

Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration

ISSN: 2167-8421 (Print) 2167-9223 (Online) Journal homepage: <https://www.tandfonline.com/loi/iafd20>

Implications of spirometric reference values for amyotrophic lateral sclerosis

Ruben P.A. van Eijk, Jaap N.E. Bakers, Michael A. van Es, Marinus J.C. Eijkemans & Leonard H. van den Berg

To cite this article: Ruben P.A. van Eijk, Jaap N.E. Bakers, Michael A. van Es, Marinus J.C. Eijkemans & Leonard H. van den Berg (2019) Implications of spirometric reference values for amyotrophic lateral sclerosis, *Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration*, 20:7-8, 473-480, DOI: [10.1080/21678421.2019.1634736](https://doi.org/10.1080/21678421.2019.1634736)

To link to this article: <https://doi.org/10.1080/21678421.2019.1634736>



© 2019 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.



Published online: 04 Jul 2019.



Submit your article to this journal [↗](#)



Article views: 408



View related articles [↗](#)



View Crossmark data [↗](#)

RESEARCH ARTICLE

Implications of spirometric reference values for amyotrophic lateral sclerosis

RUBEN P.A. VAN EIJK^{1,2}, JAAP N.E. BAKERS³, MICHAEL A. VAN ES¹,
MARINUS J.C. EIJKEMANS² AND LEONARD H. VAN DEN BERG¹

¹Department of Neurology, Brain Centre Rudolf Magnus, University Medical Centre Utrecht, Utrecht, the Netherlands, ²Biostatistics & Research Support, Julius Centre for Health Sciences and Primary Care, University Medical Centre Utrecht, Utrecht, the Netherlands, and ³Department of Rehabilitation, Brain Centre Rudolf Magnus, University Medical Centre Utrecht, Utrecht, the Netherlands

Abstract

Objective: Spirometry is commonly used as screening tool for respiratory insufficiency in neuromuscular diseases. Despite the well-known effects of reference standards on spirometric outcomes, its standardization is overlooked in current guidelines. We aim to illustrate the effect of spirometric reference values on prognostication, medical decision-making, and trial eligibility in the applied setting of amyotrophic lateral sclerosis (ALS). **Methods:** We selected 4,651 patients with 32,022 FVC measurements from the PRO-ACT dataset. The FVC estimates were standardized according to five reference standards: Knudson '76, Knudson '83, ECSC, NHANES III, and GLI-2012. (Generalized) linear mixed-effects and Cox proportional hazard models were used to evaluate longitudinal patterns and time-to-event outcomes. **Results:** The mean population %predicted FVC varied between 78.5% (95% CI 78.0–79.1) and 88.5% (95% CI 87.9–89.1). The unstandardized liters provided the worst fit on the survival data (AIC 20573, c-index 0.760), whereas the GLI provided the best fit (AIC 20374, c-index 0.780, $p < 0.001$). The mean population rate of decline in %predicted FVC could vary as much as 11.4% between reference standards. The median time-to-50% predicted FVC differed by 2.9 months between recent (14.5 months, 95% CI 14.4–16.1) and early reference standards (17.4 months, 95% CI 16.1–18.2). **Conclusion:** Independent of technique, device, or evaluator, spirometric reference values affect the utility of spirometry in ALS. Standardization of reference values is of the utmost importance to optimize clinical decision-making, improve prognostication, enhance between-center comparison and unify patient selection for clinical trials.

Keywords: Amyotrophic lateral sclerosis, forced vital capacity, spirometric reference standards, global lungs initiative

Introduction

Reduced pulmonary function is a common feature in COPD (1), heart failure (2), and neuromuscular disorders (3). Spirometric reference standards are used to compare a patient's pulmonary function to a healthy population of a similar age, body height, sex, and ethnicity (4). To illustrate, two patients with identical forced vital capacities (FVCs, e.g. 3.84L) may have different predicted FVCs (e.g. 3.25L and 5.75L) due to morphological differences, which result in different standardized measures (i.e. %predicted FVC of 118% and 67%, respectively). Absolute spirometric measurements

(i.e. liters) are, consequently, of little value in clinical practice. Importantly, the reference value for an individual patient depends heavily on the applied reference standard (5). As a consequence, variability between reference standards may considerably affect the patients' spirometric results and alter clinical decisions (6).

Spirometry is commonly used as screening tool for respiratory insufficiency in progressive neuromuscular diseases such as amyotrophic lateral sclerosis (ALS) (7). The end stage of ALS is characterized by weakening of the diaphragm

Correspondence: Leonard H. van den Berg, Department of Neurology, Brain Centre Rudolf Magnus, University Medical Centre Utrecht, Heidelberglaan 100, 3584 CX, Utrecht, the Netherlands. Tel: +31 (0) 88 75 554 94. Fax: +31 (0) 88 75 554 94. E-mail: lberg@umcutrecht.nl

(Received 22 April 2019; revised 3 June 2019; accepted 13 June 2019)

ISSN 2167-8421 print/ISSN 2167-9223 online © 2019 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way.

