Measurement of Botulinum Toxin Activity: Evaluation of the Lethality Assay

L. Bruce Pearce,* Gary E. Borodic,† Eric R. First,* and Robert D. MacCallum*  

*Department of Pharmacology, Boston University School of Medicine, Boston, Massachusetts 02118; †Department of Ophthalmology, Massachusetts Eye and Ear Infirmary, Harvard Medical School, Boston, Massachusetts 02114

Received November 1, 1993; accepted May 9, 1994


The use of the mouse lethality assay for the estimation of the biologic activity of botulinum toxin was evaluated. The relationship between the number of animals, number of doses, and duration of the assay used to estimate the LD50 and the precision of the assay was investigated. The results of these studies demonstrated that the LD50 for botulinum toxin can be estimated with a high degree of precision (±5%). The precision of the assay is not increased by using more than a 5-dose 50-animal assay or extending the duration of the assay beyond 72 hr. Estimates of the LD50 obtained at 48 hr were only slightly less precise but underestimated the LD50 by 15%. Analysis of the commercially available preparations of botulinum toxin with the mouse LD50 assay revealed significant discrepancies between the units of toxin in these preparations. In addition, a 2.67-fold difference in the relative potency of the two preparations of botulinum A toxin was observed using a regional chemodenervation assay that measures paralysis. The mouse LD50 assay could not detect this large difference in the potency of the two approved clinical preparations of botulinum toxin. The results of these studies demonstrate that although the mouse LD50 assay can be used to estimate the number of units of botulinum toxin with a high degree of precision this assay alone is not an adequate method for assessing the preclinical biological potency of botulinum toxin. © 1994 Academic Press, Inc.

Botulinum toxin is a paralytic neurotoxin that exists as seven different serotypes (A, B, C, D, E, F, and G) elaborated by a number of bacterial species belonging to the genus Clostridium (Hattheway, 1989). The type A toxin, produced by Clostridium botulinum, initially received attention because of the toxic and lethal effects in humans associated with foodborne, infant, and wound botulism. Since the pioneering work of Schantz and Scott in the 1970s and 1980s, clinical interest in botulinum A toxin has turned from toxicology to therapeutics (Schantz and Scott, 1981; Schantz and Johnson, 1992). This is due to the beneficial pharmacological properties associated with low-dose injection of a crystalline preparation of botulinum A toxin isolated by Schantz and Johnson, 1992). This potent neurotoxin has become widely investigated for its therapeutic potential in the treatment of a variety of neuromuscular disorders (Simpson, 1981; Jankovic and Brin, 1991; Borodic et al., 1991; Schantz and Johnson, 1992). Intramuscular injection of nanogram quantities of botulinum toxin is the treatment of choice for a number of clinical indications including blepharospasm, spasmodic dysphonia, hemifacial spasm, and adult-onset spasmodic torticollis. Clinical indications for the use of this reversible chemodenervating agent are growing in number with investigations of the use of this toxin in the management of cerebral palsy, jaw dystonia, stuttering, occupational limb spasms, neurogenic bladder, and spasticity.

The expanding medical importance of botulinum toxins has increased the need for, and placed a premium on, the precise analysis of biologic activity contained in preparations of the toxin for both clinical use and laboratory investigation. A variety of methods have been employed to measure the biologic activity associated with botulinum neurotoxins (Notermans and Nagel, 1989; Mayorga et al., 1987; Sugiyama et al., 1975). The primary method for the measurement of botulinum toxin activity is estimation of the median lethal dose, LD50, in mice as described by Schantz and Kautter (1977). The median lethal dose has been determined for a very large number of chemicals, solvents, cosmetics, and drugs. However, estimation of the LD50 is notoriously variable and the routine use of this test is no longer justified (DePass, 1989). A number of regulatory agencies have abandoned requiring the routine use of the classical LD50 for toxicity testing. In this climate of declining acceptance of the LD50 as a biological constant, the LD50 is still used to define the fundamental unit of biologic activity of botulinum toxin. One of the goals of this report was to evaluate the use of the estimated LD50 as a reliable measure of the biologic activity of botulinum toxin.

