Prognostic value of phrenic nerve conduction study in amyotrophic lateral sclerosis: systematic review and meta-analysis

Cláudia S Silva[1], Filipe B Rodrigues[1,2], Gonçalo S Duarte[1], João Costa[1,2,3,4], Mamede de Carvalho[1,5]

1 – Faculdade de Medicina-Instituto de Medicina Molecular, Universidade de Lisboa, Lisbon, Portugal.

2 – UCL Huntington's Disease Centre, UCL Queen Square Institute of Neurology, University College London, London, United Kingdom

3 – Center for Evidence-Based Medicine, Faculdade de Medicina, Universidade de Lisboa, Lisbon, Portugal

4 – Portuguese Collaborating Center of the IberoAmerican Cochrane Network, Faculdade de Medicina, Universidade de Lisboa, Lisbon, Portugal

5 –Department of Neurosciences and Mental Health, Centro Hospitalar Universitário LIsboa-Norte, Lisbon, Portugal

Keywords: amyotrophic lateral sclerosis; meta-analysis; phrenic nerve motor response; systematic review; survival.

Running title: Phrenic nerve in ALS

Number of words: 3939

Corresponding author:
Mamede de Carvalho
Institute of Molecular Medicine, Faculty of Medicine, University of Lisbon
Av. Professor Egas Moniz, 1648-028 Lisbon, Portugal
Email: mamedemg@mail.telepac.pt
Phone and Fax: + 351 21 7805219

Acknowledgements

We are grateful to GAPIC – Gabinete de Apoio à Investigação Científica, Tecnológica e Inovação (Scientific Research, Technology and Innovation Support Office), Faculty of Medicine, University of Lisbon, for the financial support provided to Claudia S Silva. This research was supported by UID/BIM/50005/2019, project funded by Fundação para a Ciência e a Tecnologia (FCT)/ Ministério da Ciência, Tecnologia e Ensino Superior (MCTES) through Fundos do Orçamento de Estado

Declaration of interest

The authors report no conflict of interests.

Abstract

Objectives

To assess the prognostic value of phrenic nerve conduction (PNC) in amyotrophic lateral sclerosis (ALS).

Methods

We conducted a systematic review to identify studies reporting on PNC, and mortality and/or forced vital capacity (FVC) in patients with ALS. We searched Medline, EMBASE, and Web of Science. Two independent authors selected studies and extracted data. Risk of bias was assessed using the QUIPS tool. Hazard-ratios and correlation coefficients were pooled using a random effects generic inverse-variance model. Evidence quality was evaluated with GRADE.

Results

In the pooled analysis, patients with CMAP-amplitude equal or below 0.4mV are 2.021 more likely to die over the studied period (95%CI 1.161 to 3.522; I^2=55.9%; 338 participants). CMAP-amplitude showed a moderate positive correlation with FVC (r=0.400, 95%CI= 0.226 to 0.550; I^2=69.77%; 381 participants). However, there was a weak negative correlation between CMAP-latency and FVC (r=-0.235; 95%CI= -0.447 to -0.024; I^2=15.92%; 112 participants).

Conclusions

There is moderate-quality evidence that CMAP-amplitude of the PNC is correlated with FVC. Results favour a predictive value for mortality, but the risk of bias is high.

Significance

PNC is a simple test that should be considered to assess respiratory function in ALS, especially in patients with bulbar involvement or cognitive impairment.

Highlights

1 – Investigation of phrenic nerve conduction is a simple and quick test

2 - Phrenic nerve amplitude probably predicts survival and is probably correlated to FVC

2 – This test should be considered to assess respiratory function in ALS.

1. Introduction

Amyotrophic lateral sclerosis (ALS) is a rapidly progressive neurodegenerative condition that involves the upper and lower motor neurons, typically affecting the limb, bulbar, and respiratory muscles (Haverkamp *et al.*, 1995). Respiratory failure (RF) frequently arises in late-stage disease, although it can be the presenting feature in about 3% of patients (de Carvalho *et al.*, 1996). Hypoventilation with hypoxaemia and hypercapnia (Fallat *et al.*, 1979) often precipitated by respiratory infection, aspiration pneumonia or bronchial impaction, (Paulukonis *et al.*, 2015) is the most frequent cause of death in ALS. The median overall survival after symptom onset is approximately 3 years. (Paulukonis *et al.*, 2015). Accurate assessment of pulmonary function is then critical to detect early abnormalities, in order to estimate prognosis and to set treatment strategy (Pinto and Carvalho, 2014).

