Movement Disorders
Vol. 12, No. 6, 1997, pp. 1013–1018
© 1997 Movement Disorder Society

DYSBOT: A Single-Blind, Randomized Parallel Study to Determine Whether Any Differences Can Be Detected in the Efficacy and Tolerability of Two Formulations of Botulinum Toxin Type A—Dysport and Botox—Assuming a Ratio of 4:1

Cristina Sampaio, *Joaquim J. Ferreira, †Fernanda Simões, ‡Maria J. Rosas, ‡Marina Magalhães, ‡Ana P. Correia, ‡António Bastos-Lima, †Raimundo Martins, and *Alexandre Castro-Caldas

*Institute of Pharmacology and General Therapeutics, Faculty of Medicine, University of Lisbon, Lisbon; *Department of Neurology, Faculty of Medicine, Santa Maria Hospital and University of Lisbon, Lisbon; †Department of Neurology and Neurosurgery, S. João Hospital, Porto; and ‡Department of Neurology, St. António Hospital Porto, Portugal*

Summary:

Background: Elston and Russell discovered a difference in the biological potency of the English formulation of botulinum toxin type A or BTX-A (Dysport) and the American formulation (Botox). Potency of both is expressed in LD50 mouse units, but because of assay differences, these units are not equivalent. Since the first warning by Quinn and Hallet on the clinical importance of this issue, it has been impossible to reach a consensus on the conversion factor for the potency of these formulations.

Objective: To test the hypothesis that the conversion factor for the clinical potency of Dysport to Botox is approximately 4:1. DYSBOT is an acronym that results from adding "DYS" from Dysport with "BOT" from Botox.

Patients and Methods: *Design:* A single-blind, randomized, parallel comparison. A total of 91 patients with blepharospasm or hemifacial spasm were randomized to treatment with Dysport or Botox using a fixed potency ratio of 4:1. *Clinical evaluations:* The patients were evaluated at baseline (day of the treatment), 1 month after treatment, and whenever the effect was judged to be fading. Objective and functional rating scales were used as quantitative measures of the change in clinical status. Adverse reactions were collected using a systematic questionnaire.

Results: Using this ratio between products, both Dysport and Botox groups produced similar clinical efficacy and tolerability. For patients showing a positive response without the need of a booster, the duration of effect was 13.3 ± 5.9 weeks for the Dysport group and 11.2 ± 5.8 weeks for the Botox group. Of 48 patients, 11 (23%) needed booster treatment in the Dysport group compared with five (12%) of 43 in Botox group. Adverse events were noted in 24 (50%) of 48 patients in the Dysport group and 20 (47%) of 43 of the Botox-treated group.

Conclusions: Using a 4:1 conversion ratio for Dysport and Botox, similar results were obtained for the two treatments in an appropriately powered study, suggesting that this conversion factor is a good estimate of their comparative clinical potencies.
Key Words: Botox—Botulinum toxin type A—Drug potency—DYSBOT—Dysport—Equivalence of formulations.

In 1985, Elston and Russell (1) discovered a difference in the biological potency between the English formulation of botulinum toxin type A or BTX-A (Dysport) and the American formulation (Botox) even though the potency is expressed in LD_{50} mouse units (U) for both. A lack of awareness of this difference may have disastrous consequences in clinical practice. If there is no correction of the respective doses, patients may experience either effects several times more intense than the usual if they are switched from Dysport to Botox, or a failed treatment if they are transferred from Botox to Dysport. This has implications on safety as well as cost, since the treatments are expensive and need to be optimized.

