Neurological Sciences (2018) 39:847–850
https://doi.org/10.1007/s10072-018-3267-8

ORIGINAL ARTICLE

An electroglottographical analysis-based discriminant function model differentiating multiple sclerosis patients from healthy controls

George D. Vavougios [1,2] · Triantafyllos Doskas [1,2] · Kostas Konstantopoulos [3,4]

Received: 28 September 2017 / Accepted: 31 January 2018 / Published online: 15 February 2018

Abstract

Dysarthrophonia is a predominant symptom in many neurological diseases, affecting the quality of life of the patients. In this study, we produced a discriminant function equation that can differentiate MS patients from healthy controls, using electroglottographic variables not analyzed in a previous study. We applied stepwise linear discriminant function analysis in order to produce a function and score derived from electroglottographic variables extracted from a previous study. The derived discriminant function's statistical significance was determined via Wilk's λ test (and the associated p value). Finally, a 2×2 confusion matrix was used to determine the function's predictive accuracy, whereas the cross-validated predictive accuracy is estimated via the "leave-one-out" classification process. Discriminant function analysis (DFA) was used to create a linear function of continuous predictors. DFA produced the following model (Wilk's $\lambda = 0.043$, $\chi 2 = 388.588$, $p < 0.0001$, Tables 3 and 4): D (MS vs controls) = 0.728*DQx1 mean monologue + 0.325*CQx monologue + 0.298*DFx1 90% range monologue + 0.443*DQx1 90% range reading − 1.490*DQx1 90% range monologue. The derived discriminant score (S1) was used subsequently in order to form the coordinates of a ROC curve. Thus, a cutoff score of − 0.788 for S1 corresponded to a perfect classification (100% sensitivity and 100% specificity, $p = 1.67e^{-22}$). Consistent with previous findings, electroglottographic evaluation represents an easy to implement and potentially important assessment in MS patients, achieving adequate classification accuracy. Further evaluation is needed to determine its use as a biomarker.

Keywords Electroglottography · Linear discriminant function analysis · Multiple sclerosis

Introduction

Dysarthrophonia is a symptom in many neurological diseases, affecting the quality of life of patients. Dysarthrophonia research is increasing in stroke [1–3], multiple sclerosis [4], Parkinson's disease [5], amyotrophic lateral sclerosis [6], and myasthenia gravis [7].

Electroglottography (EGG) is an indirect, non-invasive quantitative technique used in ENT to measure the duration of the relative vocal fold contact patterns within the glottal cycle and its waveforms are produced when the vocal fold contact increase as electrical impedance decreases [8]. So, the slope of the EGG wave is higher (hills) when the vocal folds are in contact (closed phases) and lower (troughs) when the vocal folds are apart (open phases).

In a previous study [4], we used EGG data derived from a cohort of MS patients who were pair-matched with healthy controls (age and sex) to determine univariate predictors of phonatory impairment. In this study, we aim to produce a discriminant function equation that can differentiate MS patients from healthy controls, using EGG variables not analyzed in the previous study [6].

✉ Kostas Konstantopoulos
c.constantopoulos@euc.ac.cy

[1] University of Thessaly, Biopolis, 41110 Larissa, Greece

[2] Department of Neurology, Athens Naval Hospital, Deinokratous 70, Athens, Greece

[3] Health Sciences Department, Speech Therapy, European University Cyprus, 6 Diogenous Street, Engomi, P.O. Box: 22006, 1516 Nicosia, Cyprus

[4] Cyprus Institute for Neurology and Genetics, Nicosia, Cyprus

Methods

The Athens Naval Hospital's ethics committee approved the study protocol in the study from which the current data and variables were taken. In that study [4], all participants were informed about the aim and methods and signed an informed consent form.

The produced DFA model included a number of variables (predictors) in 128 participants (64 MS vs. 64 pair-matched healthy controls) [4].