The fundamental unit (unit, U) of biologic activity of botulinum toxin has been defined as the LD50 (1.0 LD50 =
1.0 U) for the toxin in a population of mice. Whether estimation of the LD50 alone is an adequate predictor of the clinical potency of various preparations of botulinum toxin is unclear. Comparison of the clinical potency observed for the two preparations of botulinum A toxin approved for use in the United States (Botox) and England (Dysport) raises this important question. Recommended doses for the treatment of blepharospasm with Botox are 10–30 U/eye, whereas 100–200 U/eye are recommended for Dysport. This large disparity in the required dosing is also seen with the treatment of torticollis where total doses of 100–280 U versus 875–1200 U have been reported for Botox and Dysport, respectively (Clarke, 1992). Both of these preparations are standardized using the classical mouse assay, the LD50 determination (Schantz and Johnson, 1990; Quinn and Hallet, 1989). Accordingly, the number of units (LD50s) of botulinum A toxin contained in the commercially available preparations of the toxin were analyzed and compared to determine if the differences in the reported clinical potency of the two preparations can be attributed to interlaboratory differences in the LD50 determinations used to define the two clinical preparations.

Preparations of botulinum A toxin were also compared by measuring the paralysis of the hind limb of a mouse produced following im administration of toxin. Utilizing this assay, a large difference in the potency of these two preparations of the toxin was revealed. These and other results will be discussed in terms of the appropriateness of the LD50 for estimating the preclinical biological activity of botulinum toxin.

**METHODS**

CD1, virus- and antibody-free, white male mice weighing 18–22 g were obtained from Charles River Laboratories. Type A porcine skin gelatin was obtained from Sigma Chemical Co. (St. Louis, MO). Botulinum toxin type A (100 U/vial) was provided in a freeze-dried formulation from Associated Synapse Biologics (Boston, MA). Botox was obtained from Allergan Pharmaceuticals (Irvine, CA) and Dysport was obtained from Porton Products Ltd. (Berksire, England).

**The lethality assay.** Samples of lyophilized botulinum toxin were solubilized in 0.5 ml of sterile 0.2% type A gelatin in 30 mM sodium phosphate buffer, pH 6.2. Schantz (1964) has recommended the addition of gelatin to buffer to stabilize dilute preparations of the toxin. This resuspension was allowed to equilibrate for 15 min prior to dilution. The reconstituted toxin was diluted into culture tubes containing 0.2% gelatin in 30 mM sodium phosphate buffer, pH 6.2, to give either 5 or 10 dilutions that contained between 0.45 to 1.8 U of toxin per 0.5 ml. Assays in which 25 or 50 mice were injected, 5 dilutions of toxin, and 5 or 10 mice were used per dilution, respectively. Dilutions were increased in approximately a geometric progression to achieve a symmetric design (Finney, 1978). Several different series of dilutions of toxin were used and the ratio between successive doses was approximately 1:2.5. The dose range was between 0.6 and 1.6 units. Ten dilutions were used for the 100 animal assays; the ratio of successive doses was 1:1.2 to 1:1.7. The relatively small ratio of successive doses was needed because of the steepness of the dose–response curve for toxin-induced lethality. The dose range for the 100 animal assay was between 0.45 and 1.8 units. Diluted samples of toxin were administered by ip injection to mice weighing 18–22 g. Following injection the mice were observed for 5 days.

The percentage of death was determined at each dose of toxin and probit analysis was performed on the data using the probit (Bliss, 1938) program provided with the statistical package, SPSS-X (Chicago, IL). This program estimates the best line by regression analysis and the values for the intercept and slope are estimated by the maximum likelihood method. The 95% fiducial confidence intervals for the estimates of the LD50 (Trevan, 1927) were determined as described by Finney (1971). A Pearson chi goodnes of fit test was used and if this estimate was significant, a heterogeneity factor was used in the calculation of the confidence limits. Unless stated otherwise, a freeze-dried formulation of botulinum type A toxin supplied by Associated Synapse Biologics, Inc. was used in the experiments evaluating the lethality assay.

**The hind limb regional chemodenervation assay.** Paralysis of the mouse right hind limb produced by im injection of botulinum toxin was evaluated in a manner similar to that described by Sugiyama et al. (1975). Toxin was diluted into culture tubes containing 0.2% gelatin in 30 mM sodium phosphate buffer, pH 6.2, and 0.1 ml of diluted toxin preparations was injected into the gastrocnemius muscle of the right hind limb of mice weighing 18–22 g. The degree of paralysis was evaluated on a daily basis for 12 days. Paralysis was scored: 0 for no evidence of paralysis, 1.0 for partial paralysis of the limb, and 2 for complete paralysis.

