

The Role of Facemask Spirometry in Motor Neurone Disease (MND)

Banerjee SK¹, Davies MG¹, Sharples LD², Smith IE¹

¹ Respiratory Sleep and Support Centre, Papworth Hospital NHS Trust, Papworth Everard, Cambridgeshire, CB23 3RE, UK;

² MRC Biostatistics Unit, Robinson Way, Cambridge, CB2 0SR, UK.

On line data supplement

Introduction:

Motor Neurone Disease (MND) is a degenerative disorder affecting the cerebral cortex, brain stem and spinal cord. Fifty percent of people with MND die within 3 years of the onset of symptoms and 90% within 5 years (1). Respiratory failure is the most frequent cause of death in people with MND. Respiratory muscle weakness may cause ventilatory failure or patients may succumb to respiratory tract infection due to unsafe swallowing or an ineffective cough.

Vital capacity and maximal inspiratory and maximal expiratory mouth pressures are the methods most commonly used to assess respiratory muscle impairment (2). The Forced Vital Capacity (FVC) declines with time in patients with MND and is an indicator of disease progression such that both the FVC at diagnosis and the rate of decline in FVC are predictors of survival (3, 4). An FVC of < 50% predicted is proposed by the National Institute of Clinical Excellence (NICE, UK) as an indication of the need for evaluation for non-invasive ventilation (5). However patients with motor neurone disease who have

facial or bulbar muscle weakness may not effectively seal their lips around the tube which is usually employed in spirometry, especially during forced manoeuvres (2,6). In one study of the correlation between measures of respiratory muscle function and hypercapnia in MND, some subjects used a mask for spirometry (7). No details of the numbers using the mask were reported.

A study comparing measurements of spirometry using a conventional tube or a face mask in healthy volunteers showed the face mask to be a straight forward tool to measure respiratory function in healthy persons (8). Spirometric measurements with the tube were higher than those with the mask with a difference in FVC of 100-200 mls ($p < 0.001$). This small under reading may not be of clinical significance in normal individuals but could be important in individuals with lower volume ranges, as seen in many patients with MND. However in a pilot study using a mask in MND patients (who struggled with a tube) we have shown the converse with higher values of FVC when compared to tube spirometry (9).

The objective of the current study was to identify if mask spirometry was more or less accurate than conventional tube spirometry in an unselected MND clinic population with sub group analysis according to bulbar involvement. The second objective was to identify if it could be used interchangeably with tube spirometry in all patients.

Supplementary Methods:

Trial Design and participants

Ethical approval for the study was granted by the regional research ethics committee (Cambridgeshire 3 Research Ethics Committee, REC reference number 10/H0306/9). Informed consent was obtained from all participants. Consecutive patients with a confirmed diagnosis of MND referred for respiratory assessment or under regular follow up in a specialist respiratory support clinic were invited to participate in the study. The exclusion criteria included ventilator dependence, tracheostomy or a second neurological diagnosis. Subjects with contra indications for spirometry as per the 'ATS/ERS Task Force: Standardisation of lung function testing' were also excluded from the study (10).

FVC is the maximal volume of air exhaled with maximally forced effort from a maximal inspiration (1~~10~~). Since the true value of FVC is maximal, it should not be possible for a different technique to over estimate it providing that the equipment is correctly calibrated. In patients with MND, the "within patient standard deviation" whilst performing spirometry was assumed to be 10% ~~(11)~~. We set the sample size to show that the 95% confidence interval on the difference between the mask and tube measurements is bounded by a minimum clinically important difference of +/-5%. Thus, using established methods ~~(12)~~ with 80% power we required 42 patients with both tube and mask measurements of lung function (1~~22~~). In a pilot study, only 15 of 26 patients (58%) tested could provide spirometry using both a mask and a tube. However these subjects were selected from those who had struggled with tube spirometry (9). We assumed that 70% of our patients would be able to provide measurements using both a mask and tube so that recruitment of 60 patients would ensure that 42 were available for analysis.

To quantify bulbar involvement, we used the revised Amyotrophic Lateral Sclerosis Functional Rating Scale (ALSFRS r score), a validated rating instrument for monitoring

the progression of disability in patients with MND. It is a score from 0 to 48 and consists of 12 questions, each rated from 0 to 4. The first 3 questions are related to bulbar functions and a total score of 12 for these questions indicates no bulbar symptoms. Decreasing values on these 3 questions (b ALSFRS r score) indicate progressively greater bulbar symptoms. Fattori et al, showed clinically significant swallowing difficulty in patients when the b ALSFRS r score was less than 9 (133). In our study, we have proposed *a priori* that a score of less than 9 defines the subgroup with important bulbar symptoms.