Several test have been investigated to ascertain the optimal measure of pulmonary function in people with ALS (Kaplan and Hollander, 1994; de Carvalho *et al*., 2019). Pulmonary function tests (PFT) are the standard technique used across most diseases, being non-invasive and widely available. Among these, forced vital capacity (FVC) is the one mostly used in ALS management and research, as it can predict hypercapnia (Kaplan and Hollander, 1994; Pinto *et al.*, 2009b) and prognosis (Czaplinski *et al.*, 2006; Fallat *et al.*, 1979). However, PFT have notable limitations, namely that the technique depends on patient cooperation, which can be disturbed in people with apathy, depression, or other behavioural changes, all of which are relatively common in ALS (Phukan *et al*., 2007; de Carvalho *et al*., 2019). Additionally, bulbar involvement is frequent in ALS and may cause facial weakness, which further impairs an accurate assessment by PFT. Altogether, this indicates that predictive value of FVC is compromised in patients with cognitive changes and marked bulbar involvement (de Carvalho *et al*., 2019).

Phrenic nerve conduction study (PNC) is an alternative technique that records the phrenic compound muscle action potential (CMAP) applying a percutaneous electrical stimulation in neck (de Carvalho *et al.*, 2018). Crucially, PNC does not depend on patient collaboration and

can be used irrespective of bulbar involvement. PNC records the phrenic CMAP applying a percutaneous electrical stimulation in neck (de Carvalho *et al.*, 2018). CMAP amplitude elicited by phrenic nerve stimulation correlates with diaphragm dysfunction (de Carvalho *et al.*, 2018) and is predictive of both hypoventilation (Pinto *et al.*, 2009b) and survival in ALS (Pinto *et al.*, 2012). Moreover is sensitive to detect changes over a short period of disease progression (Pinto *et al.*, 2009a). Nevertheless, like other electrophysiological techniques, PNC is not as widely available as PFT, is subject to technical pitfalls and is operator dependent.

The aim of this systematic review and meta-analysis was to ascertain the prognostic value of PNC in people with ALS.

2. Methods

The protocol of this study followed the PRISMA-P guidelines (Moher *et al.*, 2015) and was registered at the PROSPERO database (CRD42017079438). Reporting followed the MOOSE (Stroup *et al.*, 2000) and PRISMA guidelines (Liberati *et al.*, 2009). Statistical data reporting followed the SAMPL guidelines (Lang and Altman, 2015).

2.1. Eligibility criteria

2.1.1. Types of studies

We considered observational studies that compared the results of the index test, PNC, with the reference standard, PFT. We included studies in which data have been collected either prospectively or retrospectively from consecutive series of people with ALS, followed in any setting. No restrictions were made based on a minimal quality standard, minimal sample sizes, number of diseased cases, language, publication status, or data of publication.

2.1.2. Participants

We included adults with a diagnosis of definitive, probable or probable laboratory-supported ALS, as defined by the modified *El Escorial* criteria (Brooks *et al.*, 2000), of all ages and in any setting.

2.2. Information sources and search strategy

Electronic identification of reports was conducted in Medline (via Ovid), EMBASE (via Ovid), and Web of Science, using the strategies outlined in Appendices B, C and D, from inception until October 2017. Text words and database subject headings (i.e.: MeSH and EMTREE) were used in our search strategy.

In order to identify additional published studies we crosschecked the reference lists of studies included for full text revision, and contacted experts in the field. We excluded unpublished results, with the exception of conference proceedings providing that these provide sufficient data.

2.3. Study selection

Two independent review authors (CSS, FBR) performed the first selection based on title and abstract. Each author identified the studies requiring full text review. All studies identified as potential eligible studies were subject to full text review (CSS, FBR). Disagreements were solved by discussion or by a third author (GD). Both procedures were performed with Covidence®.

2.4. Data extraction and management

Two independent reviewers (CSS, FBR) extracted data from the studies included in this review using a pre-piloted standardised electronic form. Disagreements were resolved by consensus or with the help of a third reviewer (GD). The third reviewer (GD) also double-checked.

2.5. Outcomes and prioritisation

The primary outcome was the assessment of the prognostic value of PNC in overall survival. According to the type of data available, this outcome was assessed using one or more of:

- Hazard ratios for mortality using different cut-off thresholds of the CMAP amplitude obtained by PNC.
- Mortality rates using different cut-off thresholds of the CMAP amplitude obtained by PNC. Unfortunately no mortality rate data were available.