**Comparison of the clinical formulations of botulinum type A toxin.** The two clinical preparations of botulinum type A toxin, Botox and Dysport, were compared using both the lethality and mouse hind limb assays. Estimation of the LD50 for these two formulations involved analysis of the activity in three vials of Botox and four vials of Dysport. These estimations were performed on two groups of vials with different expiration dates for both formulations of A toxin. Comparison of the two formulations of botulinum A toxin using the hind limb assay involved a total of three determinations performed on the toxin contained in one vial of Botox and a total of three determinations performed on the toxin in two vials of Dysport.

**RESULTS**

The majority of the data presented in this report were analyzed retrospectively. This reduced the number of animals used strictly for the purposes of this report. The data used in the analysis of the LD50 assay were obtained from assays performed using 100, 50, and 25 animals. The results presented below for estimations of the LD50 using 40, 30, 20, and 15 animals were extracted from the 50 and 25 animal assays. This was accomplished by eliminating the observations at one or more of the doses used in the 50 or 25 animal assays.

**Precision of the lethality assay: Effect of number of animals and doses.** The results of LD50 determinations using 100 (A), 50 (B), and 25 (C) mice per assay are shown in Figs. 1A–1C. The scatter diagrams in the left panels contain the data from all the experiments used in this report to evaluate the lethality assay. In the right panels are log dilution versus probit plots of data obtained from several experiments (n ≥ 4) in which the same series of dilutions of test materials were used. In the case of the 25-animal assay all of the data were obtained using the same dilutions of toxin containing samples. The scatter diagram for the 25-animal assay reveals the greater spread of the individual data points.
FIG. 1. Probit analysis of the lethality assay results. Botulinum type A toxin (Associated Synapse Biologies, Inc.) activity in freeze-dried formulations was estimated using the mouse lethality assay. White CD1 male mice weighing 18-22 g were administered 5 to 10 doses of between 0.45 and 1.8 or 0.6 and 1.6 units, respectively, of botulinum A toxin by ip injection. Animals were monitored for 5 days whereupon the experiment was terminated. In the panels on the left the data for the 100 (A, n = 8), 50 (B, n = 9), and 25 (C, n = 10) animal experiments are expressed as the percentage of death observed on Day 5. The probit analysis of data for 100-, 50-, and 25-animal determinations are shown in the panels to the right. Data are shown for only those experiments in which the same series of dilutions were employed. The data shown in this figure represent the results of a total of 27 independent estimates of the LD50 and n refers to the number of estimates for each panel.

not be calculated for all of the assays and thus the number of experiments from which these averages were calculated are shown versus the total number of assays analyzed. The data show that the confidence intervals increase as the number of animals and doses are decreased. Clearly, when the number of mice used in each estimation of the LD50 were increased from 50 to 100, no improvement in the 95% confidence interval was obtained. This represented a doubling of both the total number of animals and doses, whereas when the number of animals and doses were reduced by 25% from 40 to 30, the 95% confidence intervals could be calculated in only 4 out of 9 assays.

The data presented in Table 1 also show the results of reducing the total number of animals and doses on the 95% confidence interval when 5 animals per dose were used. The reliability of individual estimates of the LD50 using only 5 animals per dose was substantially decreased. Confidence intervals could not be calculated for the 3-dose, 15-animal assay. Comparison with the equivalent assays using 10 animals per dose illustrates the effect of reducing the number of animals at each dose. The scatter diagrams (Figs. 1B and 1C) for the 50- and 25-animal assays illustrate the reduced precision that results from reducing the number of animals at each dose. This increase in the 95% confidence limits is consistent with the observed more shallow slope for the corresponding log dilution vs probit plot.

The data presented in Table 1 show the effect of changing both the number of animals and doses on the precision of multiple estimates of the LD50. These data were obtained from the results of assays using either 5 or 10 animals per dose. The results show that the standard deviation increases

---

**TABLE 1**

The Effect of Changing the Number of Animals and Doses on the 95% Confidence Intervals