Quinn and Hallet (2) raised this concern in 1989 and, since then, several attempts have been made to find an appropriate conversion factor for the potencies of the two formulations. Different experimental paradigms have been used in these attempts (3,4), and an exact consensus

Received March 1, 1996; revisions received January 3 and February 28, 1997. Accepted March 3, 1997.
Address correspondence and reprint requests to Dr. C. Sampaio at Instituto de Farmacologia e Terapeutica Geral, Faculdade de Medicina de Lisboa, Hospital de Santa Maria, 1600 Lisboa, Portugal.

has not been reached (5–7). Causes for the differences in biological potency are debatable. Some authors argue that the vehicle used for dilution in the animal assay is an important factor (3) and that, without a carrier protein at the very low dilution used during the assay, some of the activity may be lost. If these data are to be believed, Botox may underestimate the number of units in the vial by some threefold. Others state that the reason is the process of BTX-A formulation (6), and still others claim that the differences are artifacts caused by the intrinsic variability of the assay used (8). From a clinical viewpoint, the difference in potency is real and of concern, and some easily used conversion factor is important. Based on a previous pilot study (9), we conducted a single-blind, randomized, parallel, comparative trial between the two formulations—Dysport and Botox—to test the hypothesis that, at a conversion factor of approximately 4:1, the clinical effects are indistinguishable.

PATIENTS AND METHODS

Setting

Participating were movement disorder outpatient clinics in three teaching hospitals: Santa Maria Hospital (coordinator) in Lisbon and Santo António Hospital and S. João Hospital in Porto.

Population

The patients studied were those with a diagnosis of blepharospasm or hemifacial spasm who were referred to these clinics and who fit the following inclusion criteria: absence of a treatable cause of the syndrome; no previous treatment with botulinum toxin; age >18 years; capability of attending the scheduled visits; and willingness to provide informed consent. Exclusion criteria were existence of systemic disease that could confound the evaluations; concomitant oral medications that could interfere with the action of BTX-A, for example, aminoglycosides, baclofen, dantrolene, or diazepam; known hypersensitivity to any of the components of BTX-A formulations; pregnant women or women of fertile age without effective contraception; and inability to comply with the scheduled visits.

Objective

The objective was to compare the efficacy and the tolerability of two different formulations of BTX-A, assuming a conversion factor between their biological potencies of 4, that is, 4 LD_{50} U of Dysport equivalent to 1 LD_{50} U of Botox. The factor of 4:1 was chosen on basis of a previous pilot study that suggested that the conversion factor would be slightly less than 4:1 (9).

Design

Patients were randomized into two groups. One group received treatment with Dysport formulation according to the following predefined scheme: blepharospasm, 10 LD_{50} U per point, five points per eye, 100 LD_{50} U total dosage; and hemifacial spasm, 10 LD_{50} U per point, five points around the affected eye and two points administered to the lower face, 70 LD_{50} U in total. The other group received Botox according to the following predefined scheme: blepharospasm, 2.5 LD_{50} U per point, five points per eye, 25 LD_{50} U total dosage; and hemifacial spasm, 2.5 LD_{50} U per point, five points around the affected eye and two points in the lower face, 17.5 LD_{50} U in total. The dilutions were made according to the instructions provided by the respective companies in their package-insert leaflet. Each vial of Dysport (500 LD_{50}) was diluted with 2.5 mL saline, and each vial of Botox (100 LD_{50}) was diluted with 4 mL saline.

Randomization

The randomization was completed according to a table of random numbers. The randomization process was centralized and accessible by phone.

Clinical Evaluations

The evaluations were single blinded. Patients were evaluated at baseline (day of the treatment) and 1 month after the treatment—at this visit, the treatment could be judged as failed or insufficient on clinical grounds and the decision was made as to whether a booster injection was needed. If the effect was judged to be insufficient, a booster was recommended. If the treatment was judged effective, however, the patient was scheduled to be evaluated again when symptomatic benefit had started to fade.

The percentage change in the blepharospasm-rating scale (10) was used as a quantitative measure of the modification of the clinical status. The blepharospasm-rating scale takes into account a number of factors: (a) the number of muscles involved; (b) the situation that determines the involuntary movements (at rest or only during activity); (c) the influence of factors (for example, sunlight or watching television); (d) the frequency of involuntary movements graded by an ordinal score from 0 to 5; (e) the severity of involuntary movements, discriminating the upper face from the lower face, each segment graded according to an ordinal score from 0 to 4 for the upper face and from 0 to 3 for the lower face; and (f) disability measured as a percentage of normal function. The different factors taken into account by the blepharospasm-rating scale represent several subscores of the scale. We considered *global improvement* the per-

centage reduction of the global score, which means all subscores added. The disability measured as a percentage of normal function was designated the *functional score.* Finally, the subscores relating to the frequency of involuntary movements and intensity of these movements added represented the *efficacy score.*

Adverse reactions were documented at each visit by the completion of a systematic questionnaire that specifically surveyed complete ptosis, incomplete ptosis, facial paresis, and blurred vision.