For the primary outcome, a subgroup analysis was to be performed according to the main ALS phenotypical expressions, as well as a sensitivity analysis by excluding studies deemed to be at high risk of bias. The secondary outcomes were the disease severity, based on the correlation between FVC, our index test, and the CMAP amplitude, as well as the correlation between FVC and CMAP latency.

2.6. Assessment of risk of bias

The risk of bias of included studies was evaluated independently by two review authors (CSS, FBR) using the QUIPS (Quality in Prognosis Studies) tool (Hayden *et al.*, 2006; 2013).

Six domains were deemed critical for assessing biases sufficiently large to distort the findings of prognosis research: (1) study participation; (2) study attrition; (3) prognostic factor measurement; (4) outcome measurement; (5) study confounding; and (6) statistical analysis and reporting. For each domain, three to seven “prompting items” were used to rate the adequacy of reporting by a study as “yes”, “partial”, “no”, or “unsure”; an overall rating for each domain is assigned as “high”, “moderate”, or “low” risk of bias. Disagreements were solved by discussion or with consultation of a third review author (GD) in case of persisting disagreement.

2.7. Statistical analysis and data synthesis

To pool hazard ratios, we used a random effects generic inverse-variance method.

To pool the effect sizes of correlations coefficients of two continuous variables, we first transformed the coefficient into Fisher's *z* according to formula A.1. For the two studies (Bokuda *et al.*, 2014; Jenkins *et al.*, 2016) that presented a simple (i.e. univariable) linear regression, the correlation coefficient, *r*, was calculated as the square root of the coefficient of determination, R^2. The variance of z was computed according with formula A.2., and the standard error (SE) according to formula A.3. The meta-analyses was conducted using the random effect inverse-variance approach. The final results were back-transformed according to formula A.4. (Borenstein *et al.*, 2009).

We assessed statistical heterogeneity with I^2 (Higgins, Thompson, 2002). A p-value <0.05 was considered statistically significant. All statistical analyses were done using Stata® (College Station, TX) 15.0 software.

2.8. Confidence in cumulative evidence

We evaluated the quality of the evidence using the grading of recommendations assessment, development and evaluation (GRADE) working group methods extended to prognosis factor research (Huguet *et al.*, 2013).

Our evaluation was based on five domains that may decrease quality: (1) study limitations; (2) inconsistency; (3) indirectness; (4) imprecision; and (5) publication bias; and two factors that may increase quality: (1) moderate or large effect size; and (2) exposure-response gradient (Huguet *et al.*, 2013).

1. Study limitations: according with the QUIPS tool for the risk of bias, outcomes are rated as: (1) no serious limitations for studies with low risk of bias for most of the bias domains; (2) serious limitations for studies at moderate or unclear risk of bias for most of the bias domains; (3) very serious limitations for studies at high risk of bias with respect to almost all of the domains.

2. Inconsistency: the quality of evidence can be downgraded if (1) the points of effect of the studies cross the line of no effect and their confidence intervals show minimal or no overlap; or (2) I^2 is substantial (≥50%).
3. Indirectness: the quality of evidence may be downgraded when: (1) the participant population; (2) the prognosis factor(s); and/or (3) the outcomes considered in the primary studies do not fully represent the review question defined in the systematic review.
4. Imprecision: the quality of evidence may be downgraded if: (1) the sample size included in the meta-analysis is insufficient; and/or (2) there is no precise estimate of the effect size in the meta-analysis, due to an excessively wide confidence interval that overlaps the value of no effect and contain values implying that the prognostic factor is associated with protection or increased risk.
5. Publication bias: it should be considered for downgrading, unless the prognostic factor has been investigated in a large number of cohort studies.

And two factors that may increase quality:

1. Moderate or large effect size: if there is a moderate or large pooled effect of the meta-analysis.
2. Exposure response gradient: exists when elevated levels of prognostic factor (amount, longevity, intensity, etc) lead to a larger effect size over lower levels of the factor.

The overall quality or research evidence was defined as ‘high’, ‘moderate’, ‘low’, or ‘very low’.

3. Results

3.1. Study selection

A total of 345 references (MEDLINE 120, EMBASE 170, and Web of Science 55) were retrieved through the electronic search (Figure 1). After automatic and manual deduplication, 233 references had their title and abstract screened. Of these, we excluded 204. The full-texts of 29 references were examined for eligibility. In total, 21 studies were excluded, 12 due to inappropriate study design for this purpose (case reports, editorials, non-systematic reviews), four had duplicated results, three related to inadequate outcomes, one due to wrong patient population, and one due to insufficient information, persisting after contact with the authors. Eight studies met the inclusion criteria and were included (Figure 1). No additional records were obtained.