DOI: 10.1080/21678421.2019.1634736

musculature, which results in a progressive decline in lung function and finally death. Spirometry plays, therefore, a central role in the clinical care for patients with ALS and is a primary predictor of survival time (8). Specific cutoff values for spirometric measures are used to guide the timing of noninvasive ventilation (NIV) (7), as decision tool for gastrostomy placement or as eligibility criterion for clinical trial participation (9).

Despite the significant efforts to standardize the operational procedures of spirometry, standardization of reference values is overlooked in the current guidelines for ALS (e.g. the NICE 2016 (10), AAN 2009 (11), and EFNS 2011 (12)). The variability in predictions between spirometric reference standards is, nevertheless, well known (5). It is, therefore, surprising that there is currently no guidance for spirometric reference standards in ALS. The lack of standardization of reference values could have implications for patients and ALS-oriented research settings. This study, therefore, aims to illustrate the effect of spirometric reference values on prognostication, medical decision-making, and trial eligibility, and ultimately provides a basis for standardization of spirometric reference values in ALS.

Materials and methods

Study population

Data for this study originated from the open-access PRO-ACT database (version Dec. 2015). PRO-ACT contains data for 10,731 individuals from 23 ALS clinical trials performed over the past 20 years, is IRB-approved, and uses solely anonymized data; individual trials within PRO-ACT cannot be traced (13). All subjects provided their consent during trial participation. In order to assess the effect of different spirometric reference standards on clinical endpoints, we excluded those patients without follow-up time or missing demographic data regarding sex, age, height, or ethnicity. For each patient, we extracted the FVC values in liters.

Study design, outcomes, and variables

We retrospectively determined the predicted FVC in liters according to five reference standards: Knudson '76 (14), Knudson '83 (15), European Community for Steel and Coal (ECSC) (16), National Health and Nutrition Examination Survey (NHANES) III (17), and GLI-2012 (4). The Knudson '76, '83, and ECSC reference standards are still used in ALS clinics (*personal communication*) and ALS clinical trials (e.g. the Ceftriaxone trial used Knudson '83 (18), whereas the Xaliproden trials were using ECSC (19)), despite their known disadvantages (5). The GLI-

2012, the most recent standard and based on >74,000 global control subjects (4), has been endorsed by both the American Thoracic Society (ATS) and European Respiratory Society (ERS) and will probably supersede previous reference standards (5). The observed FVC was standardized according to the prediction per reference standard and expressed as percentage from normal (%predicted FVC). Survival time was defined as the time from trial inclusion until death from any cause. Patients who remained alive during the trial were censored after their last follow-up visit. The severity of dyspnea symptoms was assessed using item 10 of the revised ALS functional rating scale (ALSFRS-R).

Statistical analysis

We used linear mixed-effects models (LME) to evaluate the longitudinal patterns of decline in FVC, as described elsewhere (20). Using the longitudinal sample size framework provided by Edland, we calculated the number of patients required to detect a 25% reduction in the rate of decline with 80% power after 18 months of follow-up and a quarterly visiting scheme (20). Subsequently, we analyzed the predictive value of each reference standard in predicting survival time using Cox proportional hazard (PH) models. Predictive performance and model fit were evaluated with the Concordance statistic (C-statistic) and Akaike information criterion (AIC), respectively. All Cox PH models were adjusted for the following predictors: age at randomization, treatment arm, Δ FRS (ALSFRS-R at randomization – 48/symptom duration), body mass index (BMI), site of symptom onset, and diagnostic delay (8,21). Missing data in any of the covariates were handled by creating multiple imputed datasets, a procedure described in more detail elsewhere (9). Finally, we calculated for each individual the time to reach a specific %predicted FVC cutoff value (e.g. <80%). Kaplan–Meier curves were used to calculate the median time to reach the defined cutoff value. In addition, a generalized (logistic) LME was used to assess the longitudinal correlation between symptoms of dyspnea and the probability of obtaining a %predicted FVC below the cutoff value. The GLI-2012-predicted FVC values were calculated using the R package *rspi* (version 0.1, Lytras T, 2017) (4). (Generalized) LME models were fitted using the (*glmer* function (*lme4*, version 1.1-18-1) (22).

Results

In total, complete baseline information was available for 4,651 patients; their baseline characteristics are given in Table 1. The average number of longitudinal FVC (L) measurements was 6.9 (total 33,296 measurements); the total follow-up time was 4,701 person-years, during which 1398 deaths

occurred (12-month survival since trial enrollment of 76.1% [95% CI 74.8%–77.5%]). Depending on which reference standard was applied, the mean population %predicted FVC varied between 78.5% (95% CI 78.0–79.1) and 88.5% (95% CI 87.9–89.1); mean difference of 10.0% (95% CI 9.8–10.1, p value < 0.001). Figure 1 provides the individual differences between reference standards

Table 1. Characteristics of patients at baseline.