Variables

The primary variables of this study were defined as both the *duration of effect,* defined as the interval between the day of treatment and the date the patient reported a waning of effect, and the *number of boosters needed.* These have been judged to provide the most sensitive measures, and the sample size for the study had been set based on the duration of effect. Secondary variables were *latency of effect,* defined as the interval between the day of treatment and the start of symptomatic relief; *clinical efficacy,* measured as the relative improvement (percentage) in the rating scales; and *frequency of adverse reactions.*

Statistics

The sample size of ~20 patients/group was calculated assuming a difference between treatments in the duration of effect (9) of 1.7 months with a variance of 2.3 months, and an α error equal to 0.05 and a β error equal to 0.10. The sample size was doubled ($n_{total} = 80$) after 1 year, when an interim analysis showed no differences and an increase in the power of the study was considered advisable. Accordingly, the study was powered to show a difference of 3.8 weeks in the duration of effect with a sensitivity of 80% or a difference of 5.0 weeks with a sensitivity of 95%.

The null hypothesis for duration of effect and latency of effect was tested with the nonpaired Student's *t* test. The null hypothesis for the frequency of boosters and frequency of adverse reactions was tested with the chi-squared test, and the null hypothesis for clinical efficacy was tested with the Kruskal–Wallis test.

Ethics

The protocol was approved by the Ethics Committee of the Lisbon Faculty of Medicine/Hospital Santa Maria.

RESULTS

A total of 91 patients were enrolled. The principal characteristics of the patients studied and their distribution by the groups are summarized in Tables 1 and 2. The baseline characteristics were comparable between the two groups, and none of the parameters differed significantly between the two groups. The distribution of the sexes and the duration of the disease (Table 1) suggested some asymmetry, but these did not reach statistical significance and it is unknown whether they might have any clinical significance. Because of their clinical presentation, some patients required a slight modification to the scheme of treatment as defined in the protocol. In some cases of hemifacial spasm, there was no need to inject the lower face. In some cases of blepharospasm, muscles other than the orbicularis oculi needed treatment. All of the patients enrolled were therefore analyzed in the intention-to-treat analysis, but, in addition, those who complied exactly with the protocol were included in the "on treatment" analysis. In the blepharospasm group, all except two patients in the Botox group did comply with the protocol. In the hemifacial spasm group, 18 patients did not comply with the protocol: 12 of the Dysport group and six of the Botox group.

TABLE 1. *Demographic and clinical characteristics of the patients studied at baseline*

	Dysport (n = 48)	Botox (n = 43)
Age (mean ± SD) (years)	58.2 ± 11.0	63.2 ± 12.7
Sex (male/female)	10/38	17/26
Severity at baseline		
Efficacy score	7.1 ± 2.0	7.9 ± 2.0
Global score	14.9 ± 3.5	15.7 ± 3.7
Duration of the disease (mean ± SD) (years)	6.13 ± 7.25	3.99 ± 4.17

SD, standard deviation.

The frequency of booster treatments was 11 (24%) of 48 in the Dysport group and five (12%) of 43 in the Botox group ($\chi^2 = 0.16$, $p = 0.26$). The dose used in the booster treatments was similar, but not always the same, as the one used in the first treatment: For blepharospasm treated with Botox, the dose was in all cases 25 LD_{50} U; for hemifacial spasm, it was 7.5 LD_{50} U. For blepharospasm treated with Dysport, the dose was in all cases 100 LD_{50} U; whereas for hemifacial spasm, the mean dose was 45.0 ± 7.1 LD_{50} U.