At enrolment a full clinical history including the ALSFRS r score, was recorded for all subjects. They also had a physical examination and measurements taken for maximal inspiratory (MIP) and expiratory pressures (MEP) at the mouth and sniff nasal inspiratory pressure (SNIP) using a Micro respiratory pressure meter, handheld device (Care Fusion, Germany) (5).

Spirometry

FVC and Forced Expiratory Volume in one second (FEV₁) were measured using a Micro plus hand held spirometer (Micro Medical, Chatham, Kent). Calibration checks were performed daily with a 1-L syringe. Subjects were connected to the spirometer via a conventional tube that is a cardboard tube or via a non-compressible face mask (a Leardal child mask number 4). The spirometer with the mask was held and pressed against the subject's face by the investigator while the conventional device was held by the subject. The order in which tests were performed was randomized based on computer generated random numbers, stored in sealed envelopes, to avoid order effects. At least three measurements were performed with each technique and two

readings with ~~in 10% of each other~~ the smaller being within 10% of the larger were deemed acceptable.

Statistical Analysis

Unless otherwise stated, frequencies (percentages) and means (standard deviations) are used to summarise the data as appropriate. We report 95% confidence intervals to describe the precision of estimated mean differences between the mask and tube measurements. McNemar's test was used to compare the number of patients who were able to provide spirometry results and the number of patients with FVC>50% using the two methods. Order effects of the absolute spirometric measurements and the differences between mask and tube measurements were assessed using paired Student t-tests. Since no effects were found they were not considered in the final analysis. Paired Student t-tests were used to compare mean differences in spirometry between the two methods. To compare spirometry measurements from the two methods the analysis was based on Bland-Altman methods (144). Linear regression models were fitted to assess how well mask measurements of FVC could be predicted from the tube measurements, both overall and separately for the pre-defined subgroups defined by presence or absence of bulbar symptoms (bulbar symptoms when bulbar ALSFRS-R score < 9).

Supplementary Results:

We screened seventy three consecutive subjects for potential enrolment. Thirteen were excluded for the following reasons; 4 declined consent, 5 were on continuous NIV, 1 had developed dementia, 1 had a mini tracheostomy tube in situ, 1 had recurrent pneumothoraces and 1 had recent cataract surgery. Among the remaining 60 subjects

there were 36 (60%) men, mean age 64.7 (SD 10.5) years, mean BMI 26.7 (SD 5.5) kg/m². Thirty-three (55%) were non smokers and 27 were current smokers. The median (inter-quartile range) for time between diagnosis of MND and test completion was 558 (728) days. The mean b ALSFRS r score was 7.8 (3.3) and 30/60 (50%) subjects had a score of < 9. Seventeen (28%) subjects had been issued NIV with 14 reporting regular nocturnal use. Thirty four (57%) subjects had been prescribed riluzole.

Results were available for all 60 patients using the mask and 54 patients using the tube (McNemar p<0.001). From the total of 60 subjects, 51 (85%) managed SNIP measurements and 45 (75%) inspiratory and expiratory mouth pressure measurements.

Table 1 contains the mean FVC measured by mask and tube in the study population. Table 2 gives comparative information regarding FVC between tube and mask in subjects with varying degrees of bulbar involvement and shows greater accuracy of mask spirometry with worsening bulbar scores.

Inspection of the Bland Altman plot comparing tube and mask FVC (fig 1A) shows more points below y=0 (mask value greater) than above. The mask gave the higher reading in 36 cases, was equal to the tube in 1 and lower in 17. The mask gave significantly greater values for FVC than the tube, the mean difference was 0.19 litres (95%CI 0.07, 0.31; p=0.0024). ~~The mask gave significantly greater values for FVC than the tube, the mean difference was 0.19 litres (95%CI 0.07, 0.31; p=0.0024). On linear regression the difference between the FVC measurements of tube and mask decreased as the mean FVC increased, with the two measurements coinciding at about 3.3 litres (difference in FVC = -0.56 + 0.17*average in FVC). Both the intercept and slope were significantly~~