	PRO-ACT (N = 4651)
Age at enrollment, years	56 (12)
Sex	
Male	62% (2880)
Female	38% (1771)
Race	
Caucasian	97.6% (4540)
African-American	1.8% (82)
Other	0.6% (29)
Site of symptom onset	
Bulbar	21% (982)
Other	79% (3667)
Symptom duration, months	24 (18)
Diagnostic delay, months	12 (9)
Body Mass Index (BMI), kg/m ²	25.8 (4.7)
Forced vital capacity, liters	
Overall	3.30 (1.12)
Males	3.82 (1.01)
Females	2.47 (0.71)
Forced vital capacity, %predicted	
Knudson 1976	86.2% (21.2)
Knudson 1983	85.4% (22.1)
ECSC 1993	88.5% (21.4)
NHANES 1999	78.5% (19.1)
GLI 2012	80.4% (19.4)

Data are given in mean (SD) or % (n).

as measure of disagreement, which reveals a clear difference between males and females: the median disagreement for males was 8.4% (IQR 3.3%) *vs.* 14.3% (IQR 5.1%) for females (p value < 0.001). The mean disagreement between reference standards, irrespective of gender, increased by 1.57% (95% CI 1.50–1.64) per 10% increase in %predicted FVC (p value < 0.001).

The NHANES III and GLI reference standards had a strong correlation (Pearson r 0.997, p value < 0.001), whereas older standards showed a larger deviation from the GLI (Figure 2a–d). As noninvasive ventilation (NIV) referral or trial eligibility is often based on a specific cutoff (9), the cumulative proportion of patients below a certain cutoff is given in Figure 2(e). When trial eligibility is based on, for example, an 80% FVC, 52.8% and 33.6% of the PRO-ACT cohort would be ineligible under NHANES III and ECSC, respectively. Interestingly, the ineligible population (i.e. FVC < 80%) under NHANES III has a 12-month survival of 65.1% (95% CI 63.1%–67.2%); under ECSC, it is 58.2% (95% CI 55.6%–60.9%, p < 0.001 for survival difference). This suggests that, depending on the reference standard, a different population is selected, which could affect the generalizability and between-trial comparability.

The effect of each reference standard on the predictive performance for survival time and variability over time is given in Tables 2 and 3. The unstandardized liters provided the worst fit on the survival data (AIC 20573, c-index 0.760), whereas the GLI provided the best fit (AIC 20374, c-index 0.780, p < 0.001). In the longitudinal data,

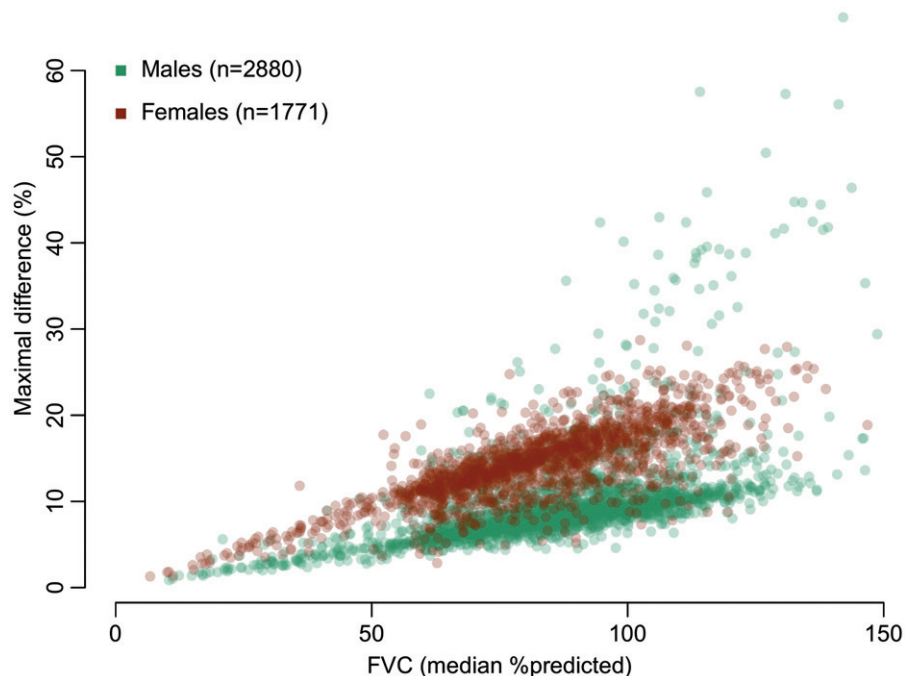


Figure 1. Disagreement in %predicted FVC between reference standards at baseline. For each patient, we calculated the maximal difference in %predicted FVC between two reference standards as measure of disagreement. The median %predicted is the median value of the five reference standards.

