

Test Method Nomination to the ICCVAM

Nomination of Alternative Methods to Replace the Mouse
LD₅₀ Assay for Botulinum Toxin Potency Testing

Nomination Sponsor:
The Humane Society of the United States
2100 L Street, NW
Washington, DC 20037

Contact Person:
Martin L. Stephens, Ph.D.
MStephens@hsus.org
301-258-3040

October 31, 2005

Table of Contents

- Preface
- I. Introduction
 - A. Background
 - B. Priority considerations for ICCVAM review of alternatives to the mouse LD₅₀ assay
- II. Scientific and regulatory rationale for the proposed test method(s)
 - A. Scientific rationale (basis) for the proposed test method(s)
 - B. Regulatory rationale (applicability) for the proposed test method(s)
- III. Proposed test method protocol(s)
 - A. Overview of alternative methods for the potency assessment of botulinum toxin products
 - B. Mouse hind-limb assay
 - C. Abdominal ptosis assay
 - D. Mouse phrenic nerve-hemidiaphragm system
 - E. SNAP-25/endopeptidase assay
 - F. Endopep-MS assay
 - G. In vitro cell-based assays
- IV. Test substances
- IV. Reference method data and protocol(s)
 - A. Reference method data
 - B. Reference method protocol(s)
- VI. Test method data and results
 - A. Test method data
 - B. Test method results
- VII. Accuracy of the proposed test method
- VIII. Reliability (repeatability/reproducibility) of the proposed test method
- IX. Test method data quality
- X. Other scientific reports and reviews
- XI. Animal welfare considerations (refinement, reduction, and replacement)
 - A. History of FRAME and HSUS activities
 - B. Animal welfare considerations specific to the proposed test methods
- XII. Practical considerations (e.g., test method study costs, time needed to perform a study, ease of transferability of the test method among laboratories)
- XIII. References

APPENDICES

Appendix A: Methods and/or protocols for proposed test method(s)

Appendix B: Methods and/or protocol(s) for generation of reference method data

Appendix C: Copies of other relevant manuscripts and documents

Abbreviations

BoNT, *Clostridium botulinum* neurotoxin

BoNT/A, BoNT serotype A

BoNT/B, BoNT serotype B

CAMR, Centre for Applied Microbiology & Research

CDC, Centers for Disease Control and Prevention

DOD, Department of Defense

ECVAM, European Centre for the Validation of Alternative Methods

EDQM, European Directorate for the Quality of Medicines

EP, European Pharmacopeia

EPA, Environmental Protection Agency

EU, European Union

FDA, US Food and Drug Administration

FRAME, Fund for the Replacement of Animals in Medical Experiments

HSUS, Humane Society of the United States

ICCVAM, Interagency Coordinating Committee on the Validation of Alternative Methods

kDa, kilodalton (a measure of molecular mass)

LD₅₀, lethal dose 50%; the dose of a material that causes death in 50% of the test subjects

MLA, mouse lethality assay

MPA, mouse protection assay

MS, mass spectroscopy

NIBSC, National Institute for Biological Standards and Control

NICEATM, NTP Interagency Center for the Evaluation of Alternative Toxicological Methods

NIH, National Institutes of Health

NTP, National Toxicology Program

OECD, Organization for Economic Cooperation and Development

SNARE, N-ethyl maleimide-sensitive factor, attachment receptors (proteins essential for regulated exocytosis)

LC, light chain

SNAP-25, synaptosome-associated protein of 25 kDa

USDA, United States Department of Agriculture

VAMP, vesicle associated membrane protein

Preface

The Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) document, *Guidelines for the Nomination and Submission of New, Revised, and Alternative Test Methods*, defines a Test Method Nomination as “a test method proposed to ICCVAM for review and evaluation for which a complete test method submission is not available” (NTP, 2003). One of the examples of a Test Method Nomination provided in the ICCVAM guidelines is “test methods that are recommended for a workshop or other activities.”

The Humane Society of the United States (HSUS) is hereby submitting a Test Method Nomination for “Alternative Methods to Replace the Mouse LD₅₀ Assay for Botulinum Toxin Potency Testing.” In this nomination document, we discuss a number of potential refinement and replacement methods to the mouse LD₅₀ assay for botulinum toxin potency, and we recommend that these methods be assessed and prioritized for prevalidation and validation studies. We believe that an expert workshop would be an important initial step in this effort.

The ICCVAM nomination guidelines state that “priority is given to test methods that may provide for improved prediction of adverse human, animal, or ecological effects, and those that might reduce, refine, or replace animal use” (NTP, 2003). The unintended variability in potency of botulinum toxin-based products stemming from the use of the mouse bioassay would be reduced with the use of a more analytical type of assay. The alternative methods would also provide results more quickly, and could be used for additional applications where the mouse LD₅₀ is still the standard method for measuring botulinum toxin activity or antibodies, such as in diagnosis, vaccine efficacy, detection of neutralizing antibodies, and environmental sampling (food and water) related to bioterrorism threats. These potential applications of the proposed alternative methods are likely to provide for the improved prediction of adverse human effects.

The various alternative test methods presented in this nomination clearly would, if successfully validated, reduce, refine, and replace animal use for an *in vivo* test that causes a prolonged and distressful death for many mice. The proposed alternatives to the mouse LD₅₀ test also effectively address other ICCVAM prioritization criteria such as being applicable to regulatory testing needs and to multiple agencies/programs. Additional discussion on how the proposed alternative methods address the ICCVAM prioritization criteria is provided in the Introduction section of this document.