~~different from zero suggesting that the relative accuracy of the two methods varies with the value of the FVC measurement.~~

The scatter plot (Fig 1B2) shows FVC results for the mask against those for the tube, with a 45° line of agreement super-imposed. The mask gave higher measurements of FVC for patients with low FVC values but slightly lower values for patients with higher FVC values. A linear regression was fitted to predict mask results from tube results. The resulting regression equation was:

$$\text{Mask (litres)} = 0.61 + 0.80 \text{ Tube (R}^2 = 90\%) \text{ (litres)}$$

The intercept was significantly greater than zero, confirming that the estimated mask records higher FVC for low values, and the slope was significantly less than one, so that estimated mask measurements are lower for values of FVC above approximately 3 litres.

Effect of bulbar disease (subgroup analysis)

All 30 subjects with bulbar ALSFRS-R score ≥ 9 were able to record both mask and tube FVC measurements. Six of the 30 subjects with b-ALSFRS r score < 9 could not perform tube measurements, but all 30 had a result using the mask (McNemar $p < 0.0001$). Five patients with preserved bulbar function and 4 patients with B ALSFRS r score < 9 (86%) were unable to record SNIP.

For patients with both FVC measurements the mask was higher or equal in 21/24 (88%) of cases with bulbar disease and 15/30 (50%) with non-bulbar disease (Fisher's exact test 0.004). Figure [1C3](#) shows the difference between measurements plotted against the average, split by whether or not the patients have bulbar disease (bulbar ALSFRS-R score < 9, right) or not (bulbar ALSFRS-R score ≥ 9, left). Patients with bulbar disease had higher mean FVC measurements with the mask but the non-bulbar group did not. The mean difference between mask and tube FVC measurements was 0.32 litres (95%CI 0.15, 0.49), p=0.001 in the bulbar group and 0.09 litres (95%CI -0.08, 0.25) in the non-bulbar group, p=0.283.

FVC measured using the face mask was preferred by 44 (73%) subjects irrespective of bulbar or non- bulbar involvement.

Supplementary Discussion:

Respiratory muscle strength is a key prognostic factor in motor neurone disease. A decline in respiratory muscle strength assessed by vital capacity, mouth pressures and SNIP has been shown to correlate with time to death (1). Survival may be only a few weeks when patients with MND develop ventilatory failure ([155](#), [166](#)) but prompt treatment with NIV can improve survival for patients with predominantly non bulbar disease, and improve quality of life in the remainder ([145](#)). Present NICE guidelines propose surveillance for ventilator failure every 3 months. ~~A balance needs to be struck between frequent clinic visits, which are disruptive to the patient, and the risk of missing the opportunity to start NIV, which may prolong survival and improve quality of life.~~ Better predictors of ventilator failure could allow better planning of such follow up visits.

In the past, tube and mask interfaces have been treated as interchangeable (7, 174). However a study of spirometry in healthy volunteers showed measures with a tube were higher than with a mask with a difference in FVC of 100-200 mls (8). It was concluded that measurements obtained using the conventional tube and the mask could not be used interchangeably and that new reference values were needed for the mask interface. We have found that tube and mask connections to the spirometer cannot be used interchangeably in people with MND.

In our study the mask gives higher values than the tube when FVC is less than 3 litres (bulbar and non-bulbar) and for all patients with moderate to severe bulbar involvement regardless of the value of FVC. All 60 patients were able to perform FVC using a mask, whilst only 54 (90%) managed with the tube. Six patients, who produced a mean FVC of 1.45 (0.68) litres with the mask could not record anything with tube spirometry. ~~However at higher forced expiratory volumes, tube spirometry tended to produce higher values.~~ The potential under reading of FVC at higher volumes means that mask cannot completely replace tube as an interface for measuring spirometry in patients with MND.

Other authors have proposed that SNIP measurements predict mortality in MND better than spirometry (6). In this study, ~~where only conventional tube spirometry was performed~~ SNIP was achieved by 96% of subjects and spirometry by only 84%, within 6 months of death. ~~This was not the experience of another group using mask spirometry (7).~~ In our study 100% of subjects could perform mask spirometry compared to just 85% for SNIP. It is of note that bulbar involvement can interfere with SNIP measures. In our study and that of Chaudri (7) the percentage with bulbar involvement was 50% (moderate to severe) and 52.5% respectively while in the study with better results for SNIP the proportion with bulbar involvement was just 31% (6).