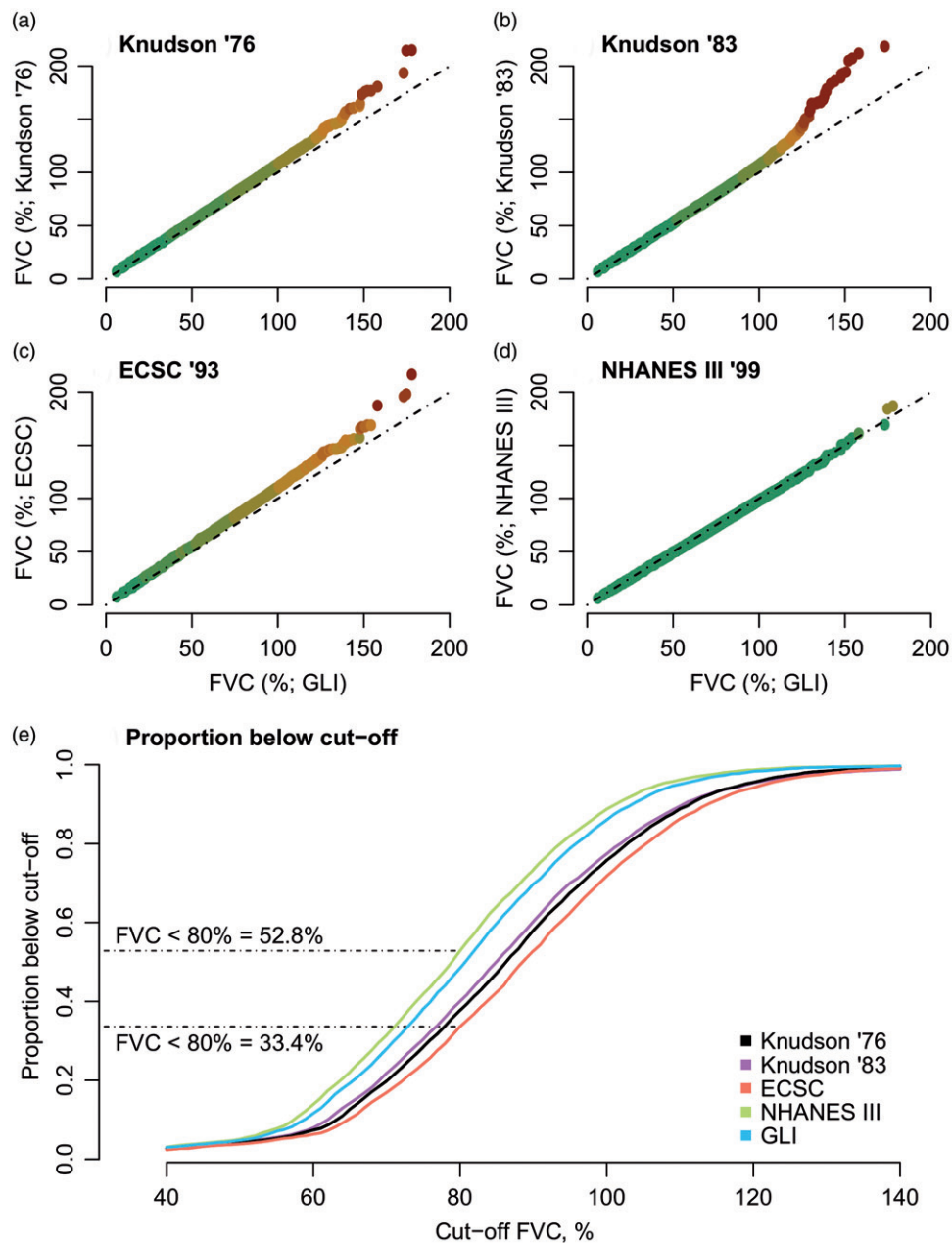


Figure 2. Quantile-Quantile plots of the different reference standards. (a–d) Quantile-Quantile plots of each reference standard with the GLI as reference distribution. Ideally, two reference standards would provide similar estimates, resulting in a straight line (dashed-line). The colors represent the deviation from the ideal line (*green* less than 2.5% deviation, *red* more than 25% deviation). (e) Cumulative proportions of patients below FVC cutoff values in PRO-ACT.

Table 2. Multivariate Cox proportional hazard models per reference standard.

Model	AIC	C-statistic	Hazard ratio
Unadjusted liters	20573	0.760	–
Knudson (1976)	20393	0.778	0.74 (0.72–0.76)
Knudson (1983)	20428	0.774	0.75 (0.72–0.77)
ECSC (1993)	20383	0.780	0.74 (0.72–0.76)
NHANES III (1999)	20380	0.779	0.71 (0.69–0.74)
GLI (2012)	20374	0.780	0.72 (0.69–0.74)

Models were corrected for age, treatment allocation PRO-ACT, ALSFRS-R slope, BMI, site of onset and diagnostic delay. Estimated hazard ratios are per 10% increase in predicted FVC. AIC = Akaike information criterion; C-statistic = concordance statistic.

