Alternatives to the LD_{50} assay for botulinum toxin potency testing: Strategies and progress towards refinement, reduction and replacement

Dorothea Sesardic and Rose Gaines Das

National Institute for Biological Standards and Control

Corresponding author: Dorothea Sesardic Ph.D
Principal Scientist, Division of Bacteriology, NIBSC
Blanche Lane, South Mimms, Potters Bar, Hertfordshire, EN6 3QG, UK
Phone: +(44)-1707-641447, Fax: +(44)-1707-641054, dsesardic@nibsc.ac.uk or tsesardic@nibsc.ac.uk

Abstract
Therapeutic preparations of botulinum toxin products are calibrated by a mouse LD_{50} assay, and the labelled unit of activity is expressed in mouse LD_{50} units, defined as the median lethal intraperitoneal dose at a defined time-point. Potency and stability testing are required at several stages of the production process. Testing at the final lot stage requires high precision, so a large number of animals are used. LD_{50} assays are severe, and are at the front line for replacement on humane grounds. The Humane Society of the United States submitted a Test Method Nomination for "Alternative Methods to Replace the Mouse LD_{50} Assay for Botulinum Toxin Potency Testing" to ICCVAM, and a workshop organized by ICCVAM/NICEATM/ECVAM was held in Washington DC, USA, in November 2006. Several alternative methods, developed and validated at NIBSC over the past 15 years, were reviewed at that workshop. The workshop participants concluded that the validation of a refined in vivo local paralysis assay could provide a realistic short-term strategy for replacing the LD_{50} assay. Transferability studies with a UK testing laboratory are in progress. It was further anticipated that one or more of the in vitro methods, such as an enzyme cleavage assay, could be validated for particular applications, and then used to confirm consistency of production. The approach adopted at NIBSC has led to the elimination of the LD_{50} assay during independent confirmatory testing of the final lot product, and refined and alternative methods are included in the European Pharmacopoeia monograph, for use, subject to validation.

Keywords: endopeptidase, flaccid paralysis, LD_{50}, therapeutic botulinum toxin

Current licensed products and expression of potency
Therapeutic applications of botulinum toxin have increased steadily over the last 20 years, following the FDA's approval, in 1989, of Oculinum (now Botox, Allergan Inc.) for the treatment of strabismus and blepharospasm. Following the MHRA's approval of Dysport (Ipsen Biopharm Ltd) in the UK in 1995, a new version of Botox, incorporating purer bulk toxin, was licensed in 1997. Further products specifically targeting cosmetic applications were launched more recently, such as Vistabel for the treatment of glabellar lines. Therapeutic applications and the conditions for which these bacterial neurotoxins are used continue to increase, and licensed products now include pure type A neurotoxin, Xeomin (Merz Pharma, Germany), and products similar to Botox, such as Prosigine (Lanzhou Institute for Biological Products, China) and Meditoxin (Medy-Tox Inc., Korea).

In common with other biological products used in human medicine, the biological activity of the active component must be determined for each new batch and its activity followed by using suitable assay systems during the production process. Typically for these products, information on activity and stability is required for the concentrated bulk active toxin, and again at the final lot stage of the production process. A high-precision assay is required, particularly at the final lot stage, to confirm the amount of active substance in the product before it enters the market. Due to its high sensitivity, the mouse LD_{50} assay was adopted by all manufacturers, as a way of expressing product potency per vial. However, a standardised approach was not adopted, so the units used in labelling are product-specific and non-interchangeable (Sesardic, et al., 1994, McLellan, et al., 1996). This poses problems in product use, and has potential safety consequences as new products enter the global...
Although the mouse LD$_{50}$ assay is the primary method used prior to marketing approval by national and international regulatory authorities, a number of alternative methods have been developed, and approaches have been adopted which can reduce and refine the distress caused to the test animals, and can even replace their use. The European Pharmacopoeia (Ph Eur) monograph on Botulinum toxin type A (BoNT/A) for injection (01/2005:2113) supports the use of alternative methods, subject to their validation.

In response to a nomination from the Humane Society of the United States, a workshop, organised by the Interagency Coordinating Committees on the Validation of Alternative Methods (ICCVAM), the National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM), and the European Centre for Validation of Alternative Methods (ECVAM), was held on 13 –14 November 2006, in Silver Spring, MD, USA. The participants included scientists from leading governmental and academic institutions, national and global regulatory authorities, private industry, and the animal welfare community.

The objective of the meeting was to review the current state of alternative methods, and included those developed and used at NIBSC over the last 15 years for the independent confirmation of activity in some of the licensed products.

All the presentations made at the workshop are available on the NICEATM-ICCVAM website, at: http://iccvam.niehs.nih.gov/methods/biologics/botdocs/biolowkshp/wkshp_pres.htm.