3.2. Study characteristics

The study characteristics are summarized in Table 1.

3.2.1. Methods

All included studies were observational: one was cross-sectional (Pinto *et al.*, 2016), four were prospective cohorts (Jenkins *et al.*, 2016; Pinto *et al.*, 2017; Singh *et al.*, 2011; Yamauchi *et al.*, 2014), one was retrospective cohort (Pinto *et al.*, 2012), and one was a case-controls study (Sathyaprabha *et al.*, 2010). For Bokuda *et al*. (2014) only the abstract was available and it was not possible to assess if the cohort was either retrospective or prospective. All were published in English, single-centre, and set at tertiary hospitals. Three studies took place in Portugal, two in India, two in Japan, and one in the USA.

3.2.2. Participants

The included studies involved a total of 635 participants. The main inclusion criteria entailed adults (i.e. above 18 years old), able to give informed consent, with a diagnosis of definite or probable ALS, as defined by the modified *El Escorial* criteria. Exclusion criteria included patients with implantable pacemakers, cardiac insufficiency, lung disease, polyneuropathy,

diabetes mellitus, dementia, or malignancy. The participants' mean age ranged between 49.7 and 61.5 years, with a mean disease-duration between 16.0 and 27.6 months. Bulbar-onset ranged between 22.5% and 41.9% of the participants. In two studies (Jenkins *et al.*, 2016; Yamauchi *et al.*, 2014) participants initiated non-invasive ventilation during the duration of the study.

3.2.3. Index and reference tests

All participants in all included studies underwent PFT. From these, FVC was assessed by spirometry in the sitting up position, and expressed as a percentage of the predicted lung capacity adjusted for gender, weight, height and race.

PNC techniques were comparable across studies. They were performed with percutaneous bipolar electrical stimulation of the phrenic nerve at the neck level (posterior to the lateral border of the sternocleidomastoid muscle), with the exception of one study, Yamauchi *et al.*, (2014), in which stimulation was applied at the supraclavicular fossa. Recording through surface electrodes on xiphoid process (active electrode) and costal margin of the midclavicular line (reference electrode). The ground electrode was placed over the sternum or ipsilateral arm. The latency of CMAP was measured from the stimulus to the onset of potentials and expressed in milliseconds (ms). The amplitude was measured through peak-to-peak of CMAP, and expressed in miliVolt (mV).

Clinical evaluation was performed with the ALS-Functional Rating Scale-Revised (ALS-FRS-R), including the respiratory subscore (0-12, which includes dyspnoea, orthopnoea, need of ventilatory assistance).

3.3. Outcomes

Three studies reported the hazard ratios for mortality (Bokuda *et al.*, 2014; Jenkins *et al.*, 2016; Pinto *et al.*, 2012). However, one of these did not contained the enough information to be included in the analysis (Jenkins *et al.*, 2016). Seven studies correlated the amplitude of CMAP with FVC (Bokuda *et al.*, 2014; Jenkins *et al.*, 2016; Pinto *et al.*, 2016; Pinto *et al.*,

2017; Sathyaprabha *et al.*, 2010; Singh *et al.*, 2011; Yamauchi *et al.*, 2014). Three studies correlated the latency of CMAP and FVC (Pinto *et al.*, 2017; Sathyaprabha *et al.*, 2010; Singh *et al.*, 2011). For Pinto *et al.* (2017) study, correlation coefficients for FVC-CMAP amplitude, and FVC-CMAP latency were obtained through contact with the authors.

3.4. Synthesis of results

In the pooled analysis, participants with CMAP amplitude equal or below 0.4mV are 2.021 more likely to die over the studied period (95%CI 1.161 to 3.522; I^2=55.9%; 2 studies; 338 participants, Figure 2).

In the pooled analysis, amplitude of CMAP showed a moderate (Cohen, 1988) positive correlation with FVC (r=0.400, 95%CI= 0.226 to 0.550; I^2=69.77%, 7 studies, 381 participants, Figure 3). On the other hand, there was a weak (Cohen, 1988) negative correlation between latency of CMAP and FVC (r=-0.235; 95%CI= -0.447 to -0.024; I^2=15.92%; 3 studies; 112 participants, Figure 4).

3.5. Risk of bias across studies

The overall risk of bias across studies was moderate to high (Figure 5).