NHANES III exhibited the lowest between-patient variability over time (2.04%) and would be the most sensitive measure to detect a given treatment effect (Table 3). Interestingly, the mean monthly rate of decline in % FVC could vary as much as 11.4% between reference standards (i.e. –2.81% ECSC *vs.* –2.49% NHANES III).

Finally, we estimated the median time-to-referral for noninvasive ventilation defined as predicted FVC <50% or <80% (Figure 3(a+c)) (23,24). Overall, results for Knudson '76 & '83 were similar to the ECSC (*results not shown*). The median time-to-50% predicted FVC was the shortest when using NHANES III (14.5 months, 95% CI

Table 3. Longitudinal linear mixed models per reference standard.

Model	Baseline	Baseline variance	Slope (<i>Monthly</i>)	Slope variance	Required sample size
Unadjusted liters	3.31 L	1.14 L	-0.10 L	0.09 L	370
Knudson (1976)	86.38%	21.23%	-2.73%	2.25%	356
Knudson (1983)	85.57%	22.12%	-2.69%	2.22%	358
ECSC (1993)	88.70%	21.49%	-2.81%	2.32%	358
NHANES III (1999)	78.71%	19.17%	-2.49%	2.04%	352
GLI (2012)	80.58%	19.49%	-2.55%	2.09%	354

Sample size calculations were based on a 25% reduction in the rate of decline in %FVC after 18 months (quarterly follow-up) with 80% power and a two-sided alpha of 5%.

