.72</td>
<td></td>
</tr>
<tr>
<td>Sehr et al[32], 2019</td>
<td>n = 15 RRMS</td>
<td>CNS only</td>
<td>Plasma</td>
<td>Simoa</td>
<td>Spearman's</td>
<td>r=0.74</td>
<td></td>
</tr>
<tr>
<td>Staffaroni et al[33], 2018</td>
<td>n=49 sporadic CJD</td>
<td>CNS+PNS</td>
<td>Plasma</td>
<td>Simoa</td>
<td>Pearson's</td>
<td>r=0.8</td>
<td></td>
</tr>
<tr>
<td>Steinacker et al[34], 2018</td>
<td>n=37 bvFTD</td>
<td>CNS only</td>
<td>Serum</td>
<td>Simoa</td>
<td>Spearman's</td>
<td>r=0.706</td>
<td></td>
</tr>
<tr>
<td>Weydt et al[35], 2016</td>
<td>n=95 (n=19 controls, n=12 asymptomatic ALS, n=64 symptomatic ALS)</td>
<td>CNS+PNS</td>
<td>Serum and plasma</td>
<td>ECL</td>
<td>Spearman's</td>
<td>r=0.79</td>
<td></td>
</tr>
<tr>
<td>Wilke et al[36], 2016</td>
<td>n = 112 (n = 41 FTD, n = 25 ALS, n = 46 controls)</td>
<td>FTD: CNS only, ALS: CNS+PNS</td>
<td>Serum</td>
<td>ECL</td>
<td>Pearson's</td>
<td>FTD: r=0.72 ALS: r=0.58 controls: r=0.53</td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: CSF = cerebrospinal fluid, NfL = neurofilament light chain protein, CNS = central nervous system, PNS = peripheral nervous system, MS = multiple sclerosis, RRMS = relapsing-remitting multiple sclerosis, PMS = progressive multiple sclerosis, ECL = Electrochemiluminescence immunoassay, Simoa = single-molecule array digital immunoassay, ELISA = enzyme-linked immunosorbent assay, ALS = Amyotrophic lateral sclerosis, PPMS = Primary progressive multiple sclerosis, HD = Huntington’s disease, AD = Alzheimer’s disease, PD = Parkinson’s disease, MSA = multiple system atrophy, PSP = progressive supranuclear palsy, CBS = corticobasal syndrome, GBS = Guillain-Barré syndrome, APD = atypical Parkinsonism disorders, MCI = mild cognitive impairment, FTD = frontotemporal dementia, bvFTD = behavioural variant frontotemporal dementia, CJD = Creutzfeldt–Jakob disease, OND = other neurological disorder
References:

12 Gong ZY, Lv GP, Gao LN, Lu Y, Guo J, Zang DW. Neurofilament Subunit L Levels in the Cerebrospinal Fluid and Serum of Patients with Amyotrophic Lateral Sclerosis. Neurodegener Dis Published Online First: 2018. doi:10.1159/000488681


20 Mariotto S, Farinazzo A, Magliozzi R, Alberti D, Monaco S, Ferrari S. Serum and cerebrospinal neurofilament light chain levels in patients with acquired


### Supplemental Table 2: Quality assessment of the included studies: a) Cross-sectional studies and observational cohorts and b) Before-after studies with no control group