So, variables such as the duration of each cycle (period) and the fundamental frequency of the vocal folds during vibration (fundamental frequency = 1/time for period) were examined. Based on the duration of each cycle, another variable (CQx, closed quotient expressed as percentage) was defined as the 70% of the peak wave width divided by the time for total period and multiplied by 100. Figure 1 shows a typical electroglottographic cycle (A) and the selection of a word with its concomitant cycles (B) in one MS patient exhibiting ataxic dysarthria.

The EGG data in every variable were produced as histograms with means, standard deviations, ranges, etc. or as scatterplots which showed the variations of closed quotients. Specifically, five voice variables (either frequencies or closed quotients) were employed in the present study.

- DFx1 90% range in reading = range of the fundamental frequency values of the vibrating vocal folds above and below the average between which 90% of the observed frequencies lied
- DQx1 mean in monologue = mean of closed quotients during monologue
- DQx1 90% range in reading = range of closed quotients during reading
- DQx1 90% range in monologue = range of closed quotients during monologue
- CQx1 mean in monologue = mean degree of irregularity as expressed on a scatterplot of closed quotients.

Statistical methods

Univariate analyses

The One Sample Kolmogorov-Smirnov was used to assess data normality. Results are presented as mean ± standard deviation (SD) for continuous variables if normally distributed and as median (interquartile range) if not. Qualitative variables are presented as absolute number (percentage). Correlation was determined via the Pearson's *R* coefficient where applicable.

Multivariate analyses

Discriminant function analysis (DFA) was used to create a linear function of continuous predictors. In linear DFA with a binary (categorical) dependent variable, a single discriminant function is derived. In terms of application, predictive accuracy, it is similar with binary logistic regression, if not more accurate [9]. Finally, since it produces a score that can be used in subsequent classification attempts (and can be hence cross-validated), DFA represents an attractive approach in developing a multivariate predictive model. Stepwise DFA, as used in our study, determines the optimal number of predictors, selecting the most correlated among the input to be included in the final model. In this study, the DFA model creates a score based on the formula below:

$$D_i = a_1x_1 + a_2x_2 + \ldots. + a_nx_n + c_i$$

Where:

- D_i: the discriminant function score for the ith function.

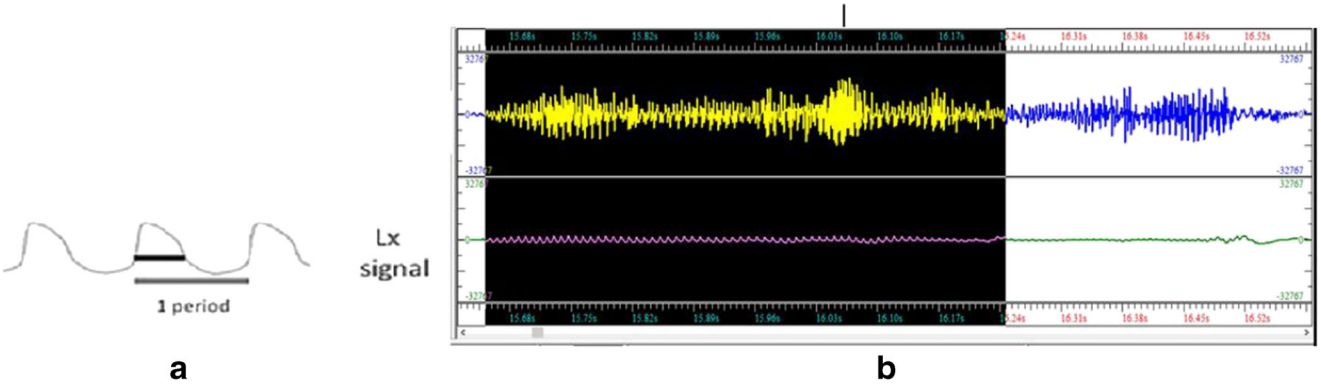

Fig. 1 Electroglottographic demonstration of a normal cycle (closed phase with a bold line) (**a**) in a MS patient exhibiting ataxic dysarthria (**b**). The electroglottographic signal is depicted in pink and green (the speech signal is depicted in yellow and blue color). The patient produced 2 words in a continuous speech of a standard reading passage (black area)