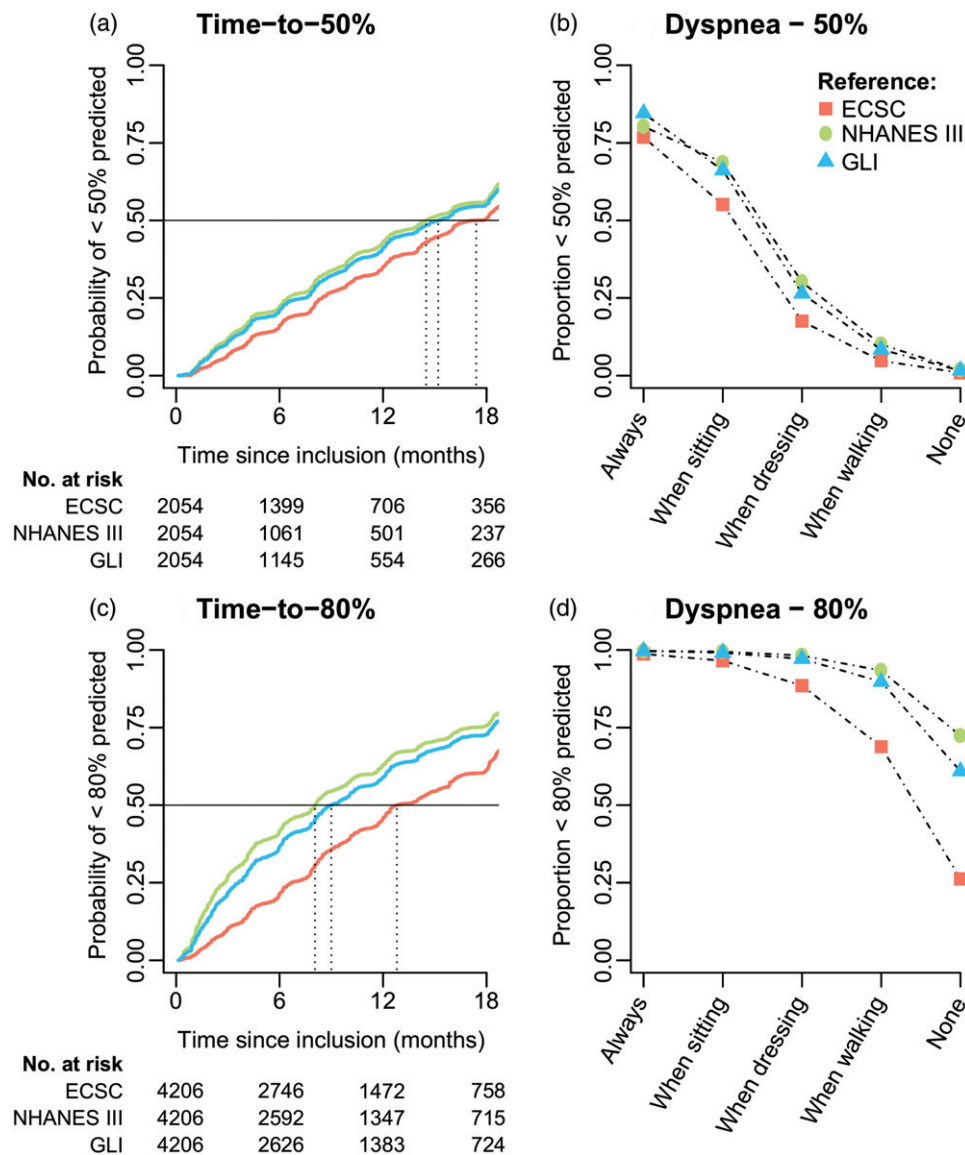


Figure 3. Kaplan-Meier curves of referral criteria and the association with symptoms of dyspnea. (a + c) Time-to-50% or 80% is defined as the time from trial inclusion to the first FVC measurement below 50% or 80% predicted, estimated in those patients whose FVC was higher than 50% or 80% at inclusion ($N=4,206$ and $N=2,054$, respectively). (b + d) Generalized linear mixed-effects model (logistic) with FVC dichotomized (<50% vs. $\geq 50\%$ or <80% vs. $\geq 80\%$) as function of ALSFRS item 10 (Dyspnea) in 4,528 patients with 32,955 matched ALSFRS – FVC data points. Overall, results for Knudson '76 & '83 were similar to the ECSC (*results not shown*).

14.4–16.1), which was 2.9 months earlier than the ECSC (17.4 months, 95% CI 16.1–18.2). This difference was 2.1 months (95% CI: 0.2–3.0) for males and 3.3 months (95% CI: 2.0–4.6) for females. A similar pattern was seen when applying

the 80% cutoff, with a median time of 8.1 (95% CI 7.8–8.4) for NHANES III and 12.8 (95% CI 12.3–14.3) for ECSC. Overall, the difference between GLI and NHANES III was minimal. Figure 3(b + d) reveal the association between

symptoms of dyspnea and proportion of FVC measurements less than 50% or 80%. For the 50% cutoff, symptoms have an approximate linear relationship with the proportion of patients failing the threshold with a minimal difference between standards. Nevertheless, 15.4%–23.2% of the patients with severe symptomatology (always dyspnoeic) still have a %predicted FVC above 50%. Interestingly, 72.4% and 60.9% of the patients, under NHANES and GLI, respectively, fail the 80% cutoff without any symptomatology, whereas this is only 26.2% for de ECSC.

Discussion

In this study, we assessed the impact of spirometric reference standards on clinical and research-oriented settings for patients with ALS. Independent of technique, device, or evaluator, there are important differences between reference standards that could considerably affect medical decision-making, prognostication, patient counseling, and patient selection for clinical trials. The timing of referral to a respiratory care clinic is affected by the applied reference values. Moreover, the comparability of clinical research is obscured when investigators use different standards for study eligibility or when treatment efficacy is based on the respiratory rate of decline. Standardizing the use of spirometric reference standards in both clinical and research-oriented settings is, therefore, of the utmost importance to optimize the utility of spirometry.

Despite the known differences between spirometric reference standards (5), standardization of spirometric reference standards receives surprisingly little attention in ALS. Our results not only scrutinize the differences between reference standards, but also reveal how the variability in predicted FVC could considerably affect real-world decisions. To illustrate: A female patient (62-year-old, 1.68 m, Caucasian, FVC 2.43L) could obtain an 82.5% predicted FVC at the neurologist (using ECSC), whereas this could be 69.0% at the physiotherapist (using NHANES III). This sudden 13.5% drop may erroneously lead to activation of additional care and cause unnecessary psychological distress. In addition, the %predicted FVC is part of a prognostic tool used for patient counseling (8). In this case, a 13.5% difference in FVC (HR 0.99), while keeping other factors constant, falsely predicts a 14.5% increased risk of death during follow-up. Importantly, this difference is solely caused by a systematic difference in reference standards and not due to a true difference in pulmonary function. Between-technician differences and other technical concerns may further inflate this bias. In addition, our results suggest essential differences between reference standards in the timing of referral to specialized respiratory

care (e.g. NIV initiation). This is of particular importance considering that NIV improves overall survival (25) and the suggestive evidence for a beneficial effect of early NIV initiation (23,24). The GLI-2012 and NHANES provide more conservative estimates, which may result in a timelier referral as compared to Knudson '76, '83, or ECSC (Figure 3). Moreover, from the same figure it can be seen that a considerable proportion (~25%) of the patients with severe dyspnea symptoms still has a %predicted FVC >50%. This highlights an important limitation of spirometry in ALS and emphasizes the need to consider both symptomatology and spirometry in clinical settings.