For study participation the risk of bias was moderate, with high risk in one study (Bokuda *et al.*, 2014) due the lack of information, and five studies with a moderate risk (Pinto *et al.*, 2016; Pinto *et al.*, 2017; Sathyaprabha *et al.*, 2010; Singh *et al.*, 2011; Yamauchi *et al.*, 2014). The information about the adequacy of study participation by eligible individuals was not available in none of the eight included studies. The period (Bokuda *et al.*, 2014; Pinto *et al.*, 2016; Pinto *et al.*, 2017) and the place of recruitment (Bokuda *et al.*, 2014; Pinto *et al.*, 2017; Yamauchi *et al.*, 2014) were not described in three studies.

The risk of attrition bias in the longitudinal studies was high, mainly because none of the studies reported information about the participants who drop-out/were lost to follow-up.

The prognostic factor measurement item presented a moderate risk of bias. None of the longitudinal studies included reported a method to incorporate missing data, and the proportion of the data on the prognostic factor available for analysis was unknown in two studies (Bokuda *et al.*, 2014; Pinto *et al.*, 2017).

The risk of bias concerning outcome measurement was low. Nevertheless, method and setting of outcome measurement was not the same in all participants in study by Jenkins *et al.* (2016). It was not clear if the method and setting of outcome measurement was the same in all participants in Singh *et al.* report (2011), as patient with bulbar subtype were submitted to a week of training to improve the technical quality of their PFT.

The study confounding risk of bias was high, as none of the eight included studies incorporated all possible confounders (i.e. respiratory symptoms, type of onset, ALSFRS-R, use of NIV) in the study design or data analysis.

Finally, the risk of bias concerning the study analysis and reporting was high, as it was not clear if the conceptual model framework or the statistical analysis of the eight studies were appropriate to answer our question.

3.6. Quality of evidence

The overall quality of evidence was low to moderate. Table 2 details the GRADE approach to the quality of the available evidence. For all three outcomes, the quality of evidence was downgraded due to the fact that the evidence comes from studies with moderate to high risk of bias. For mortality and correlation between FVC and CMAP amplitude, the quality of evidence was downgraded due to a significant heterogeneity across the studies, ascertained in the pooled analysis (I^2 = 55.9% and I^2= 69.77%, respectively). Additionally, for the correlation between FVC and CMAP latency, the quality of evidence was downgraded due to insufficient sample size included in the meta-analysis (Nº of participants = 112). The quality of evidence was upgraded for mortality due to high effect size in the meta-analysis (HR 2.02).

4. Discussion

Overall, we found a low-to-moderate quality evidence suggesting that PNC has meaningful value in ALS. This conclusion is based on eight observational studies, enrolling a total of 635 patients with ALS. PNC probably predicts mortality in ALS. Only two of the eight included studies, enrolling 338 patients, reported hazard ratios for mortality (Bokuda *et al.*, 2014; Pinto *et al.*, 2012). Our pooled analysis showed an approximately two-fold higher probability of death in patients with a CMAP amplitude below the 0.4mV cut-off (the same cut-off value for CMAP amplitude was applied in both studies). Similarly, an abnormal FVC value (<75%) was found to be significantly associated with poor survival, with a Hazard ratio of 1.68 (Czaplinski et al., 2006), and a vital capacity decline of 10% increased mortality risk in ALS, with a Hazard ratio of 1.31 (Enache et al., 2017). Significant statistical heterogeneity (I^2 55.9%) was found among these studies, which was probably driven by differences in methodological and clinical features between studies. One of them (Bokuda *et al.*, 2014) was published as an abstract leading to a high risk of bias, due to a lack of information regarding the study design and data analysis. Both studies used amplitude of CMAP for the survival analysis, which seems to be associated with a more accurate prediction of hypoventilation and survival than latency (Pinto *et al.*, 2012; Pinto *et al.*, 2009b). Although, no single test has been shown to correlate well with respiratory failure in ALS (Miller *et al.*, 1999), PFT are the most used measurements in clinical practice and research, in particular FVC. Therefore, in this systematic review we investigated the correlation between the FVC and PNC parameters (amplitude and latency), as a complementary tool to predict outcome in ALS patients. We found moderate evidence that CMAP amplitude is positively correlated with FVC, and that latency tend to be negatively correlated with FVC. Significant statistical heterogeneity (I^2=69.77%) was found across studies reporting the correlation between PNC amplitude and FVC. This was probably driven by differences in the patient population and methods between studies.

