Prospective observational study to assess the validity of the functional disability score in patients with blepharospasm, hemifacial spasm and synkinesis treated with botulinum toxin injection

Linda O. Okafor^, Aaron Jamison, Maribel Favor, Raman Malhotra

Corneoplastic Unit, Queen Victoria Hospital, East Grinstead, UK

Contributions: (I) Conception and design: R Malhotra, LO Okafor; (II) Administrative support: LO Okafor; (III) Provision of study materials or patients: LO Okafor, M Favor; (IV) Collection and assembly of data: LO Okafor, M Favor; (V) Data analysis and interpretation: LO Okafor; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: Miss Linda O. Okafor, F.R.C.Ophth. Corneoplastic Unit, Queen Victoria Hospital, Holtye Road, East Grinstead, RH19 3DZ, UK. Email: linda.okafor@nhs.net.

Background: Benign essential blepharospasm (BEB), aberrant facial nerve degeneration and hemifacial spasm (HFS) are all examples dystonia which, though not life-threatening, can have a significant impact on patient quality of life. The need for reliable self-rating surveys to monitor functional disability is fundamental. The Blepharospasm Disability Index (BSDI) is already a widely utilised and validated self-rating score for blepharospasm whilst the functional disability score (FDS) requires further validation. The principle aim of this study is to repeat validation of the FDS against the BSDI, which has been validated by several groups since its original description but only in patients with BEB.

Methods: A randomised blinded prospective cohort study was conducted at a single unit on 38 patients with BEB, aberrant facial nerve degeneration and HFS. Patients were blinded to complete the FDS followed by the BSDI or the BSDI followed by the FDS with a 30-minute interval.

Results: Both the FDS and BSDI were found to be reliable with high internal consistency and test-retest reliability. Both scales were also found to be moderately correlated with the Jankovic disease severity score.

Conclusions: This study is the first to use the FDS as a rating scale in patients with HFS and aberrant facial nerve degeneration. It is also the first study to formally validate the FDS as an acceptable rating scale for patients with dystonia and in particular it provides validation for its use in patients with HFS and aberrant facial nerve degeneration.

Keywords: Blepharospasm; botulinum toxin (BoNT); dystonia

Introduction

Benign essential blepharospasm (BEB) is a disorder of progressive involuntary contractions of the eyelid protractors (orbicularis oculi, corrugator, and procerus muscles) resulting in eyelid closure (1). Aberrant facial nerve regeneration (AFR), a recognized complication of peripheral facial nerve palsy, results in a synkinesis characterized by synchronous but unintended movements of certain areas of mimic muscles becoming mostly evident during spontaneous movements of the face based

^ ORCID: 0000-0001-7392-7569.

on emotional expressions (2). Hemifacial spasm (HFS) is characterized by involuntary, irregular, tonic, and clonic synchronous contraction of the muscles innervated by the ipsilateral facial nerve (1).

Several different measurement tools and grading instruments have been utilised to evaluate the effects of botulinum toxin (BoNT) on various aspects of blepharospasm, including force of eyelid closure, severity of muscle spasms and patient functional status (3-8). Today the most popular rating scales are the Jankovic Rating Scale (JRS) and the Blepharospasm Disability Index (BSDI). The JRS is the most widely used current examination-based clinical scale with two subscales that comprise severity and frequency (5-point scales ranging from 0 to 4 where 0 indicates no symptoms and 4 indicates the most severe or frequent symptoms) (3). Advantages of the JRS include its broad applicability and simplicity for both patients and physicians whilst its main disadvantage is lack of sensitivity to small changes in severity of symptoms.