#### a) Quality assessment questions for cross-sectional studies and observational cohorts

<table>
<thead>
<tr>
<th>Author</th>
<th>Year of publication</th>
<th>Q1</th>
<th>Q2</th>
<th>Q3</th>
<th>Q4</th>
<th>Q5</th>
<th>Q6</th>
<th>Q7</th>
<th>Q8</th>
<th>Q9</th>
<th>Q10</th>
<th>Q11</th>
<th>Q12</th>
<th>Q13</th>
<th>Q14</th>
<th>Quality Assessment score</th>
<th>Outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abdellhak</td>
<td>2018</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>N/A</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>7</td>
<td>Good</td>
<td></td>
</tr>
<tr>
<td>Abdellhak</td>
<td>2019</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>N/A</td>
<td>Y</td>
<td>N</td>
<td>N/A</td>
<td>N</td>
<td>7</td>
<td>Good</td>
<td></td>
</tr>
<tr>
<td>Benatar</td>
<td>2017</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N/A</td>
<td>Y</td>
<td>Y</td>
<td>N/A</td>
<td>N</td>
<td>8</td>
<td>Good</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Byrne</td>
<td>2017</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>N/A</td>
<td>Y</td>
<td>N</td>
<td>N/A</td>
<td>N</td>
<td>7</td>
<td>Good</td>
<td></td>
</tr>
<tr>
<td>Byrne</td>
<td>2018</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>N/A</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>8</td>
<td>Good</td>
<td></td>
</tr>
<tr>
<td>Dianso</td>
<td>2017</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>N/A</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N/A</td>
<td>7</td>
<td>Good</td>
<td></td>
</tr>
<tr>
<td>Fortea</td>
<td>2018</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>N/A</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N/A</td>
<td>N</td>
<td>7</td>
<td>Good</td>
<td></td>
</tr>
<tr>
<td>Gaiottino</td>
<td>2013</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>CD</td>
<td>N</td>
<td>Y</td>
<td>N/A</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>N/A</td>
<td>6</td>
<td>Fair</td>
<td></td>
</tr>
<tr>
<td>Gisslen</td>
<td>2016</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>N/A</td>
<td>Y</td>
<td>N</td>
<td>N/A</td>
<td>N</td>
<td>7</td>
<td>Good</td>
<td></td>
</tr>
<tr>
<td>Gong</td>
<td>2018</td>
<td>Y</td>
<td>Y</td>
<td>CD</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N/A</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>7</td>
<td>Good</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hakansson</td>
<td>2018</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N/A</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N/A</td>
<td>7</td>
<td>Good</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Halaas</td>
<td>2019</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>CD</td>
<td>N</td>
<td>Y</td>
<td>N/A</td>
<td>Y</td>
<td>Y</td>
<td>N/A</td>
<td>N</td>
<td>6</td>
<td>Fair</td>
<td></td>
</tr>
<tr>
<td>Hansson</td>
<td>2017</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>N/A</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>N/A</td>
<td>7</td>
<td>Good</td>
<td></td>
</tr>
<tr>
<td>Kalm</td>
<td>2017</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>CD</td>
<td>N</td>
<td>Y</td>
<td>N/A</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N/A</td>
<td>5</td>
<td>Fair</td>
<td></td>
</tr>
<tr>
<td>Kuhle</td>
<td>2016</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>N/A</td>
<td>Y</td>
<td>Y</td>
<td>N/A</td>
<td>N</td>
<td>7</td>
<td>Fair</td>
<td></td>
</tr>
<tr>
<td>Kuhle</td>
<td>2016</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>N/A</td>
<td>Y</td>
<td>Y</td>
<td>N/A</td>
<td>N</td>
<td>8</td>
<td>Good</td>
<td></td>
</tr>
<tr>
<td>Lu</td>
<td>2015</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>N/A</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N/A</td>
<td>8</td>
<td>Good</td>
<td></td>
</tr>
<tr>
<td>Mariotto</td>
<td>2018</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>N/A</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N/A</td>
<td>8</td>
<td>Good</td>
<td></td>
</tr>
<tr>
<td>Mariotto</td>
<td>2019</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>N/A</td>
<td>Y</td>
<td>Y</td>
<td>N/A</td>
<td>N</td>
<td>8</td>
<td>Good</td>
<td></td>
</tr>
<tr>
<td>Marques</td>
<td>2019</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>N/A</td>
<td>Y</td>
<td>Y</td>
<td>N/A</td>
<td>N</td>
<td>9</td>
<td>Good</td>
</tr>
<tr>
<td>Mattsson</td>
<td>2017</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>N/A</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>N/A</td>
<td>8</td>
<td>Good</td>
</tr>
<tr>
<td>Meeter</td>
<td>2016</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N/A</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>N/A</td>
<td>9</td>
<td>Good</td>
<td></td>
</tr>
<tr>
<td>Menke</td>
<td>2015</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>N/A</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N/A</td>
<td>7</td>
<td>Good</td>
<td></td>
</tr>
<tr>
<td>Mielke</td>
<td>2019</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>CD</td>
<td>N</td>
<td>Y</td>
<td>N/A</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N/A</td>
<td>6</td>
<td>Fair</td>
<td></td>
</tr>
<tr>
<td>Novakova</td>
<td>2017</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N/A</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N/A</td>
<td>8</td>
<td>Good</td>
</tr>
<tr>
<td>Pietsch</td>
<td>2018</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>N/A</td>
<td>Y</td>
<td>N/A</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N/A</td>
<td>7</td>
<td>Good</td>
<td></td>
</tr>
<tr>
<td>Rojas</td>
<td>2018</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>N/A</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N/A</td>
<td>7</td>
<td>Good</td>
<td></td>
</tr>
<tr>
<td>Rojas</td>
<td>2016</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>N/A</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>N/A</td>
<td>7</td>
<td>Good</td>
<td></td>
</tr>
<tr>
<td>Sanchez-Valle</td>
<td>2018</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>N/A</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>N/A</td>
<td>N</td>
<td>7</td>
<td>Good</td>
<td></td>
</tr>
<tr>
<td>Sehe</td>
<td>2019</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>N/A</td>
<td>Y</td>
<td>N/A</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N/A</td>
<td>8</td>
<td>Good</td>
<td></td>
</tr>
<tr>
<td>Staffanoni</td>
<td>2019</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>N/A</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>7</td>
<td>Good</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Steinacker</td>
<td>2018</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>N/A</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>8</td>
<td>Good</td>
<td></td>
</tr>
<tr>
<td>Weydt</td>
<td>2016</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>N/A</td>
<td>Y</td>
<td>N/A</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>N/A</td>
<td>7</td>
<td>Good</td>
<td></td>
</tr>
<tr>
<td>Wilke</td>
<td>2016</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N/A</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N/A</td>
<td>9</td>
<td>Good</td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: CD = Cannot determine, N/A = not applicable, NR = not reported
Question 1 (Q1): Was the research question or objective clearly stated?