The primary variable—duration of effect—was analyzed for those patients who did not need booster treatments, and the results were equivalent in the two groups (Table 3). The mean duration of effect was 13.3 weeks in

TABLE 2. *Distribution of the patients studied by diagnosis*

	Dysport	Botox
Blepharospasm	21	21
Hemifacial spasm	27	22

TABLE 3. *Analysis of the duration of effect variable (weeks)*

	Dysport	Botox	Student's *t*
Intention-to-treat analysis (mean ± SD)	12.8 ± 5.6 μCI (10.5; 14.9)	13.1 ± 11.8[a] μCI (8.1; 13.6)	p = 0.91
Difference of means	1.85 difCI (−1.53; 5.23)		
On treatment analysis (mean ± SD)	13.3 ± 5.9	11.2 ± 5.8	p = 0.27

SD, standard deviation; μCI, confidence interval of the mean; difCL, confidence interval of difference.

[a] The greater SD is due to the presence of an outlier. Without this outlier, the value comes down to 10.9 ± 5.9, but the statistical conclusion still holds, p = 0.27.

the Dysport-treated group and 11.2 weeks in the Botox group. The analysis of both primary variables showed considerable variability between patients' response to toxin in both groups, but favored the statistical and clinical equivalence of the two preparations at this ratio. Since a greater number of patients who had hemifacial spasm did not comply with the protocol compared with the blepharospasm group, the results of the hemifacial spasm group alone are presented in Table 4; the equivalence of the 4:1 ratio still holds true in this subgroup. The secondary variables showed the same tendency toward equivalence of the two formulations (Table 5).

The frequency of adverse reactions did not differ between the two groups: Dysport 50% and Botox 47%. The most prevalent type of adverse reaction was incomplete ptosis for patients with blepharospasm and facial paresis for patients with hemifacial spasm.

DISCUSSION

Our study showed that, when a relationship of 4:1 between the biological potency of Dysport and Botox is assumed, pattern of duration of effect, efficacy, and tolerability is similar. This means that, on clinical grounds, 4 U Dysport corresponds approximately to 1 U Botox.

Duration of effect rather than efficacy was chosen as the primary variable because previous studies (9) and the general literature show that the former is a more sensitive variable. Our two primary variables could be considered to be mutually exclusive, since the need for a booster injection precludes an accurate measurement of duration of effect. However, this does not jeopardize our conclusions because, when a patient needed a booster, this was taken into account in the booster analysis, and only those patients not requiring boosters contributed data to the evaluation of duration of effect. In fact, our primary variables are complementary. Both support a 4:1 relationship, which suggests that it is probably a reasonable working ratio. The secondary variables were analyzed without separating any of the effects of a second booster dose, because the booster was given only where the first injection was clinically ineffective. This approach is in keeping with normal clinical practice. The comparative study of efficacy between formulations is therefore a comparison of their general efficacy. We have shown that there were no statistically significant differences between the frequency of boosters even though the crude numbers point toward more boosters in the Dysport group. This study uses a fixed low dose for all new patients, whereas some are known to require careful upward dose titration to achieve optimal effect. It is not possible to predict potential responsiveness in these patients at their first treatment.

TABLE 4. *Results relating to the subgroup of hemifacial spasm patients*

	Duration of effect (mean ± SD)		Frequency of boosters	
	Intention to treat	On treatment	Intention to treat	On treatment
Dysport	13.0 ± 6.3 (n = 27)	13.9 ± 7.0 (n = 15)	7.4% (2/27)	13.3% (2/15)
Botox	12.8 ± 6.6 (n = 22)	13.4 ± 6.5 (n = 16)	4.5% (1/16)	6.3% (1/22)

SD, standard deviation.

The 4:1 relationship that we found is similar to the 3:1 obtained by Marion et al. (11) and Whurr et al. (12). Our study may be more persuasive than theirs because it comes from the comparison of first treatments in a population of similar disease severity and is well randomized. Marion et al. (11) studied patients with the same type of diagnosis as we did. Their population, though, was preselected, having been previously treated several times with Dysport and then titrated to an equivalent effect of Botox in successive treatments. Whurr et al. (12) compared two nonconcurrent, nonrandomized groups of patients with laryngeal dystonia. The sample size (n = 10) was too small to show any clinically meaningful difference between the two formulations. Only the very striking differences such as those occurring in the 1:1 comparison of Dysport and Botox were apparent. They found a relationship of 3:1 because it was the ratio they tested, but the study was not sensitive enough because of the small sample size.