This Test Method Nomination provides a review of existing *in vivo*, *ex vivo*, and *in vitro* methods that are in various stages of development and validation as potential alternatives to the mouse LD₅₀ assay for botulinum toxin potency testing. The HSUS is submitting this Test Method Nomination as a recommendation that the existing alternative methods be expeditiously reviewed in the format of an expert workshop organized by the ICCVAM. It is expected that ICCVAM would work in coordination with the European Centre for the Validation of Alternative Methods (ECVAM) and other appropriate stakeholders to secure the cooperation of individuals and organizations that possess relevant protocols and test data, and that following the workshop, would work with appropriate partners to validate one or more of the alternative assays. The primary goal should be replacement of the mouse LD₅₀ test with one or more non-animal methods.

I. Introduction

A. Background

Botulinum neurotoxins (BoNTs), produced by strains of the bacterium *Clostridium botulinum*, are the most deadly natural toxins known, lethal at doses as low as 0.05 µg (Arnon, 2001; Lalli, 2003). There are seven antigenically distinct types of *Clostridium* neurotoxins, types A-G, which have a similar structure and mechanism of action (Centers for Disease Control and Prevention (CDC), 1998). BoNT types A, B, E, and in rare cases, F cause disease in humans, and types C and D cause disease in animals (CDC, 1998).

Only BoNT type A and type B (BoNT/A and BoNT/B) are approved for human clinical use. In the early 1970s, the therapeutic value of the paralytic potential of BoNT for treating problems involving the extraocular muscles was first recognized (Balls and Stephens, submitted). Since that time their use has rapidly expanded, and BoNTs are now used in over 50 therapeutic and cosmetic applications (Jost and Kohl, 2001; Jankovic, 2004). The markets in which the products are approved are continually increasing, for example, Dysport® (a European BoNT/A product) is now licensed in over 60 countries for treating a variety of neuromuscular disorders (Ispen, 2004).

In 1989, the US Food and Drug Administration (FDA) approved Botox® (Allergan, Inc., Irvine, CA) for the treatment of two eye muscle disorders - misaligned eyes (strabismus) and involuntary and uncontrollable blinking (blepharospasm) (FDA, 2002). In 2000, the FDA approved Botox and Myobloc® (BoNT type B; Elan Pharmaceuticals Inc., Morristown, NJ) for cervical dystonia (a neurological movement disorder that causes head and neck contractions) (Jankovic, 2004). Botox® Cosmetic (Allergan) was approved in 2002 for treating the frown (glabellar) lines between the eyebrows, a cosmetic application (FDA, 2002). Botox and Botox Cosmetic are commonly used for other cosmetic and therapeutic applications that have not been reviewed or approved by the FDA ("off-label" use) (FDA, 2002). According to the American Society for Aesthetic Plastic Surgery (ASAPS), Botox injections are the fastest-growing cosmetic procedure in the industry, (FDA, 2002). By 2004, Botox Cosmetic accounted for 40% of the net sales of Allergan's BoNT production (\$295 M, representing about 2.8 million treatments) (Allergan, 2005; ASAPS, 2005).

The potency testing of BoNT products for therapeutic or cosmetic purposes is based on an *in vivo* LD₅₀ assay that assesses the amount of toxin required to kill 50% of a group of mice that have been injected intraperitoneally with the preparation (Bottrill, 2003). The LD₅₀ assay used to test each batch of BoNT uses at least 100 mice, and "results in diffuse muscular paralysis, impaired vision, and paralysis of the diaphragm, followed by suffocation and death" (Balls and Stephens, submitted). A variation of the LD₅₀ mouse assay is also used to assess neutralizing antibodies to BoNTs, which develop in response to vaccines and can develop in a small percentage of patients receiving BoNT injections (Hall, et al., 2004). When used to assess neutralizing antibodies the assay is typically called the mouse lethality assay (MLA). In addition to the ethical concerns, there are other problems with the mouse LD₅₀ assay. McLellan, et al. (1996) noted that the bioassay is inherently variable and has a large interlaboratory variability. The mouse LD₅₀ assay has been found inadequate for assessing large differences in the potency of different manufacturer's products, suggesting that it is inadequate at detecting some relevant factor (Pearce, et al., 1994). The mouse assay also takes 1-4 days to provide results, delaying production, diagnosis, and detection applications that rely on this assay.

Therefore, there is growing urgency to replace the mouse LD₅₀ assay for both the ethical concerns over the animal use and the scientific concerns over the continued use of a variable animal bioassay when more sensitive alternative methods have already been developed.

The ever expanding demand and production of BoNT products places greater urgency on the need to develop and validate acceptable alternative assays to the mouse LD₅₀ assay for the potency and quality control tests used to produce and market these products.

A number of alternative *in vivo*, *ex vivo*, and *in vitro* methods for BoNT potency testing are in various stages of development and validation (Table 1). These test methods, if successfully validated, would reduce, refine, and replace animal use for an *in vivo* method that causes a prolonged and distressful death for scores of mice per assay. For the past few years, The HSUS, in the United States, and the Fund for the Replacement of Animals in Medical Experiments (FRAME), in Europe, have tried to work collaboratively with BoNT product manufacturers and government agencies to encourage them to validate and replace the mouse LD₅₀ assay with more humane test methods. These advocacy efforts have had limited success. Therefore, The HSUS is now nominating this issue to the ICCVAM.