As opposed to the JRS, instruments that assess everyday activities or patient functional status are rated by the patients themselves. These scales recognise the significance of improvement in everyday activities as an outcome of therapy. The blepharospasm disability scale (BDS) emerged as a useful functional ability rating scale in the 1980s and 1990s (9,10). Despite documented reliability and validity of this scale, certain drawbacks have been identified. The BSDI was created to address these drawbacks and has been used in several BoNT studies (11). The BSDI comprises six daily activities each rated on a scale of 0 (= no impairment) to 4 (= activity not possible due to disease). Critically, there is an opportunity to grade an activity as "not applicable" for any individual item. The BSDI was developed to complement rather than an alternative to the JRS (12).

An alternative grading scale, the functional disability score (FDS), includes six clearly defined criteria with a 5-point scale of growing value. The FDS was first described and validated in a group of patients operated on for BEB resistant to treatment with BoNT (13). Though the assessment criteria are similar to that of the BSDI, more specific questions, relevant to daily activities are asked. However, the FDS has not received additional validation in trials outside of its original developers.

The principle aim of this study is to repeat validation of the FDS against the BSDI, which has been validated by several groups since its original description but only in patients with BEB. Furthermore, this study uniquely provides validation for the use of the FDS and the BSDI in patients with HFS and AFR. In particular we aim to compare the rating scales with respect to their metric properties with a focus on reliability and validity testing, in patients with BEB, HFS and AFR. We present this article in accordance with the STROBE reporting checklist (available at https://aes.amegroups.com/article/view/10.21037/aes-22-42/rc).

Methods

Participants

A total of 38 subjects with either BEB, HFS, AFR or a combination of these were recruited for the study in an outpatient setting based at the Corneoplastic Unit, Queen Victoria Hospital, East Grinstead. All patients were already receiving BoNT therapy on a rolling basis. The inclusion criteria were (I) patients of the Queen Victoria Hospital NHS Foundation Trust meeting the diagnostic criteria for BEB, HFS or AFR; (II) male or female patients aged over 18 years; (III) patients who were fluent in written and spoken English language; (IV) patients willing and able to give informed consent for participation in the study. Specific exclusion criteria were patients who had previously undergone periocular surgery.

Procedure

All participants completed the JRS, FDS and BSDI. Once obtaining informed consent, patients were randomly

---

Highlight box

Key findings
- This study demonstrates that the functional disability score is a valid and reliable tool to measure the functional disability of patients with dystonia—in particular BEB, HFS and AFR.

What is known and what is new?
- The blepharospasm disability index is already a widely utilised and validated self-rating score for blepharospasm.
- This study formally validates the functional disability score as an acceptable rating scale for patients with dystonia.

What is the implication, and what should change now?
- We recommend that the FDS is utilised more readily in clinical practice as an alternative option to the BSDI which will further validate its use.

assigned to complete the FDS followed by the BSDI or vice versa with 30 minutes in between the completion of each self-rating scale. Patients were blinded to the name of the self-rating scale. Participants were randomised to either category 1 (completing FDS followed by BSDI) or category 2 (completing BSDI followed by FDS). Each scale appeared on separate pages. A trained study team member then completed the JRS to assess severity of symptoms prior to treatment with BoNT. The total time for the assessment was approximately 30 minutes. This was then followed by routine administration of BoNT as per protocol.

JRS

The JRS, to date, is the only severity scale specifically developed for blepharospasm and widely utilised. The scale consists of two subscales that quantify intensity and frequency of eyelid muscle spasms, both based on a five-point grading scale (3). The total score ranges from 0 to 8 points (sum score) and includes two categories: severity (from 0= none to 4= severe) and frequency (from 0= none to 4= functionally "blind" due to persistent eye closure more than 50% of waking time). A single experienced rater was used to reduce interrater variability. The JRS scale has been widely criticised for its the lack of a clear definition of spasms regarding the degree of eyelid closure, the combination of examiner-based and patient-based information, and the lack of detail with regard to clinical features.