Ventilatory failure is the most frequent cause of death in MND. Measuring pulmonary function as an indicator of diaphragm function provides prognostic information but doubt about the likely time to developing ventilator failure mandates frequent surveillance visits for these patients. Whilst it might be that no single test will predict accurately when assisted ventilation will be needed (2), a more accurate measure of lung function in this patient population may contribute towards the goal of instituting NIV at the appropriate time. In Duchenne's Muscular Dystrophy (DMD), patients started prematurely on NIV were associated with an adverse outcome (18). Whilst no data for MND exists, one needs to be aware of this potential adverse outcome. Further studies to identify whether mask spirometry has better predictive power for time to ventilatory failure and the need for NIV are justified.

Supplementary References:

1. Stambler N, Charatan M et al. Prognostic indicators of survival in ALS. *Neurology* 1998; 50:66-72.
2. Lyall RA, Donaldson N, Polkey MI, Leigh PN and Moxham J. Respiratory muscle strength and ventilatory failure in amyotrophic lateral sclerosis. *Brain* (2001), 124, 2000-2013.
3. Schmidt E, Drachman D, Wiener C et al. Pulmonary predictors of survival in amyotrophic lateral sclerosis: use in clinical trial design. *Muscle & Nerve* 2006; 33: 127-132.
4. Czaplinski A, Yen A and Appel S. Forced vital capacity as an indicator of survival and disease progression in an ALS clinic population. *J Neurol Neurosurg Psychiatry* 2006; 77: 390-92.

5. NICE clinical guideline 105. Motor neurone disease: the use of non-invasive ventilation in the management of motor neurone disease, July 2010.
<http://www.nice.org.uk/guidance/CG105/>.
6. Morgan R, McNally S, Alexander M et al. Use of Sniff Inspiratory Force to predict survival in amyotrophic lateral sclerosis. *Am J Respir Crit Care Med* 2004; 171: 269-274.
7. Chaudri MB, Liu C, Kinnear WJ et al. Sniff nasal inspiratory pressure as a marker of respiratory function in motor neurone disease. *Eur Respir J* 2000; 15: 539-42.
8. Wohlgenuth M, Van der Kooi E, Hendriks J et al. Face mask spirometry and respiratory pressures in normal subjects. *Eur Respir J* 2003; 22: 1001-1006.
9. Khaliq AA, Smith IE. Facemask spirometry in patients with bulbar motor neurone disease. *Thorax* 2009; 64: A43-A46.
- ~~10. Cooper BG. An update on contraindications for lung function testing. *Postgrad Med J*. 2011 Oct;87(1032):724-33.~~
- ~~10-11. Miller MR, Hankinson J, Brusasco V et al., Crapo R and Viegi G. 'ATS/ERS Task force: Standardisation of lung function testing'. *Eur Respir J* 2005; 26: 319–338.~~
- ~~11. Polkey MI, Green M, Moxham J. Measurement of respiratory muscle strength. *Thorax* 1995; 50: 1131-35.~~
12. Jones B, Jarvis P, Lewis J A, Ebbutt A F. Trials to assess equivalence: the importance of rigorous methods. *BMJ* 1996; 313:36-39.
13. Fattori B, Grosso M, Bongioanni P et al. Assessment of swallowing by oropharyngo-esophageal scintigraphy in patients with amyotrophic lateral sclerosis. *Dysphagia* 2006; 280-86.
14. Bland J M, Altman D G. Statistical methods for assessing agreement between two methods of clinical measurement. 1986 *Lancet*; I: 307-310.

15. Bourke S, Tomlinson M, Williams T, Bullock R, Shaw P, Gibson G. 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.
16. Mustafa N, Walsh E, Bryant V et al. The effect of non-invasive ventilation on ALS patients and their caregivers. *Neurology* 2006; 66: 1211-1217.
17. Polkey MI, Green M, Moxham J. Measurement of respiratory muscle strength. *Thorax* 1995; 50: 1131-35.
18. Raphael J-C, Chevret S, Chastang C et al. Randomised trial of preventive nasal ventilation in Duchenne muscular dystrophy. *Lancet* 1994; 343: 1600-1604.