The %predicted FVC is one of the primary eligibility criteria in ALS clinical trials and a common outcome for evaluating drug efficacy (9,26). Patients are selected for clinical trials based on a specific %predicted FVC (e.g. $\geq 80\%$ (27)). As the %predicted is related to the applied reference standard (5), the reference standard exerts a considerable effect on the exclusion rate in clinical trials. In our example, the difference in exclusion rates was considerable (33.4% under NHANES III vs. 52.8% under ECSC) and the selected trial populations differed in their overall survival. This observation has considerable consequences for the comparability of trial populations and the generalizability of results (28). This becomes even more critical when one considers the systematic underreporting of reference standards in ALS clinical trials: Of the 37 randomized placebo-controlled trials using lung function either as inclusion criterion or efficacy endpoint, only three reported the applied reference standard (2 Knudson '83 and 1 ECSC) (9). This indicates that, for 92% of the clinical trials, it is not known which reference standard was used to determine eligibility or efficacy. In the end, this obscures the clinicians' ability to translate trial outcomes to clinical practice.

Our study has limitations that should be considered. First, we used an open-source database of ALS clinical trials. It is well known that trial participants differ from the general population and fast-progressing patients are likely underrepresented (9,29). The estimated time to reach a certain %predicted FVC is, therefore, overestimated. Nevertheless, our imperative was not to obtain real-world estimates of the time to reach a critical limit, but merely to illustrate the large effect of the reference standard on population characteristics and outcomes. Moreover, the PRO-ACT dataset contains data over the full range of %predicted FVC (Figure 1) and the relative differences between spirometric reference standards will not change in population-based datasets. An important consideration is that respiratory insufficiency is generally defined as a carbon dioxide pressure (PCO_2) exceeding 45 mmHg (30), an endpoint that could

not be evaluated in PRO-ACT. It would be worthwhile to assess the associations between reference standards with PCO_2 and to optimize cutoff criteria for referral to specialized respiratory care. Finally, the implementation of the GLI-2012 in clinical practice is not straightforward due to the underlying mathematical model (4). Important work to mediate the implementation has been conducted by the GLI-2012 research group (5). We have extended the implementation tools by providing an additional web-based tool for clinicians to determine the %predicted FVC according to the GLI reference standard (reactive.tricals.org).

In conclusion, our results show how spirometric reference values can considerably affect the utility of spirometry in ALS and, potentially, other neurological diseases. Similar effects may be observed in other outcomes that depend on reference standards such as muscle strength testing or other markers of respiratory function. Our results emphasize the need for the standardization of spirometric reference, which is currently overlooked in guidelines. The GLI-2012 and NHANES III differed minimally and showed the strongest associations with survival and patient-reported symptoms. Standardization of reference standards may optimize clinical decision-making, improve prognostication, enhance between-center comparisons, and unify patient selection for clinical trials.

Declaration of interest

The authors declare that they have no conflict of interest.

Funding

This study was funded by the Dutch ALS foundation (grant: *project TRICALS-Reactive*).