<table>
<thead>
<tr>
<th>Number of mice</th>
<th>N*</th>
<th>Units/vial</th>
<th>SD</th>
<th>Upper</th>
<th>Lower</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>8 (8)</td>
<td>99</td>
<td>5</td>
<td>+15</td>
<td>-12</td>
</tr>
<tr>
<td>50</td>
<td>9 (9)</td>
<td>98</td>
<td>4</td>
<td>+12</td>
<td>-15</td>
</tr>
<tr>
<td>40</td>
<td>9 (9)</td>
<td>97</td>
<td>4</td>
<td>+30</td>
<td>-21</td>
</tr>
<tr>
<td>30</td>
<td>9 (4)</td>
<td>98</td>
<td>8</td>
<td>+49</td>
<td>-19</td>
</tr>
<tr>
<td>25</td>
<td>10 (5)</td>
<td>96</td>
<td>10</td>
<td>+30</td>
<td>-22</td>
</tr>
<tr>
<td>20</td>
<td>10 (4)</td>
<td>95</td>
<td>11</td>
<td>+56</td>
<td>-39</td>
</tr>
<tr>
<td>15</td>
<td>10 (0)</td>
<td>93</td>
<td>11</td>
<td></td>
<td>*</td>
</tr>
</tbody>
</table>

* N is the number of experiments and the number in parentheses indicates the number of experiments for which 95% confidence limits could be calculated.

**TABLE 1**

- The average 95% confidence limits.
- The average units per vial: 1.0 unit = 1.0 LD50, vials containing botulinum type A toxin were obtained from Associated Synapse Biologies, Inc.
- * Indicates that 95% confidence limits could not be calculated.
only in the case of the 30-animal assay using 3 doses to estimate the LD50. In contrast, the precision of multiple estimates of the LD50 did not differ significantly if 5, 4, or 3 doses were examined when 5 animals were used per dose. However, comparing experiments in which the same number of doses were employed but 5 versus 10 animals were used shows that the precision drops by a factor of 2, with the exception of the 3-dose test.

**Duration of the lethality assay.** One of the perceived drawbacks of the lethality assay is that the assay usually takes 4 to 5 days to complete. Accordingly, estimates of the LD50 were calculated 48, 72, and 96 hr following injection to examine the relationship between the precision of the estimate of the LD50 and the duration of the assay. The results of these experiments are shown in Fig. 2 and Table 2. In Fig. 2, the data show that only a very small fraction of the total deaths occur on Days 4 and 5. Of 473 deaths only 45, <10%, occurred after 72 hr. Analysis of the LD50 values estimated on Days 3, 4, and 5 are shown in Table 2 and indicate that there was no statistically significant difference in the estimated LD50 values over the 3 days. These observations were similar for the 25-, 50-, and 100-animal assays.

**Comparison of the clinical preparations of botulinum A toxin.** The two clinical preparations of botulinum A toxin, Botox and Dysport, were compared using our modification of the mouse LD50 assay (Schantz and Kautter, 1977). The results of these experiments are shown in Table 3. The data obtained from tests performed on 4 vials of Botox gave an average of 127 ± 10 U/vial. In contrast, an average of 332 ± 6 U/vial was observed from tests performed on 3 vials of Dysport. These results indicate that 127 and 66% of the units of activity shown on the label for Botox and Dysport, respectively, were detected by this particular assay. Thus, on the basis of the comparison of these two preparations using the same LD50 assay, a 1.9-fold difference in the clinical potency of the two products would be expected. Table 3 also contains the median doses recommended for the treatment of blepharospasm and torticollis for the two commercially available products that have been corrected on the basis of the 1.9-fold difference from the labeled units. These estimates suggest that even with this correction the recommended doses for the two preparations of botulinum A toxin differ by approximately fourfold for both blepharospasm and torticollis.

To further investigate the basis for this unexplained fourfold difference in the clinical potency of the toxin preparations, these formulations were compared by directly evaluating the extent of denervation of the mouse hind limb. Botox and Dysport were compared following injection of 0.2 or 0.6 units of toxin, respectively. The units of toxin given were determined on the basis of the estimation of the LD50 performed above and not on the labeled contents of these materials. Thus, the doses of the toxins administered to the mice in the paralysis assay were estimated on the basis of the results from exactly the same assay. The results of these experiments are shown in Fig. 3. Paralysis was evaluated over 4 days for Botox and Dysport. Although the doses administered differed by threefold (0.2 and 0.6 units), the effect on the mice was approximately equivalent. The scores recorded in three experiments for Botox were 1.2 ± 0.0 as compared to 1.07 ± 0.1 for Dysport. These data suggest that in addition to the 1.9-fold labeling discrepancy derived from differences in the LD50 estimations, these two preparations of A toxin differ in potency by at least an additional factor of 2.67 ((1.07/1.2)∗(0.6/0.2)), for a total of 5.1 (1.9∗2.67)-fold.