In our study, the robustness of our conclusions regarding the 4:1 ratio depended on the power of the study. Our study was powered to show a difference in duration of effect of 3.8 weeks with a sensitivity of 80% or a difference of 5 weeks with a sensitivity of 95%. Our study demonstrates that a difference between treatments of this

TABLE 5. *Analysis of the secondary variables*

	Dysport	Botox	Statistics
Latency of effect (days)			Student's *t*
Intention-to-treat analysis	5.3 ± 6.7	4.4 ± 4.1	p = 0.45
On-treatment analysis	5.9 ± 7.5	4.4 ± 4.3	p = 0.31
Clinical efficacy			
Efficacy score			Kuskall–Wallis test
Intention to treat analysis	47.2 ± 35.9	46.5 ± 26.8	p = 0.35
On treatment analysis	40.3 ± 35.2	44.3 ± 24.3	p = 0.51
Functional score			
Intention-to-treat analysis	48.6 ± 41.9	33.8 ± 34.4	p = 0.20
On-treatment analysis	44.3 ± 42.2	31.1 ± 31.5	p = 0.27
Frequency of adverse reactions			Chi squared
Intention-to-treat analysis	24/48	20/43	p = 0.74

magnitude is not present, and does so with a fair degree of certainty.

Factors that could have affected the results in our study and other studies include the batches of product used and the dilutions. The manufacturing standards of the Food and Drug Administration and other regulatory authorities allow biological drugs to vary in potency by 25% between batches. If this was applied by the manufacturers, each vial of Dysport, although marked as 500 U, could contain somewhere between 375 and 625 U, and Botox could contain 75–125 U. These variations, of course, are very crude, and both companies manufacture vials with tighter specifications, which would make them much closer to the quantities stated on the label. In our study, however, we did not record the exact lot numbers and did not factor this additional variability into our analysis. Additionally, we took the risk of using different dilutions for each formulation because this reflects the use of these drugs in normal clinical practice. We therefore used the optimal concentration as recommended by the companies, recognizing that this could have influenced the findings. The last and perhaps the most important factor influencing our results is the large interindividual variation, which was reflected in large standard deviation of the mean. Taking all of these factors into consideration, we recognize that we cannot accept the 4:1 ratio as an exact measure, but we can accept it as a good indicator. This reasoning allows us to consider the results of Marion et al. (11), Whurr et al. (12), and our own data (9) together, which on the whole seem to be consistent with the present study and most laboratory experiments (3,4,6,7).

In this context, the study by Durif (13) is rather contradictory, since he found a factor of 5–6:1 in the comparison between Dysport and Botox. What is unusual in this is that it is the only data suggesting a ratio of >4. It is hard to accept the result of Durif's study as accurate. His study included a nonrandomized parallel comparison between Dysport and Botox with numerically unbalanced groups of n = 17 and n = 22 (Botox) versus n = 9 and n = 12 (Dysport), respectively, for patients with blepharospasm and those with hemifacial spasm. The dosages in each group were also allowed to vary widely. Duriff's study should be dismissed on these grounds and perhaps also for not having been published in an unbiased forum.

Our trial did not address the problem of the putative differences in immunogenicity between the two formulations. However, our results are also important for future studies concerning immunogenicity. Accepting the ratio of 4 LD_{50} U Dysport to 1 LD_{50} U Botox toxin–hemagglutinin complex means that the patients treated with Botox receive a quantity of complex that is 10 times greater than the quantity received by Dysport patients. This difference in the size of inoculum may have immunological significance for our patients' long-term treatment.