BSDI

The BSDI was developed to improve the BDS with respect to ease of use (12). It is a patient-rated disability scale that is disease-specific and quantifies inability to perform specific everyday activities as a result of blepharospasm. It comprises six domains rating specified activities (vehicle driving, reading, watching television, shopping, walking, and doing everyday activities), scored on a 5-point Likert scale relating to the severity of impairment (0= no impairment; 4= no longer possible due to illness), as well as a "not applicable" option. The range of scores is 0 to 24, with higher scores indicating a greater disability. The mean item score of the BSDI can also be calculated by dividing the total BSDI score by the number of items answered. It is available only in English, although the scale has been used extensively in Europe and Israel.

FDS

The FDS, like the BDS, stems from the ability to perform distinct activities of daily living, thus curtailing interview subjectivity. Six criteria confirm an accurate result though criteria like driving or work are not included in the final calculation as they are frequently not applicable to each patient due to the average age of patients with dystonia (13). The scale makes it possible to enumerate precisely the patient's socio-professional disability at a given time. Evolution of disability can therefore be monitored following treatment—the efficacy of which can in turn be evaluated. The FDS also avoids non-discriminatory questions like sunglasses wear which can also affect the accuracy of the score.

Additional points

We recognise that HFS and AFR are mostly due to pathology of the facial nerve whereas BEB is a focal dystonia however the disability scale sores have no relevance to aetiology and simply score severity of symptoms in relation to daily life.

Statistical analysis

Data was collected prospectively from clinical notes as well as the assessment questionnaires. Microsoft Excel software was utilised for data input and, whilst data analysis was completed using Statistical Product and Service Solutions version 28 (SPSS 28.0.1.0). Differences between individual groups with reference to continuous variables were assessed using the paired t-test.

Reliability testing

We utilise test-retest reliability and internal consistency as measure of reliability. Test-retest reliability involves administering the rating scale to patients and then repeating the questionnaire with the same set of patients at a future time point (14). Two separate time points were then used to compare the responses.

It was imperative to re-test those patients who usually require reinjection with Botox-A at roughly the 3-month time point as this is assumed to be when symptoms return, and reinjection is required. It is the point at which reinjection is required that we use to assess disability with the patient questionnaires. We therefore opted to only
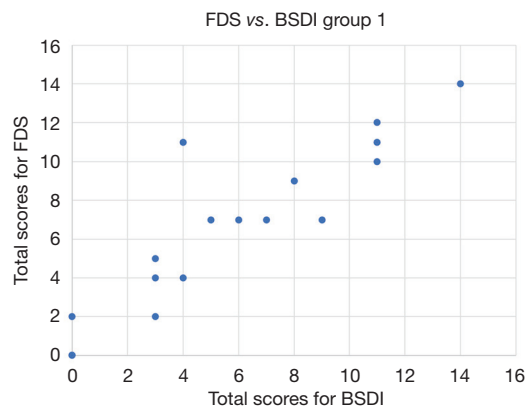
Figure 1 Regression line graph of FDS total scores and BSDI total scores. FDS, functional disability score; BSDI, blepharospasm disability index.

retest those patients who usually require their Botox-A at the 3-month interval and who did come back within that timeframe. Some patients who did not require their Botox-A at 3 months were not included in the test-retest calculation.

As the rating scales consist of tailored questions that are merged to create a score on a scale, we compare the values at the two different timepoints as a correlation. In this case Spearman rank coefficient was calculated using SPSS.

Internal consistency also known as split-half reliability refers to the general agreement between multiple items (often Likert scale items) that make up a composite score of a rating scale of a given construct. Internal consistency index of reliability is appealing because it is estimated after only one test administration and thus avoiding the problems associated with testing over numerous time periods. Cronbach's α coefficient is the most commonly used measure of internal consistency hence its utilisation alpha is the most commonly used measure of internal consistency. Cronbach's alpha is the equivalent of the average reliability across all possible groupings of split-halves. In particular, analysis software will also habitually calculate for each questionnaire item the value of Cronbach's alpha if that questionnaire item was deleted. These values can be scrutinised to evaluate whether the reliability of the scale can be enhanced by omitting any of the questionnaire items. Cronbach's alpha was calculated on SPSS.15.