**FIG. 2.** Duration of the mouse assay. White CD1 male mice weighing 18–22 g were administered 5 to 10 doses of between 0.45 and 1.8 or 0.6 and 1.6 units, respectively, of botulinum A toxin (Associated Synapse Biologies, Inc.) by ip injection. Animals were monitored for 5 days and the deaths were recorded on Days 3, 4, and 5. Data are shown for 100 (A, n = 4), 50 (B, n = 4), and 25 (C, n = 6) animal experiments. Error bars represent the standard error of the means from at least four experiments. The data were obtained from a total of 14 experiments and n refers to the number of estimates of the LD50 performed to obtain the data in each panel.
**TABLE 2**
The Effect of the Duration of the Assay on the Precision of the Estimated LD50 for Botulinum Toxin

<table>
<thead>
<tr>
<th>Day</th>
<th>Number of mice</th>
<th>Units/sample average</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>100</td>
<td>95</td>
<td>5</td>
</tr>
<tr>
<td>4</td>
<td>100</td>
<td>99</td>
<td>6</td>
</tr>
<tr>
<td>5</td>
<td>100</td>
<td>100</td>
<td>6</td>
</tr>
<tr>
<td>3</td>
<td>50</td>
<td>91</td>
<td>5</td>
</tr>
<tr>
<td>4</td>
<td>50</td>
<td>98</td>
<td>2</td>
</tr>
<tr>
<td>5</td>
<td>50</td>
<td>100</td>
<td>3</td>
</tr>
<tr>
<td>3</td>
<td>25</td>
<td>99</td>
<td>12</td>
</tr>
<tr>
<td>4</td>
<td>25</td>
<td>100</td>
<td>11</td>
</tr>
<tr>
<td>5</td>
<td>25</td>
<td>101</td>
<td>10</td>
</tr>
</tbody>
</table>

* The average units per vial; 1.0 unit = 1.0 LD50, vials containing botulinum A toxin were obtained from Associated Synapse Biologics, Inc., n ≥ 4 in all cases.

**DISCUSSION**

The results of this study demonstrate that large numbers of animals are not required to obtain very precise estimates of the number of units (LD50s) of botulinum toxin activity. On the basis of a single determination there was no apparent statistical advantage associated with using 100 versus 50 mice for the estimation. Multiple determinations using a 4-dose, 40-mouse assay gave a similar standard deviation to that observed with the 100- and 50-mouse assays. When the number of animals used per assay was decreased from 50 to 25, this resulted in approximately a doubling of the coefficient of variation (4 to 10%). Even with this doubling of the standard deviation, the 25-animal assay still provides a relatively precise estimate of activity. In addition, the average value is not significantly different from that obtained with the 50-animal assay. On the basis of the results reported herein, the 25-animal assay, using 5 mice per dose, may be adequate for most laboratory experimentation when only a single determination of the LD50 is made. However, if multiple determinations are required, the 15-animal assay with a coefficient of variation of 11% is probably adequate, whereas standardization of the units of botulinum activity in formulations of the toxin for clinical use requires greater precision. For this purpose, multiple estimates of the LD50 using either the 40- or 50-animal assay should be optimal.

Schutz and Fuchs (1982) have shown that the number of animals required to obtain significance using the LD50 can be reduced by doubling the duration of the observation period from 4 to 8 weeks. While this may be true in many cases, this is not true for botulinum toxin. Estimates of the LD50 for botulinum toxin have been customarily performed over 4 days. Analysis of data obtained from the 50-animal assay indicates that 71 ± 10% (n = 4) of the animals are dead within 48 hr while 91 ± 6% (n = 4) of the animals were dead within 72 hr following toxin injection. Comparison of the results of the probit analyses showed that the estimates of toxin content at 48 hr (80 ± 6) and 72 hr (93 ± 5) were significantly different (df = 6, n = 4, p < 0.05). Although the coefficient of variation of the estimated LD50 at 48 hr was only 7%, the estimated LD50 was a significant underestimate of the value observed at 3, 4, or 5 days. On the basis of these findings, the recommended minimum duration of the mouse bioassay, as described above, should be 3 days. However, the data available 48 hr postinjection provides an estimate of the LD50 that is within approximately 15% of the final value. Limiting the duration of the assay to 3 days does not increase the number of animals required to achieve the desired precision. Limiting the duration is important to reducing the suffering associated with measuring death as an endpoint.