Supplementary Table 1: Descriptive data: Mean and standard deviation of FVC

| | FVC mask | FVC mask (patients who also had a tube measurement) | FVC tube | Paired comparison between mask and tube |
|--------------------------------|-------------------|---|-------------------|--|
| No of cases | 60 | 54 | 54 [*] | |
| Mean (litres) | 2.24 [▫] | 2.32 [*] | 2.13 [*] | P=0.0024 |
| Standard Deviation (litres) | 1.10 | 1.11 | 1.30 | |
| Minimum (litres) | 0.58 | 0.58 | 0.37 | |
| Maximum (litres) | 4.84 | 4.84 | 5.20 | |

^{*}Six patients had measurements of FVC equal to zero using the tube. These have been coded as missing.

[▫] Mean FVC of 60 patients

^{*} Mean FVC of 54 patients

Supplementary **Table 2: Comparative data based on Bulbar ALSFRS r score**

| B ALSFRS r | N | | FVC mask | FVC tube | P value |
|------------|----|--------------------|----------|----------|---------|
| 12 | 8 | Mean (Litres) | 3.43 | 3.54 | 0.414 |
| | | Standard Deviation | 1.37 | 1.57 | |
| | | | | | |
| ≥ 9 | 30 | Mean (Litres) | 2.69 | 2.60 | 0.283 |
| | | Standard Deviation | 1.16 | 1.34 | |
| | | | | | |
| < 9 | 24 | Mean(Litres) | 1.88 | 1.56 | <0.001 |
| | | Standard Deviation | 0.87 | 1.01 | |
| | | | | | |

Fig 1 A: Bland Altman plot of FVC measurements using mask and tube in litres

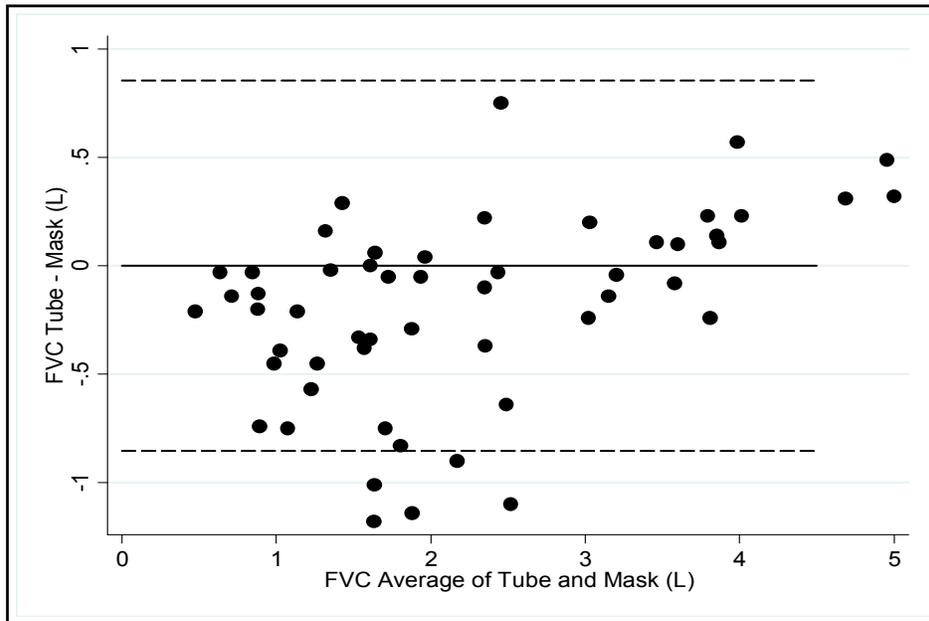


Figure 1 A. Bland Altman plot showing mask readings were higher in 36 cases, equal to the tube in 1 and lower in 17. The overall mean difference was 0.19 litres (95%CI 0.07, 0.31; $p=0.0024$) in favour of the mask. In five subjects the difference was greater than 1 litre.