B. Priority considerations for ICCVAM review of alternatives to the mouse LD₅₀ assay

The ICCVAM nomination guidelines define seven ICCVAM prioritization criteria which the National Toxicology Program (NTP) Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM) uses to evaluate the priority of test method submissions and nominations (NTP, 2003). The guidelines specifically note that “priority is given to test methods that may provide for improved prediction of adverse human, animal, or ecological effects, and those that might reduce, refine, or replace animal use” (NTP, 2003). Each of these prioritization criteria will be briefly discussed in this section.

1. Improved prediction of adverse effects

The alternatives to the mouse LD₅₀ assay have the potential to provide improved prediction of adverse health or environmental effects, compared to current test methods accepted by regulatory agencies.

Although the human clinical use of botulinum toxins has been found to be relatively safe (Naumann, 2004), variability in the assessment of the potency of BoNT products, both within the same product type and among the different BoNT products, has been cited as one problem stemming from the use of the mouse bioassay (McLellan, et al., 1996; Pearce, et al. 1997; Pearce, et al., 1995a; Jankovic, 2004; and others). The FDA has concluded that while “there is no chance of contracting botulism from Botox injections, there are some risks associated with the procedure. If too much toxin is injected, for example, or if it is injected into the wrong facial area, a person can end up with droopy eyelid muscles (ptosis) that could last for weeks” (FDA, 2002).

McLellan, et al. (1996) observed a large interlaboratory variability with the mouse bioassay in the absence of any commonly used reference preparation, “found that different assay conditions can affect potency estimates of clinical formulations of type A botulinum toxin,” and that different preparations are differentially affected by some of these assay conditions. They concluded “that these differences might well contribute to the differences observed in their clinical use” (McLellan, et al. (1996). Pearce, et al. (1995a) reported that “a discrepancy exists

between the clinical potency of equivalent international units of different formulations of botulinum A toxin for multiple clinical indications” when tested using the mouse LD₅₀ assay to quantify the biologically active toxin. Jankovic (2004) noted that there are still many unresolved issues and concerns about BoNT, including the “lack of standardization of biological activity of the different preparations....”

Potency and product variability force a compensating reliance on the clinician’s judgment and experience in using the products, which creates a source for human error. A standardized, cell-based or analytical assay for potency assessment would be expected to have less intra- and inter-assay variability than a mouse bioassay, resulting in reduced product and potency variability, and thereby providing for the improved therapeutic uses of the BoNT products. A standardized, quantitative alternative method would also reduce the inter-product variability described as the lack of comparability of ‘mouse units’.

The same alternative methods when substituted for other applications of the mouse bioassay such as potency and safety testing for therapeutic antitoxins or environmental assessments could result in the improved prediction of adverse health and environmental effects. An assay that would more quickly provide results for botulinum toxicity in humans would clearly be useful. The CDC (1998) reports that botulism is confirmed by identifying BoNT in the serum, feces, vomitus, or gastric contents of the patient, and/or in food remnants. “The only currently acceptable method for detection and identification of botulinum neurotoxin is the mouse toxicity and neutralization bioassay” (CDC, 1998). “Laboratory diagnostic testing for botulism in the United States is available only at the CDC and approximately 20 state and municipal public health laboratories” (Arnon, 2001). “The administration of antitoxin is the only specific therapy available for botulism, and evidence suggests that it is effective only if given very early in the course of neurologic dysfunction” (CDC, 1998). Thus, a validated alternative assay that provides results more quickly than the mouse assay would better protect human health.

Cell-based and analytical assays also take less time, in general, than the mouse bioassay, and would therefore better protect public health in the event of a bioterrorism incident involving BoNT. The Working Group on Civilian Biodefense recommended the development of rapid diagnostic and toxin typing techniques for recognizing and responding to a bioterrorist attack, and concluded that “assays that exploit the enzymatic activity of botulinum toxin have the potential to supplant the mouse bioassay as the standard for diagnosis” (Arnon, 2001).

2. Potential to refine, reduce, and/or replace animal use

The alternatives to the mouse LD₅₀ assay also *have the potential, compared to current test methods accepted by regulatory agencies, to refine, reduce, and/or replace animal use*. The mouse LD₅₀ assay is reported to use at least 100 mice per test, and a given batch of toxin may be tested up to three times prior to marketing. The mice are injected intraperitoneally with the toxin, and monitored for 3-4 days to determine the dose required to kill 50% of the test group (Bottrill 2003). Death is the endpoint, resulting from suffocation due to paralysis of the diaphragm musculature. Over the 3-4 day period there are additional distressful and debilitating symptoms. Given that these symptoms cause fear and panic in humans that are accidentally poisoned by BoNT, it can be argued that they cause similar feelings among the mice.

The market for botulinum products, both therapeutic and cosmetic, is expanding rapidly (Jankovic, 2004). Consequently, the potency testing for botulinum products represents an area of increasing animal use, and the lethal and distressful *in vivo* LD₅₀ assay is a prime candidate for

replacement with more humane alternatives. The availability of a number of promising alternative methods facilitates the effort proposed in this nomination to move one or more of these methods into the pipeline for further evaluation and validation.