Validity testing

Convergent validity is a subtype of construct validity and refers to how closely two tests or scales that have the same or similar constructs are related. Construct validity refers to a test designed to quantify a particular construct (i.e., intelligence) that is actually measuring that construct. Spearman rank correlation coefficient was used to assess convergent validity by correlating the total scores of the JRS and the BSDI or FDS. Moderately high correlations ($0.5 < r < 0.7$) would be suggestive of reasonable results, i.e., corresponding to the results of the other scales, but not measuring redundantly the same aspects of therapeutic efficacy. The latter would be the case with very high correlation coefficients (close to $r = 1.0$) (14).

Ethical compliance

This study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). The South West-Central Bristol Institutional Review Board approved this study with a reference number of IRAS-297057. All patients provided written informed consent, and understood and complied with the study requirements.

Results

Demographics of the study population.

The mean age of the study population comprising 38 patients was 74.36±11.73 (range, 39–91 years); these included 30 females and 8 males. Our cohort consisted of a majority of patients with BEB, HFS, HFS/AFR and AFR at 23, 8, 6 and 1 respectively. The mean Jankovic score was 3.24. The mean number of days elapsed from the last BoNT treatment and the date of study investigation was 144 days.

Consistency between FDS and BSDI

Addressing the consistency between total FDS scores and total BSDI scores, the intraclass correlation coefficient was 0.953 (>0.7, very high) [95% confidence interval (CI): 0.909–0.976, P<0.001]. Utilising the sum of the BSDI as the independent variable, linear regression analysis of the sum of the FDS and the sum of the BSDI was performed (*Figure 1*). Utilising the sum of the BSDI as the independent variable X and the sum of the FDS as the dependent variable Y the regression equation was Y= –0.232+0.990X. A *t*-test was conducted on regression coefficient 0.990, *t*=18.042 (P<0.01), and regression relation was observed between the total BSDI score and total FDS score. The coefficient of

determination R^2=0.900 and the regression model showed a good fit. Paired *t*-tests were conducted on patient responses to the FDS and BSDI for each like for like category (reading, watching television, driving, walking and everyday/household tasks) and found no significant difference between patient responses to the BSDI and FDS in each category.

Reliability analysis and validity testing

The internal consistency of the FDS for the category 1 group was 0.907 (95% CI: 0.777–0.973) and category 2 was 0.829 (95% CI: 0.709–0.966) as established by Cronbach's α coefficient. The BSDI has an internal consistency of 0.876 (95% CI: 0.744–0.952) for the category 1 patients and 0.897 (95% CI: 0.776–0.964) for category 2 patients, as established by Cronbach's α coefficient. Using the recommendations by Nunnally and Bernstein [2006] (15) for α ≥0.70, these internal consistency values were considered adequate.

To evaluate the test-retest reliability, seven of the 38 patients were asked to complete the FDS followed by BSDI for a second time, approximately 8–12 weeks after initially completing the said questionnaire. The intra-class correlation coefficient for the test-retest reliability of FDS was 0.856 (P<0.001). According to the recommendations of McGraw and Wong [1996] (16) for interpreting the effect size, this value implies a moderate to high correlation.

The concurrent validity of FDS and BSDI was assessed by correlating these questionnaires with JRS using Spearman correlation analysis. As expected, the levels of disease severity based on BSDI (r=0.588) and FDS (r=0.521) were moderately associated with the JRS scores (P<0.001 for both).