Our results emphasize another technical aspect. The orbicularis oculi muscles were injected according to a protocol that defines five different injection points: two points in the orbital portion of the superior eyelid below the eyebrow (medial and lateral), two points above and below the lateral insertion of orbicularis oculi muscles, and one on the lateral portion of the lower eyelid. With this distribution, we obtained meaningful and sustained symptomatic relief from blepharospasm with dosages of Dysport (100 LD_{50} U in total) that are several times less than those that are recommended in the package inserts (240 LD_{50} U in total), based mainly on the work of Elston (14). We are aware, though, that much lower doses of Dysport are also effective if the injections are closer to the lid margin (15). Among Botox dosages, there are less remarkable differences, since the total dosage used in our study was 25 LD_{50} U, the recommended dosage in the package inserts is 20 LD_{50} U, and the dosage used in standard practice is 12.5–30.0 LD_{50} U

(16). These findings suggest that further controlled studies are needed in order to achieve any standardization of the injection sites and dosages for treatments with BTX-A.

In conclusion, the 4:1 ratio is a fair estimate of the conversion factor of biological potency, expressed in LD_{50} U, between Dysport and Botox formulations. Further work is needed to standardize minimum and optimum effective dosages for either formulation and to determine how best to optimize other aspects of treatment with the toxin.

Acknowledgment: This study was funded by the Centro de Estudos Egas Moniz.

REFERENCES

1. Elston J, Russell RW. Effect of treatment with botulinum toxin on neurogenic blepharospasm. *BMJ* 1985;290:1857–1859.
2. Quinn N, Hallett M. Dose standardisation of botulinum toxin. *Lancet* 1989;1:964.
3. Pickett AM. Potency equivalence of botulinum toxin preparations. *J R Soc Med* 1994;87:719.
4. Borodic GE. Botulinum toxin potency: a mystery resolved by the median paralysis unit. *J R Soc Med* 1994;87:571–572.
5. Sesardic D, Das R, Corbel MJ. Botulinum toxin. *J R Soc Med* 1994;87:307.
6. Borodic GE. Therapeutic botulinum toxin. *Lancet* 1994;344:1370.
7. Pickett AM, Hambleton P. Dose standardisation of botulinum toxin. *Lancet* 1994;344:474–475.
8. Schantz EJ, Johnson E. Dose standardisation of botulinum toxin. *Lancet* 1990;335:421.
9. Sampaio C, Ferreira JJ, Castro-Caldas A. Clinical efficacy of botulinum toxin: American versus British toxin [Abstract]. In: *Proceedings of the joint meeting of the British and Portuguese Societies of Neurology, Lisbon, November 1992.*
10. Fahn S. Assessment of the primary dystonias. In: Fahn S, Munsat T, eds. *Quantification of Neurologic Deficit;* vol 19. Boston: Butterworths, 1989:241–310.
11. Marion MH, Sheey M, Sangla S, Soulayrol S. Dose standardisation of botulinum toxin. *J Neurol Neurosurg Psychiatry* 1995;59:102.
12. Whurr R, Brooks G, Barnes C. Comparison of dosage effects between the American and British botulinum toxin A products in the treatment of spasmodic dysphonia [Abstract]. *Mov Disord* 1995; 10:387.
13. Durif F. Clinical bioequivalence of current commercial preparations of botulinum toxin [Abstract]. In: *Second European botulinum toxin symposium, Cannes, March 17–9, 1995.*
14. Elston J. Idiopathic blepharospasm, hemifacial spasm and therapeutic ptosis induction. In: Moore P, ed. *Handbook of Botulinum Toxin Treatment;* vol 6. Oxford: Blackwell Science, 1995:90–100.
15. Arimideh M, Ongerboer de Visser BM, Brans JWM, Koelman JHTM, Speelman JD. Pretarsal application of botulinum toxin for the treatment of blepharospasm. *J Neurol Neurosurg Psychiatry* 1995;59:309–311.
16. Dutton JJ. Acute and chronic effects of botulinum toxin in the management of blepharospasm. In: Jankovic J, Hallet M, eds. *Therapy with Botulinum Toxin;* vol 14. New York: Marcel Dekker, 1994:199–210.