Requirements and challenges for alternative potency assays

The ICCVAM workshop concluded that the methods most suitable for replacing the LD$_{50}$ assay for the potency testing of therapeutic products would need to fulfil specific requirements, and that approaches suitable for replacing assays for the detection of toxin in the environment may not be suitable for the potency testing of products. The replacement assay for potency testing must be able to accurately quantify nanogram levels of active toxin in the finished product, which contains a high concentration of bulking or stabilising proteins. The method must provide high sensitivity, with a limit of detection close to that of the mouse LD$_{50}$ assay. In addition to providing high precision and reproducibility, it must be easy to use, and sufficiently robust and transferable for quality control in the laboratory setting. Finally, it was emphasised that any replacement strategy must take into account the knowledge of the mode of action of the toxin, as well as an understanding of the production process of the therapeutic products. Any methods considered for validation should therefore be evaluated for their relevance and suitability for their intended purpose and use.

Reduction and refinement alternatives

The workshop provided an overview of alternative methods and approaches that, if sufficiently validated, accepted and applied, could reduce the number of animals used in assays or eliminate the pain and distress associated with the current method(s). Four different approaches were considered to be relevant:

- The use of reference standards and the expression of relative potency
- The use of earlier non-lethal humane endpoints for the current in vivo botulinum assay
- The use of alternate in vivo models to measure botulinum activity without lethality
- The use of ex vivo test models prepared from humanely-euthanized animals

Absolute measurements, such as LD$_{50}$ data, are not biological constants, and they are therefore highly dependent on the assay conditions. In line with many other biological methods, the use of relative potency estimates and the inclusion of a reference standard in the assay, lead to more precise and more reproducible estimates, which provide opportunities for reductions in animal use. An international collaborative study involving 10 laboratories, confirmed the role of a common standard in improving agreement among the data from potency assays of therapeutic toxin preparations (Sesardic, et al., 2003). In addition to reducing animal use in the conventional LD$_{50}$ assay, relative potency estimates can provide greater independence from assay conditions, and can provide opportunities for the development and use of alternative methods.

For the potency assay of BoNT/A for injection, the Ph Eur monograph (01/2005: 2113) now requires that a suitable reference preparation be tested in parallel. If a calibrated reference standard is incorporated into a bioassay for potency, the unit of measurement becomes relative potency, and the value is expressed as a unit of activity defined in terms of the activity of the reference standard.

This approach is required for the expression of potency when mouse paralysis is used as an endpoint. The mouse flaccid paralysis assay (also known as the mouse abdominal ptosis assay; Sesardic, et al., 1996; Sesardic, et al., 2004) relates the activity of BoNT/A to the degree of abdominal bulging seen after the toxin is subcutaneously injected into the left inguinoocular region of a mouse. The magnitude of the paralysis is dose-dependent. This approach has been proposed as a refinement to the mouse LD$_{50}$ test, because it relies on a humane endpoint and mice given the highest toxin dose only show a loss
of about 4% of their body weight. As only a sub-lethal dose of BoNT/A is injected, the assay is more sensitive (the maximum dose is 0.2 mouse LD_{50} units, so the ED_{50} is 10-fold more sensitive than the LD_{50}). Furthermore, the paralysis endpoint evaluates localized muscle effects, rather than systemic toxicity making it more similar to the clinical use of the toxin than the lethal assay. The flaccid paralysis assay is also more rapid than a lethality test, yielding results in 24 to 48 hours, compared to 72 to 96 hours for a typical LD_{50} assay. The paralysis model measures the activity of all three functional domains of the toxin protein, and its results show excellent agreement with the LD_{50} values for different products (Sesardic, et al., 1996). The robustness of the test was confirmed when several independent observers scored animals at 24 and 48 hours on a five-point scale, according to the size of the local abdominal bulge. As no specialised equipment is required, the test is relatively easy to conduct, and has been validated at NIBSC, where it was found to be capable of providing comparable potencies by using only 20% of the animals used in comparable LD_{50} testing. The method has been included as an option in the Ph Eur monograph (01/2005:2113), and a transferability exercise with a UK testing laboratory was initiated in 2006.

Ex vivo nerve/muscle preparations can also provide refined and humane alternatives, since mice or rats are only used for the donation of tissue, although all the functions of the toxin are required for producing an effect (Goschel, et al., 1997). In this approach, the amplitude of a twitch response to electrical stimulation of the nerve is measured. The basis for determination of the potency of the toxin is the decrease in the amplitude of the twitch response after it is applied to the maintenance medium. The usual endpoint of the assay is the time until a 50% decrease in amplitude is observed. The use of ex vivo models for potency assays is rapid, as it can provide results within two hours. Also, the experimental conditions can easily be varied. However, these tests require laboratory personnel trained in the use of sophisticated and expensive equipment. Additional validation studies are needed, to fully characterize the usefulness and limitations of various ex vivo models. At present, these models remain useful only as research tools, and their acceptance for batch release testing may not be straightforward. Furthermore, the direct application of the toxin to tissue preparations in solution is far removed from intramuscular injection in vivo, which results in poor correlations with the data from the LD_{50} assay or in vivo paralysis models, particularly when products with different stabilisers are compared.