One approach that has been taken to reduce the number of experimental animals needed for estimating the LD50 is to use a survival time (time-to-death) assay. Beccari first described the time-to-death method in 1949 which was later modified by Molinengo (1979). Meier and Theakston (1986) have argued that the method of Molinengo makes possible the use of “far fewer” experimental animals and that there should be a switch to this approximate method.

**TABLE 3**
Comparison of the Clinical Preparations of Botulinum A Toxin

<table>
<thead>
<tr>
<th>Preparation</th>
<th>Labeled units</th>
<th>Actual units</th>
<th>Actual/label</th>
<th>Corrected† median dose</th>
<th>Ratio‡ (avg)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Botox</td>
<td>100</td>
<td>127 ± 10</td>
<td>1.27</td>
<td>Bleharospasm 20</td>
<td>150</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Toricollis 79</td>
<td>547</td>
</tr>
<tr>
<td>Dysport</td>
<td>500</td>
<td>332 ± 6</td>
<td>0.66</td>
<td></td>
<td>3.80</td>
</tr>
</tbody>
</table>

* Median doses are based on data from Clarke (1992) that were corrected according to the actual number of units in the vials.

† The ratio between the actual units and the labeled units.

‡ The average ratio for the two clinical indications between the median dose for Botox and Dysport.
on the basis of scientific, ethical, and economic reasons. These authors point out that with only 8–10 animals a single test will provide estimates of the LD50 within the same range of statistical variability as the classical method which employs 30 or more experimental animals. However, the linear transformation of the hyperbolic relationship between time-to-death and dose results in a plot of the ratio of time-to-death/dose versus dose which places the dependent variable time on the abscissa which violates one of the assumptions of regression analysis as pointed out by Sevcik (1987). As a result, questions have been raised as to the meaning of the regression analysis of the data and the statistical validity of this assay. Also at issue is whether the value estimated from this method is in any way truly related to the LD50 (Finney, 1985; Sevcik, 1987). The method of Molinengo (1979) differs from the methods described below for the estimation of the LD50 for botulinum toxin, in that the LD50 is determined directly from the time-to-death assay data, whereas the LD50 for botulinum toxin is determined from a standard curve relating time-to-death to dose of toxin in terms of LD50s.

The use of the time-to-death assay for estimation of the LD50 for botulinum toxin was first described by Boroff and Fleck in 1966. The time-to-death assay is thought to require fewer animals and provides information in minutes to hours versus days required with the classical LD50 determination. However, this assay is orders of magnitude less sensitive than the LD50 assay; the former has a range of between 100 to 100,000 LD50 per 0.5 ml for botulinum toxin (Schantz, 1964). Boroff and Fleck (1966) have argued that with as few as 3 mice at each point a coefficient of variation of 14% for the LD50 is observed. Kondo et al. (1984) make the assertion that when the same numbers of animals are used in the time-to-death and the classical ip methods for estimating the LD50, the time-to-death assay has a higher accuracy. DeArmon et al. (1958) observed that 420 mice were needed to obtain a coefficient of variation of 15%, and with 30 mice a coefficient of variation of 60% was observed for the estimation of the LD50 with the time-to-death assay. Schantz (1964) reported that the classical ip method is a more precise assay and that four times the number of animals are necessary with the time-to-death assay to achieve a ±12% accuracy of the estimated LD50. Lamanna et al. (1970) have compared the results of the survival times observed with the high-molecular-weight complex ($M_w$ 900,000) of the toxin versus the neurotoxin ($M_w$ 150,000). These authors demonstrated that the time to death was dependent on the molecular weight of the toxin. These results point out, not unexpectedly, that pharmacokinetic factors may play a significant role in determining the endpoint of the time-to-death assay. Kondo et al. (1984) reported that the time-to-death assay is not affected by the molecular weight of the species of toxin; rather, the serotype of the toxin affects the time to death. These authors also indicated that the time-to-death assay is so variable from day to day that simultaneous titration against a standardized preparation and expression of toxin activity in terms of relative potency was necessary. It appears that significant questions remain regarding the potential use of this assay for either the accurate estimate of the true LD50 or the conservation of experimental animals.