Other PFT, such as maximal voluntary ventilation, maximal inspiratory and expiratory pressures, and nasal inspiratory pressure during sniff (SNIP), since data was not available. However, maximal expiratory pressures decline rapidly in early stages of the disease, and both are strongly influenced by orofacial paresis (Kaplan and Hollander, 1994). SNIP, despite being more suited for ALS patients with orofacial paresis, it is only predictive of hypoventilation in spinal-onset patients (Morgan *et al.*, 2005).

The main limitation of our contribution is the heterogeneity in the objectives, methods, and data reporting in the primary studies. Variable clinical features of the population samples, such as respiratory symptoms, functional impairment, duration of the disease, region of onset, or NIV initiation, may contribute to the heterogeneity of the results. For instance, in Jenkins *et al* (2016) study, which was the study with the highest correlation between FVC and PNC CMAP amplitude (r=0.69), patients presented a wide range of disease duration (2-102 months) at evaluation, and around 21% patients were using NIV at the time of the study, which may possibly influence the results. An additional limitation, regards the meta-analysis on phrenic nerve response as predictor of mortality, which incorporated one study including the last author of the current report, but this specific author did not participate in following steps of the current analysis: study selection, data extraction, and in the assessment of the risk of bias. Sub-group analyses were not performed due to a lack of data of the individual studies. However, it would be interesting to compare spinal and bulbar-onset subgroups regarding the impact of PCN results on prognosis. This would be relevant mainly for the bulbar-onset subgroup, since these patients are results from PFT are less reliable.

PNC can complement conventional respiratory tests in ALS. FVC and slow vital capacity are the standard tests to evaluate and monitor respiratory function. However, it is well accepted that in patients with facial weakness and cognitive decline, test results are affected by air leak and poor collaboration, respectively (de Carvalho et al., 2019). We propose that PNC in ALS would be of particular importance in patients with bulbar involvement or cognitive dysfunction, since results from conventional respiratory tests are unreliable. It would be

relevant to analyse the potential impact of PNC results on NIV initiation time, taking into account that NIV improves survival in ALS (Bourke *et al.*, 2006), proportionally to duration of use (Pinto *et al.*, 1995), and has a positive influence on quality of life (Bourke *et al.*, 2003; 2006).

Overall, our results favour the utility of PNC in ALS, but further research is needed in order to validate this technique as a biomarker of pulmonary function in ALS. Larger cohort studies with the complete follow-up, unbiased case selection and complete ascertainment of the possible confounders need to be performed in order to ascertain the capacity of PNC in predicting ALS outcome. Regarding the technique itself, like other electrophysiological techniques, PNC is highly operator-dependent and its results depend on the quality of its performance and standardization of the technique is essential for the interpretation of the results.

REFERENCES

Bokuda K, Shimizu T, Hayashi K, Yamazaki T, Nakano I. Phrenic nerve motor responses and survival in amyotrophic lateral sclerosis. Clin Neurophysiol 2014;125:S292.

Borenstein M, Hedges LV, Rothstein H, Higgins J. Effect Sizes Based on Correlations. Introduction to Meta-Analysis. John Wiley & Sons, Ltd; 2009.

Bourke SC, Bullock RE, Williams TL, Shaw PJ, Gibson GJ. Noninvasive ventilation in ALS: indications and effect on quality of life. Neurology 2003;61:171-7.

Bourke SC, Tomlinson M, Williams TL, Bullock RE, Shaw PJ, Gibson GJ. Effects of non-invasive ventilation on survival and quality of life in patients with amyotrophic lateral sclerosis: a randomised controlled trial. Lancet Neurol 2006;5:140-7.

Brooks BR, Miller RG, Swash M, Munsat TL. El Escorial revisited: revised criteria for the diagnosis of amyotrophic lateral sclerosis. Amyotroph Lateral Scler Other Motor Neuron Disord 2000;1:293-9.

Cohen J. Statistical Power Analysis for the Behavioral Sciences. 2nd ed. USA: Lawrence Erlbaum Associates 1988.

Czaplinski A, Yen AA, Appel SH. Forced vital capacity (FVC) as an indicator of survival and disease progression in an ALS clinic population. J Neurol Neurosurg Psychiatry 2006;77:390-2.

de Carvalho M, Matias T, Coelho F, Evangelista T, Pinto A, Luis ML. Motor neuron disease presenting with respiratory failure. J Neurol Sci 1996;139(Suppl):117-22.

de Carvalho M, Pinto S, Swash M. Diaphragm motor responses to phrenic nerve stimulation in ALS: surface and needle recordings. Clin Neurophysiol 2018;129:349-53.

de Carvalho M, Pinto S, Swash M. Diaphragmatic neurophysiology and respiratory markers in ALS. Front Neurol 2019;10:143.