Replacement alternatives

An attractive potential replacement method for the potency testing of botulinum toxin would rely on primary neuronal cells or cell lines of a neuronal lineage. However, several important challenges exist when cell-based potency measurements on formulated products are attempted. For example, the concentration of BoNT in the final formulation is in the picomolar range, and losses due to surface adsorption can markedly affect the results, as can formulation excipients, which can interfere with the analytical methods employed. Primary cells of rodent origin have been successfully used as research tools to confirm the actions of the toxin, but remain of limited applicability for routine use as a product batch release assay. In a controlled laboratory setting, a stably-transfected cell line could possibly minimize variability. However, extensive cloning or cell modification may be required, which would make it more difficult to create a stable cell line. Various endpoints have been explored in the search for cell-based assays. Loss of intact substrate or cleavage product within the cell, confirmed by immunoassays or the measurement of exocytosis, such as inhibition of glycine release, have been used, as these reflect the activities of all three functional domains of the toxin. However, in such studies, radioactive or fluorescent tracers and the equipment associated with their measurement with tissue cultures are required. Our own experience with the differentiated SH-SY5Y and M17 neuronal cell lines, confirms that the sensitivity of the assay is directly proportional to cellular uptake, and that none of the measurements provide sensitivity equivalent to that of the mouse bioassay, for therapeutic products.

Endopeptidase assays for BoNT toxins offer several advantages, because they reflect one of the key modes of action of the toxin, i.e. an intracellular enzyme activity of the toxin L-chain. The test requires, as a substrate, a synthetic or recombinant peptide that contains the endopeptidase-binding site, usually greater than 30 amino-acids in length. BoNT endopeptidase activity is highly toxin serotype-specific, and cleaves its target polypeptide between two different amino-acid residues in each case. Furthermore, these assays can provide sensitivity comparable to that of the mouse bioassay, and offer a wide variety of modifications and detection systems. Examples of the detection systems include the detection of exposed epitope post-cleavage with BoNT, with targeted antibody detection (Hallis, et al., 1996; Ekong, et al., 1997a, Jones, et al., 2007); capillary electrophoresis (Ekong, et al., 1997b; Sesardic et al., 1997); mass spectrometry for detecting size change by MALDI-TOF-MS on HPLC-ESI/MS (Barr, et al., 2005; Kalb, et al., 2006); fluorescent polarisation (Gilmore, et al., 2005) and plasmon resonance on Biocore (Ferracci, et al., 2005). An ELISA plate format with peptide substrate...
and targeted antibodies to detect the exposed epitope, was selected and validated at NIBSC (Ekong, et al., 1997a, Gaines-Das, et al., 1999), to verify the manufacturer's claims for product potency and as a consistency test for clinical samples (Sesardic et al., 1999, Jones, et al., 2007). Rapid and quantitative dose–response curves and correlation with LD₅₀ data were established for different products. It was confirmed that the assay provided adequate precision and reproducibility, and, now in use at NIBSC, it has provided potency estimates which are comparable with those obtained by the manufacturers using the LD₅₀ assay, with the result that the use of animals was greatly reduced. Endoptidase assays only measure toxin L-chain activity, and changes to the Hc domain of the toxin are not detected. However, these assays are more likely to overestimate potency when compared to LD₅₀ values and, following their validation, could be employed, in specific circumstances or in a tiered-testing strategy, to reduce the use of mice in current protocols. The need to maintain a supply of critical reagents is a potential disadvantage, and validation studies will need to be performed, to address the question of robustness and transferability. The method has been included as an option in Ph Eur monograph (01/2005:2113) for use at the final lot stage of product testing, again subject to validation.

Three R strategies for the potency testing of botulinum toxin products

Reduction:

Reduction alternatives are possible via the expression of potency relative to a reference standard, rather than absolute LD₅₀ values so that fewer animals provide results of equivalent precision. A reduction in animal use could also be achieved by extending the shelf-lives of reference standards and optimizing assay design and analysis.

Refinement:

Refinement alternatives are possible via reliance on earlier non-lethal endpoints or via the use of a non-systemic endpoint, such as paralysis, or ex vivo models. All these approaches require fully functional toxin and the expression of potency relative to a reference standard. Refinement methods could be validated for use when fully functional assays are essential, such as for the re-calibration of reference standards for defining activity in new bulk active toxin preparations, and for the validation of replacement methods.

Replacement:

Replacement alternatives presently available, such as in vitro enzyme activity models, are most suited for use as indicators of consistency, and could save animals if applied for final lot testing in the production process.

In Conclusion, refined and alternative methods suitable for the potency testing of BoNT toxins in therapeutic products are now available. Their validation, acceptance and adoption for use will depend on the cooperation of all the stakeholders. The 6th World Congress has provided opportunities for international validation organisations, manufacturers and regulators to discuss the next steps in the validation process.

Acknowledgements

We are grateful to The Humane Society of the United States, which provided financial support for attendance of DS at the Sixth World Congress on Alternatives & Animal Use in the Life Sciences. We are also grateful to ICCVAM/NICEATM/ECVAM and the EDQM, for the interest in the alternative methods developed at NIBSC.

References