The alternative test methods presented in this nomination, if successfully validated, would clearly reduce, refine, and replace the mouse LD₅₀ assay for potency testing of BoNT products. The *in vitro* cell, molecular, and analytical assays would provide the greatest reduction in animal use and suffering. The *in vivo* assays that cause local paralysis (Table 1) are obvious refinements over a test that causes systemic paralysis leading to death, although they still cause some distress.

3. *Other priority setting issues*

The alternatives to the mouse LD₅₀ assay *are applicable to regulatory testing needs*. Both Botox and Botox Cosmetic are regulated as drugs by the FDA (FDA, 2003a; FDA, 2003b; FDA, 2004). The submission of toxin potency data is required to support marketing of BoNT products intended for human use, and the accepted test method is the *in vivo* mouse LD₅₀ assay. Regulatory agencies that require the mouse potency assay include the FDA in the US, agencies in other countries where BoNT products are produced, and other countries where the products are marketed. The alternative *in vivo* local paralysis assays and neuronal cell-based assays involve or replicate the identical mechanism of action as the *in vivo* mouse LD₅₀ assay. Once validated in comparison to the mouse bioassay, these alternative methods should provide comparable and possibly more reliable results than the mouse assay for the evaluation of BoNT potency. The molecular assays also represent the same mechanism of action that is responsible for the *in vivo* effects of BoNT, the proteolysis and inhibition of neurotransmitter release. The molecular assays do not reproduce the receptor binding step, but this has not been shown to be essential for obtaining comparable results.

The alternatives to the mouse LD₅₀ assay *are applicable to multiple agencies and programs*. US government agencies that require, request, or use the mouse LD₅₀ assay for BoNT potency, diagnostic, or detection testing include the FDA, the CDC, the United States Department of Agriculture (USDA), the Department of Defense (DOD), and the Environmental Protection Agency (EPA) (FDA, 2003a; FDA, 2003b; FDA, 2004; CDC, 1998; EPA, 1997; EPA, 2002; USDA, 2005). Some other branches of the Department of Health and Human Services, such as the National Institutes of Health (NIH), may use the assay in research and safety testing applications. Other agencies that work with, transport, or monitor human pathogens that are potential bioterrorism agents, such as the Department of Homeland Security and the Department of Transportation (DOT) may also use the mouse LD₅₀ assay for research, safety, or diagnostic testing purposes, but this has not been confirmed. There is also considerable government interest in alternatives to the mouse LD₅₀ assay in the EU (Bottrill, 2003; Balls, 2003; Balls and Stephens, submitted).

The alternatives to the mouse LD₅₀ assay *are warranted, based on the extent of expected use or application and impact on human, animal, or ecological health*. Jankovic (2004), Sesardic (2002), and others have noted the expanding numbers of clinical and cosmetic applications for BoNTs. Financial reports of the manufacturers indicate the large and increasing demand for BoNT products (Allergan, 2005). The ever increasing production of and markets for BoNT products has resulted in increasing numbers of animals being used in the potency assay. As previously mentioned, there are also other applications for alternatives to replace the MLA and LD₅₀ assay (antitoxin and toxin vaccine production, environmental monitoring, bioterrorism

monitoring, etc.). The extent of use of the LD₅₀ potency assay, the large numbers of animals required per test, and the severity of the test on the animal subjects, collectively warrant that alternatives to the mouse LD₅₀ assay are a high priority for development and validation by government authorities.

The alternatives to the mouse LD₅₀ assay *provide other advantages compared to current methods*. Other advantages of the proposed methods are discussed in Section XII.

4. *The completeness of the nomination*

We recognize that some elements of this nomination packet are incomplete, but this represents the current status of the available methods, not a lack of diligence on our part.

II. Scientific and Regulatory Rationale for the Proposed Test Method(s)

A. *Scientific rationale (basis) for the proposed test method(s)*

1. *Mechanism of action of botulinum toxin*

In a recent report on the many clinical applications of BoNT, Jankovic (2004) noted that “few therapeutic agents have been better understood in terms of their mechanism of action before their clinical application.” He describes the therapeutic effect of BoNT as being derived from “its ability to inhibit the release of acetylcholine from the presynaptic nerve terminal, causing local chemodenervation” (Jankovic, 2004). The inhibition of acetylcholine release results from proteolysis of the proteins that are involved in the release (exocytosis) of this neurotransmitter (Segelke, 2004).

Segelke (2004) described the *in vivo* mechanism of action of BoNT proteins as follows: “BoNTs are synthesized *in vivo* as 150-kDa holotoxins that are subsequently cleaved to form disulfide-linked heterodimers containing an ~50-kDa N-terminal light chain (LC) and the ~100-kDa C-terminal heavy chain. The heavy chain consists of the C-terminal receptor-binding domain, which targets the toxin to presynaptic neurons, and the N-terminal translocation domain, which induces pore formation in acidic endosomes and translocates the LC into the cytosol. BoNT LCs are zinc proteases known to cleave proteins of the soluble *N*-ethylmaleimide-sensitive factor attachment protein receptor (SNARE) complex which are critical for neurotransmitter exocytosis.”