Enache I, Pistea C, Fleury M, Schaeffer M, Oswald-Mammosser M, Echaniz-Laguna A, Tranchant C, Meyer N, Charloux A. Ability of pulmonary function decline to predict death in amyotrophic lateral sclerosis patients. Amyotroph Lateral Scler Frontotemporal Degener 2017;18:511-8.

Fallat RJ, Jewitt B, Bass M, Kamm B, Norris FH, Jr. Spirometry in amyotrophic lateral sclerosis. Arch Neurol 1979;36:74-80.

Haverkamp LJ, Appel V, Appel SH. Natural history of amyotrophic lateral sclerosis in a database population. Validation of a scoring system and a model for survival prediction. Brain 1995;118:707-19.

Hayden JA, Cote P, Bombardier C. Evaluation of the quality of prognosis studies in systematic reviews. Annals Int Med 2006;144:427-37.

Hayden JA, van der Windt DA, Cartwright JL, Cote P, Bombardier C. Assessing bias in studies of prognostic factors. Annals Int Med 2013;158:280-6.

Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. Statistics in medicine 2002;21:1539-58.

Huguet A, Hayden JA, Stinson J, McGrath PJ, Chambers CT, Tougas ME*, et al.* Judging the quality of evidence in reviews of prognostic factor research: adapting the GRADE framework. Syst Rev 2013;2:71.

Jenkins JAL, Sakamuri S, Katz JS, Forshew DA, Guion L, Moore D*, et al.* Phrenic nerve conduction studies as a biomarker of respiratory insufficiency in amyotrophic lateral sclerosis. Amyotroph Lateral Scler Other Motor Neuron Disord 2016;17:213-20.

Kaplan LM, Hollander D. Respiratory dysfunction in amyotrophic lateral sclerosis. Clin Chest Med 1994;15:675-81.

Lang TA, Altman DG. Basic statistical reporting for articles published in biomedical journals: the "Statistical Analyses and Methods in the Published Literature" or the SAMPL Guidelines. Int J Nurs Stud 2015;52:5-9.

Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gotzsche PC, Ioannidis JP*, et al.* The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. Annals Int Med 2009;151:W65-94.

Miller RG, Munsat TL, Swash M, Brooks BR. Consensus guidelines for the design and implementation of clinical trials in ALS. World Federation of Neurology committee on Research. J Neurol Sci1999;169:2-12.

Moher D, Shamseer L, Clarke M, Ghersi D, Liberati A, Petticrew M*, et al.* Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. Syst Rev 2015;4:1.

Morgan RK, McNally S, Alexander M, Conroy R, Hardiman O, Costello RW. Use of Sniff nasal-inspiratory force to predict survival in amyotrophic lateral sclerosis. Am J Respir Crit Care Med 2005;171:269-74.

Paulukonis ST, Roberts EM, Valle JP, Collins NN, English PB, Kaye WE. Survival and Cause of Death among a Cohort of Confirmed Amyotrophic Lateral Sclerosis Cases. PloS one 2015;10:e0131965.

Phukan J, Pender NP, Hardiman O. Cognitive impairment in amyotrophic lateral sclerosis. Lancet Neurol 2007;6:994-1003.

Pinto AC, Evangelista T, de Carvalho M, Alves MA, Sales Luis ML. Respiratory assistance with a non-invasive ventilator (Bipap) in MND/ALS patients: survival rates in a controlled trial. J Neurol Sci 1995;129 (Suppl):19-26.

Pinto S, Alves P, Pimentel B, Swash M, de Carvalho M. Ultrasound for assessment of diaphragm in ALS. Clin Neurophysiol 2016;127:892-7.

Pinto S, Alves P, Swash M, de Carvalho M. Phrenic nerve stimulation is more sensitive than ultrasound measurement of diaphragm thickness in assessing early ALS progression. Neurophysiol Clin 2017;47:69-73.

Pinto S, de Carvalho M. Breathing new life into treatment advances for respiratory failure in amyotrophic lateral sclerosis patients. Neurodegener Dis Manag 2014;4:83-102.

Pinto S, Geraldes R, Vaz N, Pinto A, de Carvalho M. Changes of the phrenic nerve motor response in amyotrophic lateral sclerosis: longitudinal study. Clin Neurophysiol 2009a;120:2082-5.

Pinto S, Pinto A, de Carvalho M. Phrenic nerve studies predict survival in amyotrophic lateral sclerosis. Clin Neurophysiol 2012;123:2454-9.