- $a_1 \ldots_n$: discriminant function coefficients for the y predictors, with $y \in [1 \ldots n]$.
- $x_1 \ldots_n$: discriminant function variables for the y predictors, with $y \in [1 \ldots n]$.
- c_i: the constant for the ith

Subsequently, the derived discriminant function's statistical significance was determined via Wilk's λ test (and the associated p value). Finally, a 2×2 confusion matrix was further used to determine the function's predictive accuracy, whereas the cross-validated predictive accuracy was estimated via the "leave-one-out" classification process [10].

Receiver-operating characteristic analysis

Using the discriminant function scores (determined in the previous step) as coordinates, receiver-operating characteristic (ROC) curves were constructed, and areas under the curve were calculated in order to determine optimal cutoff (MS vs. controls). For the purposes of determining diagnostic accuracy, sensitivities and specificities were also determined in this step of the analysis. For all univariate and multivariate analyses, a p value of < 0.05 was considered statistically significant. All analyses were performed using the SPSS 23.0 software (IBM, New York, USA).

Results

Demographic and EGG analyses have been previously presented in detail [4]; Table 1 represents the characteristics of the sample including EDSS and disease duration.

Table 1 Decriptive characteristics (MS duration and EDSS global score) and electroglottographic variables employed in the present study

Variable		Value
MS duration (years)		2.50 ± 4.04
EDSS		2.14 ± 2.26
DQx1 mean (monologue)	*MS*	45.35 ± 3.79
	Control	44.65 ± 4.37
DQx1 90% range (reading)	*MS*	15.60 ± 5.32
	Control	16.79 ± 4.19
DQx1 90% range (monologue)	*MS*	16.63 ± 4.52
	Control	18.65 ± 4.37
CQx mean (monologue)	*MS*	24.81 ± 10.70
	Control	22.46 ± 6.78
DFx1 90% range (monologue)	*MS*	0.83 ± 0.21
	Control	0.92 ± 0.23

MS duration: time in months since a diagnosis of multiple sclerosis. Continuous variables are presented as mean ± SD

The stepwise DFA model

The function produced is as follows:

DFA produced the following model (Wilk's $\lambda = 0.043$, $\chi^2 = 388.588$, $p < 0.0001$):

$$D_{(\text{MS vs controls})} = 0.728^{*}\text{DQx1}_{\text{mean monologue}} + 0.325^{*}\text{CQx}_{\text{monologue}} + 0.298^{*}\text{DFx1}_{90\%\text{range monologue}} + 0.443^{*}\text{DQx1}_{90\%\text{range reading}} - 1.490^{*}\text{DQx1 90\% range monologue}$$

The derived discriminant score (S_1) was used subsequently in order to form the coordinates of a ROC curve. Thus, a cutoff score of -0.788 for S_1 corresponded to a perfect classification (100% sensitivity and 100% specificity, $p = 1.67e^{-22}$). The DFA model was found to correlate with the duration of the disease ($r = 0.583$, $p < 0.00001$ but not with the EDSS global score ($r = -0.051$, $p = 0.653$).

Discussion

Consistent with previous findings [4], EGG represents an easy to implement and potentially important assessment in MS patients. In the present study, it achieved 100% classification accuracy. The correlation of the DFA score with disease duration shows that dysarthrophonia increases as the disease duration increases, a finding consistent with the everyday clinical impression. In contrast, the lack of correlation of the DFA score with EDSS global score may indicate that this score may not be sensitive to dysarthrophonia.