Different botulinum toxins cleave different components of SNARE: SNAP 25 (synaptosome-associated protein of 25 kDa) (cleaved by BoNT A and E, at different sites), VAMP (vesicle associated membrane protein)/synaptobrevin (cleaved by BoNT B, D, and F), or syntaxin and SNAP-25 (cleaved by BoNT C) (Jankovic, 2004; Segelke, 2004). Cleavage of any of these SNARE components prevents the fusion of acetylcholine synaptic vesicles with the neurons plasma membrane, blocking the release of the neurotransmitter (Jankovic, 2004).

The delay in recovery from the inhibition of neurotransmitter release is the feature of BoNTs that provides for their therapeutic utility. Kadkhodayan, et al. (2000) provided the following description for the mechanism of this effect: “Within a few days after injection of botulinum toxin A into skeletal muscle, the affected nerve terminals are no longer capable of neurotransmitter exocytosis, but newly formed sprouts release acetylcholine, forming a functional synapse. After about three months, consistent with the return of clinical function of the muscle and a wearing off response from the previous injection, the original terminal resumes exocytosis and the sprouts regress to return the neuromuscular junction to its original state. Using cerebellar

neurones, the half life of protease activity based on blockade of transmitter release was estimated to be 31 days for toxin A, 25 days for toxin C1, 10 days for toxin B, two days for toxin F, and 0.8 days for toxin F.”

Tang-Liu, et al. (2003) injected radioiodinated BoNT/A and BoNT/B into rat gastrocnemius muscle and eyelids of rabbits, and found that most of the neurotoxin did not diffuse from the injection site, and observed no detectable systemic effects or generalized toxicity.

2. *Mechanism of action of BoNT in the mouse LD₅₀ assay*

At the cellular and molecular levels, the mechanism of action of BoNT is the same in the proposed alternative methods as in the *in vivo* mouse LD₅₀ assay. However, the clinical relevance of systemic paralysis is not equivalent to the clinical effect for which BoNT products are used, which is actually better replicated by the *in vivo* alternative methods of local paralysis.

The use of the mouse LD₅₀ assay, a test for systemic toxicity, has had a poor track record as a method to evaluate potency for BoNT products (Pearce, et al., 1994; 1995a; 1995b; 1997). The problems with assay and product variability have been sufficiently discussed in other sections of this document. The variability is mentioned again here, because one of the contributing factors is related to whether the mechanism of systemic toxicity from BoNT is sufficiently different from the mechanism of the action of BoNT on an isolated muscle. “The international unit of biologic activity of botulinum toxin has been defined as the LD₅₀ in mice. This unit of biologic activity provides one measure of the toxic properties of botulinum toxins; however, it does not reflect the pharmacologic properties of botulinum toxins most important in clinical practice. Comparison of the clinical potency of the two approved preparations of botulinum toxin (Botox and Dysport) has raised the question of whether or not the lethality assay is the most accurate method for assessing the biologic activity of botulinum toxin” (Pearce, et al., 1995a). They reported that a regional chemodenervation assay, the mouse hind-limb assay, better approximated the clinical use of BoNT.

3. *Mechanism(s) of action of proposed test methods*

All of the BoNT assays nominated in this document for consideration as replacements for the mouse LD₅₀ assay, whether *in vivo*, *ex vivo*, or *in vitro*, are based on the *in vivo* mechanism of action for BoNT described in Section II.A.1. They are all considered to be functional assays that measure the biological activity of BoNTs rather than detecting the presence of the toxin proteins.

Scientists at the NIBSC (National Institute for Biological Standards and Control) claim that “advances made in understanding the mode of action of clostridial neurotoxins have provided the basis for the development of alternative mechanism-based assay methods” (Ekong, et al., 1997).

The alternative *in vivo* local paralysis assays replicate the mechanism of action of BoNT in an isolated muscle, and more closely resemble the human clinical application of BoNT than does the mouse LD₅₀ assay. The molecular assays, like the SNAP 25/endopeptidase assay, reproduce the same mechanism of action that is responsible for the *in vivo* effects of BoNT, the proteolysis and inhibition of neurotransmitter release. The molecular assays do not reproduce the receptor binding step, however, this has not been shown to be essential for obtaining comparable results. The neuronal cell-based methods, although the least developed into a useful assay format and sometimes lacking sufficient sensitivity, are promising methods for future development.

Williamson, et al. (1996) was “the first to demonstrate in physiologically relevant cells, *i.e.* in intact functioning neurons, a direct correlation between the clostridial neurotoxin-induced block

in neurotransmitter release and the cleavage of toxin-specific protein substrates, VAMP, SNAP-25, or syntaxin.... These findings confirm that the principal mechanism of action of clostridial neurotoxins is proteolytic cleavage of specific synaptic proteins necessary for neurotransmitter release.” Once validated in comparison to the mouse bioassay, these alternative methods should provide comparable and possibly more reliable results than the mouse assay for the evaluation of BoNT potency.

4. *Usefulness and limitations of proposed test methods*

Only functional assays, or ones that provide an assessment of botulinum toxin activity, have been proposed in this nomination. Immunoassays and other methods that detect the toxin proteins, but not their activity, reportedly are not useful in assessing the potency of the toxin products (Ekong, et al., 1995).