Discussion

This study is the first to evaluate the use of the FDS against the BSDI score as a rating scale for disability caused by dystonia. The potential effect on quality of life is high for patients with dystonia. Patient self-rating surveys during the course of treatment can be a useful indicator of efficacy—hence the need for reliability. Both the BSDI and the FDS are 5-point Likert scales which assess disability experienced by patients with dystonia, though the BSDI is more widely used in other treatment centres. A key difference between the two scales is that whilst the BSDI may be considered specific, the FDS is more relatable in assessing tasks such as reading where there is a focus on the duration of certain tasks, i.e., ability to watch a feature film at 2 hours vs ability to watch a half hour sitcom. A more obvious difference is that the BSDI asks about everyday activities and shopping whilst the FDS asks about house activities and work.

Consistent with the previous study of Jankovic *et al.* 2009, this study demonstrated that BSDI and the FDS, are reliable and valid screening tools for functional disability among patients with dystonia—specifically BEB, HFS and AFR (12). In this study, we showed that the BSDI and the FDS have high reliability, as evidenced by Cronbach's α-based internal consistency. Internal consistency was also acceptable despite the relatively minor number of items in the FDS, particularly if one considers that Cronbach's α is also reliant on the number of items. The test-retest reliability of both surveys was acceptable. The moderately high correlation of FDS and BSDI with JRS signified their concurrent validity for the evaluation of the severity of dystonia, thus endorsing the outcomes of previous studies (15).

This study has several strengths. By conducting the survey in a tertiary unit with long-term dystonia patients well used to completing questionnaires on symptoms we have reduced any risk of misinterpretation of the study questionnaires. Patients were asked to complete each blinded survey with an interval of 30 minutes in order to reduce the risk of bias. Having similar surveys repeated with shorter intervals (<30 minutes) could lead to enhanced variation in patient feedback and a falsely low degree of reliability due to a priming effect on subject recall of symptoms (16). Specifically, patients who are asked about symptoms repeatedly might tend to report these symptoms with greater (or, perhaps less likely, lesser) frequency/severity when asked about them a second time, particularly in rapid succession. We compared patient responses to the FDS and BSDI in each like-for-like category (reading, watching television, driving, walking and everyday/household tasks) and found a tendency for scores to go up on for the second survey administered however we could not elicit any statistical significance.

In this study we opted for surveys self-administered by the patient. In general, public health methodology suggests that to minimize bias it would be preferable to have survey instruments self-administered by the patient (if possible) (e.g., using a tablet) *vs.* survey administration by interview (17). It is possible, for example, that patients may report greater improvement in symptom scores when asked by the surgeon directly, as compared to providing this data via an instrument administered via tablet. In addition, self-administration via tablet may be more efficient and cost-effective with a lower

administrative/time burden on clinic personnel. A future consideration would be the use of tablet administration of surveys in additional studies (18,19).

There are some limitations of this study to mention. Firstly, our sample was selected from and executed in an isolated tertiary unit. Thus, the findings for FDS and BSDI may not be entirely demonstrative of other comparable settings. Additional studies could be carried out in multiple units with a varied patient population. A second limitation is the sample size whereby preliminary power calculations suggested a larger sample size would be beneficial. We recommend for future studies larger cohorts with a multi-centred approach may provide more definitive evidence. It is important to point out that due to the moderately low rate of the disease the target population for the study were patients already receiving treatment thus this may have led to some selection bias. Conducting similar studies on new dystonia patients who are inexperienced in answering surveys may also yield relevant data. It is also important to note patients had varied number of days between their last dose of BoNT and the study date which would have impacted on their degree of symptoms and therefore survey responses.

Conclusions

The FDS is a valid and reliable tool to measure the functional disability of patients with dystonia—in particular BEB, HFS and AFR. Given the ease of administration as well as accuracy, we recommend that the FDS is utilised more readily in clinical practice as an alternative option to the BSDI which will further validate its use.