The most economical use of mice for estimation of the LD50 for botulinum toxin can be achieved in several ways. First, establishing standardized and well-controlled procedures for conducting the assay is critical. Second, the choice of the number of animals needed for the estimate of the LD50 should be based on statistical criteria. The required 95% confidence intervals and/or the standard deviation of the estimates of the LD50 should be considered carefully. Finney (1978) has shown that a symmetric design and choice of doses based on the expected response rates are factors that are important to minimizing the variance of the estimated LD50.

Botulinum toxin produces its effects and eventual lethality by a single mechanism of action. Botulinum toxin irreversibly inhibits ACh release from cholinergic nerve terminals producing chemical denervation of the neuromuscular junction and synapses within the central nervous system. Botulinum toxin, in particular the light chain of the toxin, is thought to enzymatically cleave either synaptobrevin or SNAP-25 which are neuronal proteins intimately involved in neurotransmitter release (Schiavo et al., 1992, 1993;
Shone et al., 1993). Although botulism presents with a number of clinical signs, these can be attributed to the inhibition of ACh release. The nonselective actions of the toxin are probably a result of partial or complete blockade at autonomic neurotransmitter sites and ganglia, whereas death results from suffocation due to the action of the botulinum toxin on respiration. Paralysis of the diaphragm results from chemonemotropism of the toxin on autonomic innervation of the diaphragm.

In contrast to botulinum toxin, the lethality associated with many other molecules may be due to the combined effect of these solvents and chemicals on many organ systems and processes important to the maintenance of life. This is further complicated by the fact that lethality may be produced by some of these effects alone or in various combinations. Assuming that the quantal dose–response curve (log dose vs percentage mortality) is the sum of individual dose–response curves for each system that contributes to lethality and the individual variations in the sensitivities of these systems for each animal, the more factors that contribute to the lethality of a particular chemical the more shallow the slope of the dose–response curve and the more variable the results of such analyses. Chemicals that produce lethality by a variety of mechanisms are inherently difficult to characterize by a single measure of toxicity, such as the LD50. The steep log dose–response curve that is typical for the lethality assay for botulinum toxin reflects the remarkable specificity action of this toxin. The relatively high degree of selectivity and specificity of botulinum toxin contributes significantly to the reproducibility of the mouse LD50 assay.

Many factors can affect the LD50 assay such as route of administration, species, strain, age, diet, sex, housing conditions, temperature, seasonal variations, and genetics (Zbinden and Fluri-Reversi, 1981). The data presented in this report were obtained from LD50 determinations performed by one laboratory over the course of more than 1 year. The majority of the factors that affect determination of the LD50 such as the route of administration, species, strain, age, diet, sex, housing conditions, and temperature were under control. Those factors that were not under some degree of control did not dramatically affect the outcome of the determinations of botulinum toxin lethality. The small standard deviation obtained from multiple determinations of the LD50 observed in this study can be attributed, at least in part, to strict standardization of the procedures used to determine the LD50. Zbinden (1963) has shown that the variability of the LD50 within an individual laboratory can be made relatively small by instituting strict controls and standardization of the assay procedures. However, as indicated above, the observed reliability of the LD50 test for botulinum toxin is also indicative of the fact that there are fewer factors contributing to lethality and there is little variation in the sensitivity of individuals within a population. Thus, death may simply be an inherently less variable end point for botulinum toxin.

The data presented in this report do not address the problems associated with interlaboratory variation in the estimates of the LD50. This is generally perceived as a very significant problem for LD50 values used to estimate toxicity. Zbinden and Fluri-Reversi (1981) reported greater than 10-fold differences in the lowest and highest LD50 values estimated by different laboratories. This same issue was addressed specifically with regard to estimation of the LD50 for botulinum toxin by Schantz and Kautter (1977) who reported the results of a study involving 11 laboratories. The coefficient of variation of 11 estimates of the toxin content of a reference standard was ±27%. These authors argued that the majority of this variation was probably due to the differences in the sensitivities of the mice used by the different laboratories. Trevan, in 1927, recognized the importance of the sensitivities of experimental animals to the reproducibility of the estimated LD50. Finney (1978) has indicated that the variance in the estimated LD50 can be affected by both experimental design and the homogeneity of the experimental animal population.