The advantages, in general, for the proposed alternative methods are that they will provide results faster, and will provide for reduced pain and suffering in animals. A major advantage is that potency results from the proposed *in vitro* alternative methods are expected to show less variability than the mouse bioassay, thereby providing for more consistent potency of BoNT products. The variability in the proposed alternative *in vivo* and *ex vivo* methods is not known, but would be predicted to be no greater than the mouse bioassay variability. Some of the alternative methods have already been shown to have good correlation with the *in vivo* assay.

The limitations, in general, are that the *in vivo* and *ex vivo* methods still rely on using animals. Most of the alternative methods would take longer to perform (meaning more laboratory personnel time to perform the assay, but actually a shorter time to obtain the end result), and would require more skill to conduct. For example, variation in the injection site might affect the degree of response observed in the alternative *in vivo* methods.

Specific advantages and limitations for some of the proposed methods have been discussed in greater detail by researchers in their publications, and some of the details are included in the discussion of the test methods in Section III.

B. *Regulatory rationale (applicability) for the proposed test method(s)*

Botox, Botox Cosmetic, and MyoBloc are regulated as drugs by the FDA (FDA, Sept 16, 2003; FDA, Sept 24, 2003; FDA, 2004; CenterWatch, 2004), and other BoNT products are similarly subjected to regulatory approval in the countries in which they are produced and marketed. To obtain regulatory approval for marketing a BoNT product, a manufacturer has to conduct and submit data from routine safety testing including toxicological tests conducted in different species. However, once the product is approved this testing does not need to be repeated. When marketing a BoNT product, however, each new batch of product has to be tested by the manufacturer for toxin potency using the mouse LD₅₀ assay, which is the only accepted assay at this time. Bottrill (2003) notes that “a further confirmatory potency test will be carried out by one or more control laboratories in the countries in which the batch will be used.”

The mouse LD₅₀ assay for BoNT potency testing is particularly severe in that it causes a slow death by asphyxiation. With the ever increasing number of BoNT products, markets, and applications, the number of animals used in the LD₅₀ assay will continue to increase. The validation and acceptance of an alternative method(s) to the mouse LD₅₀ bioassay would support the intention of statutes established to reduce animal use, such as the ICCVAM Authorization Act

(US Congress, 2000), and, with respect to Botox Cosmetic, the Seventh Amendment to the EU Cosmetics Directive (Europa, 2003). The use of the LD₅₀ test is being phased out worldwide, as symbolized by the deletion of the LD₅₀ assay for testing the systemic toxicity of chemicals from the Organization for Economic Cooperation and Development's (OECD) Health Effects Test Guidelines (OECD, 2002).

Currently, BoNT product producers use the mouse LD₅₀ assay to evaluate product potency, however, each company has its own test protocol (Bottrill, 2003). The potency of a "mouse unit" of BoNT varies among the products from different manufacturers, and it is an ongoing dilemma to clinicians to translate equivalent doses between products such as Botox and Dysport (Poewe, 2002; Jankovic, 2004). Bigalke, et al., (2001) claim that "disagreement on the dose equivalence of the two [Botox and Dysport] preparations implies a potential risk for the patient." Since a major purpose for regulating BoNT products is to protect the public health, then the validation and acceptance of a more quantitative alternative method to the use of the mouse LD₅₀ assay for BoNT potency assessment, which would standardize potency within and across different product types, should be a sufficient regulatory rationale.

As discussed in the Introduction, the alternatives to the mouse LD₅₀ assay are applicable to regulatory testing needs. The alternative *in vivo* local paralysis, molecular (SNAP-25), and neuronal cell-based assays replicate the mechanism of action of BoNT in the *in vivo* mouse LD₅₀ assay - the proteolysis and inhibition of neurotransmitter release. The early *in vitro* assays lacked sufficient sensitivity for determining the potency of BoNT in therapeutic preparations, however, the sensitivity of current *in vitro* assays now equals or exceeds that of the mouse bioassay (Ekong, et al., 1997). The NIBSC, which tests BoNT products manufactured in the UK, routinely uses the endopeptidase assay (Bottrill, 2003), and uses a non-lethal *in vivo* test only when the *in vitro* tests are inconclusive (Stephens and Balls, submitted). Ipsen, Ltd., the European manufacture of Dysport is reportedly working with the NIBSC to develop suitable alternative batch release tests (Stephens and Balls, submitted).

III. Proposed Test Method Protocol(s)

A. Overview of alternative methods for the potency assessment of botulinum toxin products

The known potential alternatives to the mouse LD₅₀ assay for BoNT potency testing, both *in vivo* and *ex vivo* refinement and *in vitro* replacement assays, are summarized in Table 1. Pearce, et al. (1997), Bottrill (2003), and Balls (2003) have published comparative reviews of alternative methods for measuring the biological activity of BoNTs. Copies of these review articles are provided in Appendix C. Several new alternative methods that have been developed are also discussed below.

In a recent monograph on BoNT type A (European Pharmacopeia (EP), 2005), the EP noted the potential of alternative methods to substitute for the mouse LD₅₀ test by stating that: "After validation with respect to the LD₅₀ assay (reference method), the product may also be assayed by other methods that are preferable in terms of animal welfare, including one of the following: endopeptidase assay *in vitro*; *ex vivo* assay using the mouse phrenic nerve diaphragm; or mouse bioassay using paralysis as the end-point." For these alternative methods, "the potency is calculated with respect to a suitable reference preparation calibrated in mouse LD₅₀ units" (EP, 2005). (Note: the EP Monograph 5.0 was published prior to publication of the *in vitro* endopeptidase assay.) Sesardic, a researcher at the NIBSC, claims that "scientific advances in the

understanding of the mode of action of clostridial neurotoxins have now provided the basis for improving conventional testing procedures for both therapeutic toxins and antitoxins” (Sesardic, 2002).