Acknowledgments

Footnote

Reporting Checklist: The authors have completed the STROBE reporting checklist. Available at https://aes.amegroups.com/article/view/10.21037/aes-22-42/rc

Data Sharing Statement: Available at https://aes.amegroups.com/article/view/10.21037/aes-22-42/dss

Peer Review File: Available at https://aes.amegroups.com/article/view/10.21037/aes-22-42/prf

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at https://aes.amegroups.com/article/view/10.21037/aes-22-42/coif). The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. This study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). The South West-Central Bristol Institutional Review Board approved this study with a reference number of IRAS-297057. All patients provided written informed consent, and understood and complied with the study requirements.

Open Access Statement: This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: https://creativecommons.org/licenses/by-nc-nd/4.0/.

References

1. Ross AH, Elston JS, Marion MH, et al. Review and update of involuntary facial movement disorders presenting in the ophthalmological setting. Surv Ophthalmol 2011;56:54-67.
2. Bladen JC, Feldman I, Favor M, et al. Long-term outcome of flexible onabotulinum toxin A treatment in facial dystonia. Eye (Lond) 2019;33:349-52.
3. Jankovic J, Orman J. Botulinum A toxin for cranial-cervical dystonia: a double-blind, placebo-controlled study. Neurology 1987;37:616-23.
4. Lawes-Wickwar S, McBain H, Brini S, et al. A patient-initiated treatment model for blepharospasm and hemifacial spasm: a randomized controlled trial. BMC Neurol 2022;22:99.
5. Trinchillo A, Iorillo F, De Joanna G, et al. The Impact of the reclusion on patients with blepharospasm during the covid-19 pandemic. Clin Neurol Neurosurg 2022;221:107363.
6. Timlin HM, Jiang K, Ezra DG. Impact of Upper Eyelid Surgery on Symptom Severity and Frequency in Benign Essential Blepharospasm. J Mov Disord 2021;14:53-9.

7. Yahalom G, Janah A, Rajz G, et al. Therapeutic Approach to Botulinum Injections for Hemifacial Spasm, Synkinesis and Blepharospasm. Toxins (Basel) 2022;14:362.
8. Scott AB, Kennedy RA, Stubbs HA. Botulinum A toxin injection as a treatment for blepharospasm. Arch Ophthalmol 1985;103:347-50.
9. Fahn S, List T, Moskowitz C, et al. Double-blind controlled study of botulinum toxin for blepharospasm. Neurology 1985;35:271-2.
10. Lindeboom R, De Haan R, Aramideh M, et al. The blepharospasm disability scale: an instrument for the assessment of functional health in blepharospasm. Mov Disord 1995;10:444-9.
11. Goertelmeyer R, Brinkmann S, Comes G, et al. The Blepharospasm Disability Index (BSDI) for the assessment of functional health in focal dystonia. Clin Neurophysiol 2002;113:S77-S78.
12. Jankovic J, Kenney C, Grafe S, et al. Relationship between various clinical outcome assessments in patients with blepharospasm. Mov Disord 2009;24:407-13.
13. Grivet D, Robert PY, Thuret G, et al. Assessment of blepharospasm surgery using an improved disability scale: study of 138 patients. Ophthalmic Plast Reconstr Surg 2005;21:230-4.
14. Bolarinwa OA. Principles and methods of validity and reliability testing of questionnaires used in social and health science research. Niger Postgrad Med J 2015;22:195-201.
15. Nunnally JC, Bernstein IH. Psychometric Theory, 3rd edn. New York: McGraw-Hill, Inc.; 2006.
16. McGraw KO, Wong S. Forming inferences about some intraclass correlation coefficients. Psychol Methods 1996;1:30-46.
17. Jankovic J, Comella C, Hanschmann A, et al. Efficacy and safety of incobotulinumtoxinA (NT 201, Xeomin) in the treatment of blepharospasm-a randomized trial. Mov Disord 2011;26:1521-8.
18. Parkin M. Priming. In: Lavrakas PJ. editor. Encyclopedia of Survey Research Methods. London: SAGE Publications; 2008.
19. Bowling A. Mode of questionnaire administration can have serious effects on data quality. J Public Health (Oxf) 2005;27:281-91.