References

1. Lange P, Celli B, Agusti A, Jensen GB, Divo M, Faner R, et al. Lung-function trajectories leading to chronic obstructive pulmonary disease. *N Engl J Med*. 2015;373:111–22.
2. Kelley RC, Ferreira LF. Diaphragm abnormalities in heart failure and aging: mechanisms and integration of cardiovascular and respiratory pathophysiology. *Heart Fail Rev*. 2017;22:191–207.
3. Ambrosino N, Carpena N, Gherardi M. Chronic respiratory care for neuromuscular diseases in adults. *Eur Respir J*. 2009;34:444–51.
4. Quanjer PH, Stanojevic S, Cole TJ, Baur X, Hall GL, Culver BH, et al. Multi-ethnic reference values for spirometry for the 3–95-yr age range: the global lung function 2012 equations. *Eur Respir J*. 2012;40:1324–43.
5. Quanjer PH, Brazzale DJ, Boros PW, Pretto JJ. Implications of adopting the Global Lungs Initiative 2012 all-age reference equations for spirometry. *Eur Respir J*. 2013;42:1046–54.
6. Crapo RO. Pulmonary-function testing. *N Engl J Med*. 1994;331:25–30.
7. Tilanus TBM, Groothuis JT, TenBroek-Pastoor JMC, Feuth TB, Heijdra YF, Slenders JPL, et al. The predictive value of respiratory function tests for non-invasive ventilation in amyotrophic lateral sclerosis. *Respir Res*. 2017;18:144.
8. Westeneng HJ, Debray TPA, Visser AE, van Eijk RPA, Rooney JPK, Calvo A, et al. Prognosis for patients with amyotrophic lateral sclerosis: development and validation of a personalised prediction model. *Lancet Neurol*. 2018;17:423–33.
9. van Eijk RPA, Westeneng HJ, Nikolakopoulos S, Verhagen IE, van Es MA, Eijkemans MJC, et al. Refining eligibility criteria for amyotrophic lateral sclerosis clinical trials. *Neurology*. 2019;92:e451.
10. National Institute for Health and Care Guidance. Motor neuron disease: assessment and management. Available at: <https://www.nice.org.uk/guidance/ng42>; 2016. Accessed 1 May 2019.
11. Miller RG, Jackson CE, Kasarskis EJ, England JD, Forshe D, Johnston W, et al. Practice parameter update: the care of the patient with amyotrophic lateral sclerosis: drug, nutritional, and respiratory therapies (an evidence-based review): report of the Quality Standards Subcommittee of the American Academy of Neurology. *Neurology*. 2009;73:1218–26.
12. EFNS Task Force on Management of Amyotrophic Lateral Sclerosis, Andersen PM, Abrahams S, Borasio GD, de Carvalho M, et al. EFNS guidelines on the clinical management of amyotrophic lateral sclerosis (MALS)—revised report of an EFNS task force. *Eur J Neurol*. 2012;19:360–75.
13. Atassi N, Berry J, Shui A, Zach N, Sherman A, Sinani E, et al. The PRO-ACT database: design, initial analyses, and predictive features. *Neurology*. 2014;83:1719–25.
14. Knudson RJ, Slatin RC, Lebowitz MD, Burrows B. The maximal expiratory flow-volume curve. Normal standards, variability, and effects of age. *Am Rev Respir Dis*. 1976;113:587–600.
15. Knudson RJ, Lebowitz MD, Holberg CJ, Burrows B. Changes in the normal maximal expiratory flow-volume curve with growth and aging. *Am Rev Respir Dis*. 1983;127:725–34.
16. Quanjer PH, Tammeling GJ, Cotes JE, Pedersen OF, Peslin R, Yernault JC. Lung volumes and forced ventilatory flows. Report Working Party Standardization of Lung Function Tests, European Community for Steel and Coal. Official Statement of the European Respiratory Society. *Eur Respir J Suppl*. 1993;16:5–40.
17. Hankinson JL, Odencrantz JR, Fedan KB. Spirometric reference values from a sample of the general U.S. population. *Am J Respir Crit Care Med*. 1999;159:179–87.
18. Cudkowicz ME, Titus S, Kearney M, Yu H, Sherman A, Schoenfeld D, et al. Safety and efficacy of ceftriaxone for amyotrophic lateral sclerosis: a multi-stage, randomised, double-blind, placebo-controlled trial. *Lancet Neurol*. 2014;13:1083–91.
19. Meininger V, Bensimon G, Bradley WR, Brooks B, Douillet P, Eisen AA, et al. Efficacy and safety of xaliproden in amyotrophic lateral sclerosis: results of two phase III trials. *Amyotroph Lateral Scler Other Motor Neuron Disord*. 2004;5:107–17.
20. van Eijk RPA, Eijkemans MJC, Ferguson TA, Nikolakopoulos S, Veldink JH, van den Berg LH. Monitoring disease progression with plasma creatinine in amyotrophic lateral sclerosis clinical trials. *J Neurol Neurosurg Psychiatry*. 2018;89:156–61.
21. Chio A, Logroscino G, Hardiman O, Swigler R, Mitchell D, Beghi E, et al. Prognostic factors in ALS: a critical review. *Amyotroph Lateral Scler*. 2009;10:310–23.

22. Bates D, Mächler M, Bolker B, Walker S. Fitting linear mixed-effects models using lme4. *J Stat Software*. 2015; 2015;67–48.
23. Jacobs TL, Brown DL, Baek J, Migda EM, Funckes T, Gruis KL. Trial of early noninvasive ventilation for ALS: a pilot placebo-controlled study. *Neurology*. 2016;87: 1878–83.
24. Vitacca M, Montini A, Lunetta C, Banfi P, Bertella E, De Mattia E, et al. Impact of an early respiratory care programme with non-invasive ventilation adaptation in patients with amyotrophic lateral sclerosis. *Eur J Neurol*. 2018;25:556–e33.
25. 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.
26. Andrews JA, Meng L, Kulke SF, Rudnicki SA, Wolff AA, Bozik ME, et al. Association between decline in slow vital capacity and respiratory insufficiency, use of assisted ventilation, tracheostomy, or death in patients with amyotrophic lateral sclerosis. *JAMA Neurol*. 2018;75:58–64.
27. Edaravone Writing Group. Safety and efficacy of edaravone in well defined patients with amyotrophic lateral sclerosis: a randomised, double-blind, placebo-controlled trial. *Lancet Neurol*. 2017;16:505–12.
28. Rothwell PM. External validity of randomised controlled trials: “to whom do the results of this trial apply?”. *Lancet*. 2005;365:82–93.
29. Chio A, Canosa A, Gallo S, Cammarosano S, Moglia C, Fuda G, et al. ALS clinical trials: do enrolled patients accurately represent the ALS population? *Neurology*. 2011;77:1432–7.
30. Weinberger SE, Schwartzstein RM, Weiss JW. Hypercapnia. *N Engl J Med*. 1989;321:1223–31.