Table 1. Promising alternative methods to the mouse LD₅₀ test for assessing the potency of botulinum toxin products.

Name of Test	System	Endpoint	Duration	Reference
Mouse hind-limb assay	<i>in vivo</i>	local paralysis	2 days	Pearce, et al., 1994; 1995a; 1995b; Roger Aoki K., 2002; Aoki, 2001
Mouse abdominal ptosis assay	<i>in vivo</i>	local paralysis	< 1 day	Takahashi et al., 1990a; 1990b; Sesardic, et al., 1996
Mouse phrenic nerve-hemidiaphragm assay	<i>ex vivo</i>	muscle contraction	< 1 day	Dressler, et al., 2005; Bigalke et al., 2001; Bigalke, 2001; Kalandakanond and Coffield (2001; 2001);
SNAP-25/Endopeptidase assay	<i>in vitro</i>	molecular disruption of nerve transmission	< 1 day	Ekong et al., 1997; Hallis (1996); Schmidt, et al. (2001); Wictome et al., 1999; Ferracci, et al., 2005;
Endopep-MS	<i>in vitro</i>	molecular disruption of nerve transmission	< 1 day	Boyer et al., 2005; Barr, et al., 2005; Kalb, et al., 2005
Neural cell-based assays	<i>in vitro</i>	molecular disruption of nerve transmission	varies	Sesardic, et al., 2004; Williamson, et al., 1996; Welch, et al., 2000; Dong, et al., 2004; Hall, et al., 2004; Keller, et al., 2004; Chaddock, et al., 2004
Human foot muscles extensor digitorum brevis	<i>in vivo</i>	muscle contraction	1-7 days; up to 90 days	Bigalke, 2001; Bigalke, et al., 2001; Eleopra et al., 2002; Jost, et al., 2005

The mouse bioassay is currently used for a number of semi-quantitative assessments, including toxin potency in BoNT products, neutralizing antibodies, toxin type identification, vaccine development, toxin detection in food and water, and diagnostic evaluations. Some of the

alternative methods discussed in this section have been developed specifically for one of these applications, but could potentially be modified for use for the other applications.

Many of the alternatives in Table 1 are discussed in greater detail in the sections that follow. These alternative methods, although thought to be comprehensive, are not meant to exclude the possibility that an expert workgroup might identify additional alternatives that should be considered. Some of the methods that are discussed were developed as alternatives to the MLA for detecting BoNT neutralizing antibodies. The MLA uses many mice in a distressful test; many MLS alternative methods are in development and may, with modification, represent useful alternatives to the LD₅₀ BoNT potency assay.

B. Mouse hind-limb assay

The *in vivo* mouse hind limb assay assesses local paralysis produced by intramuscular injection of BoNT (Pearce et al., 1994, 1995a, 1995b). Various dilutions of the toxin were injected into the gastrocnemius muscle of the right hind leg of mice. The degree of paralysis was evaluated (scored) daily for 12 days. Pearce, et al., (1994) compared the results obtained with the hind-limb assay to those from the *in vivo* LD₅₀ assay for different dilutions of Botox and Dysport. The LD₅₀ assay failed to predict differences in the clinical potency of the two BoNT formulations, while the hind limb paralysis model detected the large difference in potency that is known to exist between these formulations.

Pearce, et al., (1995a) quantified the regional paralysis in the mouse hind limb produced by BoNT to “define a new pharmacologic/biologic unit of activity that more accurately reflects the mechanism of action of botulinum toxin in the clinical setting.” The median paralysis unit (MPU) or ED₅₀ unit from the hind-limb assay was used to define the potencies of Botox and Dysport and accurately reflected the results from retrospective clinical studies. The data suggested that the MPU may be a more appropriate measure of the biologic activity in therapeutic formulations of botulinum toxin” (Pearce, et al., 1995a). “Although experiments were not performed using less than ten animals per dose, the observed data suggest that statistically satisfactory estimates of biologic activity could be obtained with fewer than ten animals per dose.”

Aoki used the hind-limb assay for similar studies (Aoki, 2000; Roger Aoki, 2002). The muscle weakening efficacy, duration, and safety margin of botulinum toxin type B (Myobloc) to botulinum toxin type A (Botox) were evaluated using the mouse hind limb assay (RogerAoki, 2002). Mice received a single hind limb intramuscular injection of BoNT/A or BoNT/B and were scored (0-4 scale using the digit abduction scoring (DAS) assay) on the magnitude and duration of muscle weakening efficacy. Peak DAS scores were plotted and intramuscular ED₅₀ values were calculated. BoNT/A produced muscle weakness at lower doses and had a longer duration than BoNT/B. The mean dose that was lethal in 50% of mice was lower for BoNT/A. The *in vivo* differences found were consistent with the different clinical profiles for BoNT/A and BoNT/B (Roger Aoki, 2002).