An important limitation of the present study involved that the examiner was not blind to the status of the patients during the voice/speech assessment. Further research is needed to test this method in MS patients employing the double-blind design, during or immediately after the diagnosis in order to reliably quantify mild degrees of dysarthrophonia not necessarily perceivable to the human ear.

Conclusions

Consistent with previous findings [4], EGG evaluation represents an easy to implement and potentially important assessment in MS patients. The correlation of the DFA score with disease duration may indicate that this score may be a sensitive factor to correlate with the underlying neurodegenerative processes as the disease progresses. Further evaluation is needed to determine the use of EGG as a biomarker in the early diagnosis of the disease. Along the same lines, new research is needed to use EGG for the differential diagnosis in patients with dysarthrophonia in other neurological diseases (for example early amyotrophic lateral sclerosis vs. myasthenia gravis).

References

1. Beliavsky A, Perry JJ, Dowlatshahi D, Wasserman J, Sivilotti MLA, Sutherland J, Worster A, Émond M, Stotts G, Jin AY, Oczkowski WJ, Sahlas DJ, Murray HE, MacKey A, Verreault S, Wells GA, Stiell IG, Sharma M (2014) Acute isolated dysarthria is associated with a high risk of stroke. Cerebrovasc Dis Extra 4(2): 182–185. https://doi.org/10.1159/000365169
2. Urban PP, Marx J, Hunsche S, Gawehn J, Vucurevic G, Wicht S, Massinger C, Stoeter P, Hopf HC (2003) Cerebellar speech representation: lesion topography in dysarthria as derived from cerebellar ischemia and functional magnetic resonance imaging. Arch Neurol 60(7):965–972. https://doi.org/10.1001/archneur.60.7.965
3. Urban PP, Rolke R, Wicht S, Keilmann A, Stoeter P, Hopf HC, Dieterich M (2006) Left-hemispheric dominance for articulation: a prospective study on acute ischaemic dysarthria at different localizations. Brain 129(3):767–777. https://doi.org/10.1093/brain/awh708
4. Konstantopoulos K, Vikelis M, Seikel JA, Mitsikostas DD (2010) The existence of phonatory instability in multiple sclerosis: an acoustic and electroglottographic study. Neurol Sci 31(3):259–268. https://doi.org/10.1007/s10072-009-0170-3
5. Midi I, Dogan M, Koseoglu M, Can G, Sehitoglu MA, Gunal DI (2008) Voice abnormalities and their relation with motor dysfunction in Parkinson's disease. Acta Neurol Scand 117(1): 26–34. https://doi.org/10.1111/j.1600-0404.2007.00965.x
6. Robert D, Pouget J, Giovanni A, Azulay JP, Triglia JM (1999) Quantitative voice analysis in the assessment of bulbar involvement in amyotrophic lateral sclerosis. Acta Otolaryngol 119(6):724–731
7. Konstantopoulos K, Christou YP, Vogazianos P, Zamba-Papanicolaou E, Kleopa KA (2017) A quantitative method for the assessment of dysarthrophonia in myasthenia gravis. J Neurol Sci 377:42–46. https://doi.org/10.1016/j.jns.2017.03.045
8. Fourcin AJ. Laryngographic assessment of phonatory function (1981). The American Speech-Language-Hearing Association, ASHA Reports: 116–127
9. Antonogeorgos G, Panagiotakos DB, Priftis KN, Tzonou A (2009). Logistic regression and linear discriminant analyses in evaluating factors associated with asthma prevalence among 10- to 12-years-old children: divergence and similarity of the two statistical methods. Int J Pediatr: 952042. https://doi.org/10.1155/2009/952042
10. Natsios G, Pastaka C, Vavougios G, Zarogiannis SG, Tsolaki V, Dimoulis A, Seitanidis G, Gourgoulianis KI (2016) Age, body mass index, and daytime and nocturnal hypoxia as predictors of hypertension in patients with obstructive sleep apnea. J Clin Hypertens (Greenwich) 18(2):146–152. https://doi.org/10.1111/jch.12645