Question 2 (Q2): Was the study population clearly stated and defined?

Question 3 (Q3): Was the participation rate for eligible persons at least 50%?

Question 4 (Q4): Were all the subjects recruited from similar populations? Were inclusion and exclusion criteria pre-specified and applied uniformly to all participants?

Question 5 (Q5): Was a sample size justification, power description, or variance and effect estimates provided?

Question 6 (Q6): For the analyses in this paper, were the exposure of interest measured prior to the outcome being measured?

Question 7 (Q7): Was the timeframe sufficient so that one could reasonably expect to see an association between exposure and outcome if it existed?

Question 8 (Q8): For exposures that can vary in amount or level, did the study examine different levels of the exposure as related to the outcome (e.g., categories of exposure, or exposure measured as continuous variable)?

Question 9 (Q9): Were the exposure measures (independent variables) clearly defined, valid, reliable, and implemented consistently across all study participants?

Question 10 (Q10): Was the exposure assessed more than once over time?

Question 11 (Q11): Were the outcome measures clearly defined, valid, reliable, and implemented consistently across all study participants?

Question 12 (Q12): Were the outcome assessors blinded to the exposure status of participants?

Question 13 (Q13): Was loss to follow-up after baseline 20% or less?

Question 14 (Q14): Were key potential confounding variables measured and adjusted statistically for their impact on the relationship between exposure and outcome?
### b) Quality assessment questions for before-after studies with no control group

<table>
<thead>
<tr>
<th>Author</th>
<th>Year of publication</th>
<th>Q1</th>
<th>Q2</th>
<th>Q3</th>
<th>Q4</th>
<th>Q5</th>
<th>Q6</th>
<th>Q7</th>
<th>Q8</th>
<th>Q9</th>
<th>Q10</th>
<th>Q11</th>
<th>Q12</th>
<th>Quality Assessment score</th>
<th>Outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bergman</td>
<td>2016</td>
<td>Y</td>
<td>Y</td>
<td>CD</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>NR</td>
<td>Y</td>
<td>N/A</td>
<td>Y</td>
<td>N/A</td>
<td>Y</td>
<td>7</td>
<td>Good</td>
</tr>
<tr>
<td>deFlon</td>
<td>2019</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>CD</td>
<td>Y</td>
<td>Y</td>
<td>NR</td>
<td>Y</td>
<td>N/A</td>
<td>Y</td>
<td>N/A</td>
<td>7</td>
<td>Good</td>
</tr>
</tbody>
</table>

**Abbreviations:** CD = Cannot determine, N/A = not applicable, NR = not reported

**Question 1 (Q1):** Was the study question or objective clearly stated?

**Question 2 (Q2):** Were eligibility/selection criteria for the study population pre-specified and clearly described?

**Question 3 (Q3):** Were the participants in the study representative of those who would be eligible for the test/service/intervention in the general or clinical population of interest?

**Question 4 (Q4):** Were all eligible participants that met the pre-specified entry criteria enrolled?

**Question 5 (Q5):** Was the sample size sufficiently large to provide confidence in the findings?

**Question 6 (Q6):** Was the test/service/intervention clearly described and delivered consistently across the study population?

**Question 7 (Q7):** Were the outcome measures pre-specified, clearly defined, valid, reliable, and assessed consistently across all study participants?

**Question 8 (Q8):** Were the people assessing the outcomes blinded to the participants' exposures/interventions?

**Question 9 (Q9):** Was the loss to follow-up after baseline 20% or less? Were those lost to follow-up accounted for in the analysis?

**Question 10 (Q10):** Did the statistical methods examine changes in outcome measures from before to after the intervention? Were statistical tests done that provided p values for the pre-to-post changes?

**Question 11 (Q11):** Were outcome measures of interest taken multiple times before the intervention and multiple times after the intervention (i.e., did they use an interrupted time-series design)?

**Question 12 (Q12):** If the intervention was conducted at a group level (e.g., a whole hospital, a community, etc.) did the statistical analysis take into account the use of individual-level data to determine effects at the group level?