Pinto S, Turkman A, Pinto A, Swash M, de Carvalho M. Predicting respiratory insufficiency in amyotrophic lateral sclerosis: the role of phrenic nerve studies. Clin Neurophysiol 2009b;120:941-6.

Sathyaprabha TN, Pradhan C, Nalini A, Thennarasu K, Raju TR. Pulmonary function tests and diaphragmatic compound muscle action potential in patients with sporadic amyotrophic lateral sclerosis. Acta Neurol Scand 2010;121:400-5.

Singh D, Verma R, Garg RK, Singh MK, Shukla R, Verma SK. Assessment of respiratory functions by spirometry and phrenic nerve studies in patients of amyotrophic lateral sclerosis. J Neurol Sci 2011;306:76-81.

Stroup DF, Berlin JA, Morton SC, Olkin I, Williamson GD, Rennie D*, et al.* Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group. JAMA 2000;283:2008-12.

Yamauchi R, Imai T, Tsuda E, Hozuki T, Yamamoto D, Shimohama S. Respiratory insufficiency with preserved diaphragmatic function in amyotrophic lateral sclerosis. Intern Med 2014;53:1325-31.

Table 1 - Studies characteristics. n, number of participants; SD, standard deviation; N/S, not specified; ALSFRSR, Amyotrophic Lateral Sclerosis Functional Rating Scale-Revised; NIV, non-invasive ventilation; *Age at onset

Study	Country	Study design	n	Age (years) mean ± SD (range)	Male n (%)	Disease duration (months) mean ± SD (range)	ALSFSR-R mean ±SD (range)	Bulbar onset n (%)	NIV
Bokuda 2014	Japan	Longitudinal	84	N/S	N/S	N/S	N/S	N/S	N/S
Jenkins 2016	USA	Longitudinal prospective	100	60.8 (34-84)	32 (43)	23.9 (2–102)	32.5 (8–45)	31 (31.0)	Yes
Pinto 2016	Portugal	Cross-sectional	42	58.4± 11.1 (34-77)*	20 (47.6)	17.8 ± 13.6 (5-58)	34.98 ± 3.1 (27–39)	11 (26.2)	No
Pinto 2017	Portugal	Longitudinal prospective	40	58.2 ± 10.0 (37- 77)*	20 (50.0)	16.0 ± 11.9	35.03 ± 3.4	9 (22.5)	No
Pinto 2012	Portugal	Longitudinal retrospective	254	60.9 ± 11.2 (28-80)	132 (52.0)	N/S	31.3 ± 5.6 (13–40)	79 (31.1)	No
Sathyaprabha 2010	India	Case-control	29	51.4 ±10.7 (30–68)	20 (69.0)	27.6 ± 34.3	N/S	8 (27.6)	No
Singh 2011	India	Longitudinal prospective	43	49.7±14.9	32 (74.4)	16.3 ± 15.7	44	18 (41.9)	No
Yamauchi 2014	Japan	Longitudinal prospective	43	61.5± 12.8	21 (48.8)	16.4 ± 9.8	31.77± 7.2	(51.2)*	Yes

Table 2 - GRADE table summary findings. CI, Confidence interval; HR, Hazard ratio; r, correlation coefficient; *, bulbar involvement. Explanations: a. Downgraded due to study limitations; b. Downgraded due to inconsistency: large I^2; c. Downgraded due to imprecision: optimal information size not met; d. Upgraded due to large effect size.

Outcome **№ of participants (studies)**	**Relative effect** **(95% CI)**	**Certainty**	**What happens**
Mortality **№ of participants: 338 (2 studies)**	HR 2.02 (1.16 to 3.52)	⨁⨁⨁◯ MODERATE [a, b, d]	We have moderate confidence that PNC is associated with a 102% relative increase in predicting mortality. PNC probably predicts mortality.
Correlation of FVC and CMAP amplitude **№ of participants: 381 (7 studies)**	r 0.40 (0.23 to 0.55)	⨁⨁⨁◯ MODERATE [a, b]	We have moderate confidence that CMAP is positively correlated with FVC. CMAP amplitude levels on PNC are probably positively correlated with FVC.
Correlation of FVC and CMAP latency **№ of participants: 112 (3 studies)**	r -0.23 (-0.42 to -0.02)	⨁⨁◯◯ LOW [a, c]	We have low confidence that latency is negatively correlated with FVC. CMAP latency on PNC may be negatively correlated with FVC.

Figures