Fig 1 B: Scatter plot of spirometry results for mask against tube (measurements in litres)

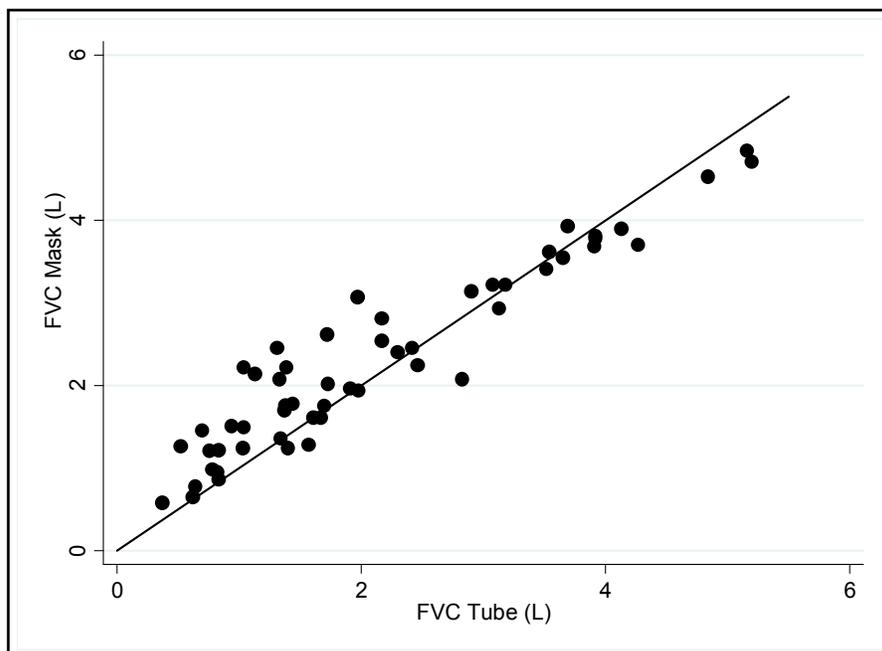


Figure 1 B. Linear regression fitted to predict mask results from tube results gives: Mask = $0.61 + 0.80\text{Tube}$ ($R^2 = 90\%$). The intercept is > 0 and slope < 1 , i.e. estimated mask FVC is higher for low values and lower for high values (above approximately 3 litres).

Fig 1 C: Bland Altman plot comparing bulbar with non bulbar measurements of FVC using mask and tube.

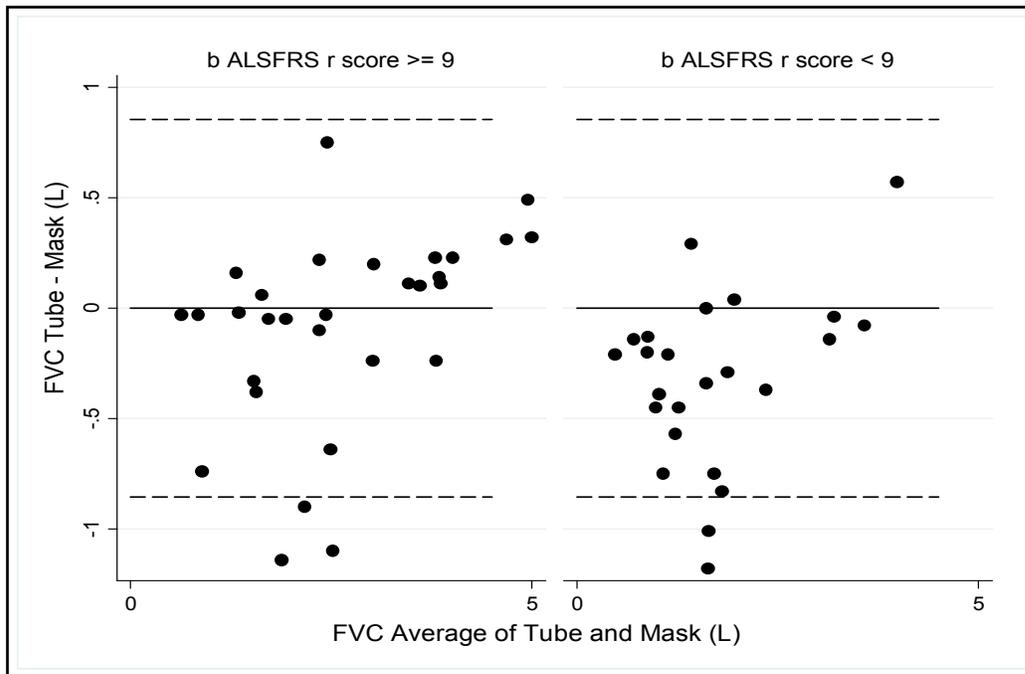


Figure 1 C. For patients with both FVC measurements the mask value was higher or equal in 21/24 (88%) of cases with bulbar disease and 15/30 (50%) with non-bulbar disease (Fisher's exact test 0.004).