Schantz and Kautter (1977), have encouraged the use of a reference standard preparation of botulinum toxin. This is important for normalizing the results from different laboratories. Schantz and Johnson (1990) have further argued that the standardized assay (Schantz and Kautter, 1977) should be adopted for the characterization of preparations of botulinum toxin to be used clinically. Taking this approach, the estimated LD50 values can be expressed as relative potencies, potency ratios, using an agreed-upon reference sample of botulinum toxin. The suggestion by Schantz and Johnson (1990) was in response to a report by Quinn and Hallet (1989) calling for standardization of the “biological potency” reported by the two suppliers of botulinum A toxin for clinical use.

It is possible that the large differences in the clinical potency of the two preparations of botulinum A toxin might be attributed, at least in part, to interlaboratory differences in the mouse LD50 assay. The results of direct comparison of Botox and Dysport revealed that these two preparations differ by a factor of 1.9 (1.27/0.66). These results punctuate the importance of the suggestion of Schantz and Johnson (1990) to utilize a standard reference preparation of toxin to normalize estimates of the number of units (LD50) of toxin in preparations designated for clinical use. This difference (1.9-fold) in the actual number of units of botulinum toxin activity does not completely account for the larger differences in the recommended clinical doses of these two preparations. In contrast, the results of the regional hind limb chemodenervation paralysis model revealed differences (2.67) in the potency of Dysport and Botox. Combining the differences in the two preparations gives an overall 5.1-fold difference in potency which is consistent with the
large differences observed clinically for the marketed preparations of botulinum A toxin.

The reason for the differences in the behavior of the two A toxins in the lethality and hind limb assays is not immediately obvious. However, these differences may stem from one or more of the fundamental differences in the A toxin preparations. The two commercial type A toxins were obtained from different strains of Clostridium botulinum and the methods for purification of the two A toxins were different. No doubt these factors contribute to differences in the final composition of the two preparations. Another possibility is that the route of administration is critical. It is not unreasonable to assume that the same acceptors (receptors) and the same mechanisms of inhibition of ACh release are involved in mediating the actions of the toxins when injected ip and im. What could be different is the stability of the toxins when injected into muscle that is not manifest following ip administration. This is suggested from preliminary results (data not shown) indicating that the LD50 was increased following im administration as compared to ip injection. Although the present data do not resolve these questions, comparison of the regional chemodenervating effects of these two A toxin preparations clearly indicates that the different strains or the methods for purification and perhaps formulation have resulted in the different potencies of these preparations. Elucidation of the molecular basis for the differences in the regional chemodenervation produced by these toxins may provide the ability to engineer or select for those attributes of toxin composition or formulation that optimize clinical potency and efficacy.

The inability of the LD50 test to predict the marked differences in the clinical potency of preparations of botulinum toxin indicates that death as an end point is not sensitive to these differences. The paralysis model, in contrast, provides clear evidence of the marked difference in the potency of two preparations when doses were administered that were equivalent according to the mouse LD50 assay. Thus, under conditions where the two preparations of botulinum A toxin have been normalized according to the LD50 assay conducted in our laboratory, a large difference in potency of the denervating activity was still observed. The authors propose that characterization of the biologic activity of botulinum toxin by the LD50 assay alone is no longer warranted. At this early stage in the development of the paralysis model assay it is premature to suggest that this method might replace the LD50 estimate of biologic activity. However, at the very least, an additional parameter derived from the paralysis assay or equivalent should be specified along with the LD50 to accurately define the biologic activity of botulinum toxin.

The data presented in this report suggest that estimation of the LD50 can be a very precise method for defining botulinum toxin activity. The observed precision is probably related to two factors: First, strict control of the assay conditions and, second, the intrinsic specificity and selectivity of action of botulinum toxin associated with death as an end point. This does not, however, support the position that death as an end point is an optimum measure of the biologic activity of the botulinum toxin most important to both clinical efficacy and potency. The results of the comparison of the two approved preparations using the LD50 assay and the hind limb paralysis model suggest that estimation of the mouse LD50 is not an accurate nor adequate measure of the biological potency of clinical preparations of botulinum toxin. The paralysis assay is under further investigation for use in the assessment of the biologic activity of botulinum toxins.

ACKNOWLEDGMENTS

This work was supported by a grant from Associated Synapse Biologics, Inc. to L.B.P. L.B.P. is a Sloan Foundation Fellow.

REFERENCES

Mayorga, L. S., De Centorbi, O. P., De Milan, C. C., and Gimenez, D. F.