C. Abdominal ptosis assay

The *in vivo* mouse abdominal ptosis assay also assesses local paralysis (Takahashi, et al., 1990a; 1990b; Sesardic, et al., 1996). Takahashi, et al. (1990a) observed a palsy of the abdominal muscle that develops after subcutaneous injections of low levels of BoNT into the mouse at the inguino-crural region. One mouse LD₅₀ unit could be detected at 6 h, and 0.1 LD₅₀ unit at 24 h.

The severities of symptoms were converted to scores, and statistical analysis of the scores established an assay method for titration of low levels of the toxin. A wide range of doses, 0.075 to 38.4 LD₅₀ units, gave a linear dose-response curve. The method required fewer animals and was more reproducible than the mouse lethality assay.

Sesardic, et al. (1996) enhanced the abdominal ptosis method of Takahashi, et al. (1990a, 1990b), and validated its use for testing the potency of BoNT preparations. They quantified the flaccid paralysis of muscles following subcutaneous injection of BoNT at the inguino-crural region. They noted that the abdominal ptosis method was simple to perform and to score (provided easily differentiated degrees of flaccid paralysis). Except for the positive controls, the highest dose of BoNT used was less than 0.25 LD₅₀ units, and no mice suffered systemic effects. Results showed that potency estimates in the refined assay had excellent agreement with estimates obtained using the currently required method [LD₅₀ assay], demonstrating that an alternative, more humane bioassay for potency testing “gives valid estimates equivalent to those currently in use” (Sesardic, et al., 1996). Bottrill (2003) noted that the NIBSC routinely uses this assay for batch release purposes.

D. Mouse phrenic nerve-hemidiaphragm assay

A variety of neuromuscular preparations, consisting of muscle tissue and attached nerves, have been used to evaluate BoNT activity (Pearce, et al., 1997; Bottrill, 2003). These animal tissue assays are based on measuring the force of muscle contraction following electrical stimulation of the tissue in the absence and presence of the BoNT. The mouse phrenic nerve hemidiaphragm assay has been identified as the most promising of the isolated tissue methods as an alternative to the mouse LD₅₀ assay (Bottrill, 2003).

The *ex vivo* mouse phrenic nerve hemidiaphragm assay uses an isolated tissue responsible for respiration, and according to Sesardic, et al. (2004) “is becoming established as a clear alternative to *in vivo* assays for research purposes, and may eventually become accepted for batch release testing.” Sesardic reports that “like *in vivo* assays, it [the hemidiaphragm model] requires the fully functional toxin and is the most sensitive functional model for detecting neutralizing antibodies (0.5 mU/ml).” Dressler, et al., (2005) also found that the hemidiaphragm model was useful for measuring neutralizing BoNT titers quantitatively and with adequate sensitivity and specificity.

Bigalke, et al. (2001) and Bigalke (2001) compared Dysport and Botox preparations using the mouse phrenic nerve hemidiaphragm assay and the results obtained in human volunteers using the foot muscles *extensor digitorum brevis*. In the method described by Bigalke, et al., (2001), the phrenic nerve-hemidiaphragm was excised from mice and placed in an organ bath containing Krebs-Ringer solution. The phrenic nerve was continuously electrostimulated and isometric contractions were recorded. The resting tension of the diaphragm was measured, and the incubation medium was then exchanged for the toxin-containing solution. Depending on the toxin concentration, the amplitude decreased slowly over time. The time required for a 50% decrease in the amplitude was used to construct dose-response curves. After a latency period there was progressive decline in the contraction amplitude down to zero. The time to paralysis, i.e., the elapsed time to attain half maximal twitch, depended strictly on the toxin concentration.

Kalandakanond and Coffield (2001a) reported that “SNAP-25, synaptobrevin II, and syntaxin I, the intracellular substrates of botulinum toxin originally identified in nontarget tissues, were present in a recognized mammalian target tissue, the mouse hemidiaphragm....[and] were

cleaved by incubation of the intact hemidiaphragm in botulinum serotypes A, C, and D, respectively. Additional studies (Kalandakanond and Coffield, 2001b) demonstrated “that cleavage of SNAP-25 by botulinum serotype A fulfills the requirements of the multistep model of botulinum toxin action that includes receptor-mediated endocytosis, pH-dependent translocation, and zinc-dependent proteolysis.”

E. SNAP-25/endopeptidase assay

BoNT type A is a zinc-dependent endopeptidase that cleaves proteins required for neurotransmitter release (Dong, et al., 2004). One of the protein substrates for the BoNT/A endopeptidase activity is SNAP-25. A potential *in vitro* replacement test method, the SNAP-25/endopeptidase assay, measures the action of BoNT/A on its target molecule, SNAP-25 (Ekong et al., 1997). The following assays, which are based on the molecular mechanism of action of the BoNT endopeptidase cleavage of its substrate molecule, are perhaps the most promising alternatives to the mouse LD₅₀ assay at this time.

Some of the early *in vitro* assays based on the cleavage of immobilized endopeptidase peptide substrates, while rapid, specific, and measuring biological activity of the neurotoxin, were not as sensitive as the mouse bioassay and therefore not realistic replacements (Wictome, et al., 1999). One of the first-generation endopeptidase assays for BoNT/A was developed by Hallis, et al. (1996) where a fragment of SNAP-25 was used as a solid-phase peptide substrate for BoNT/A. The assay system was found to be specific for BoNT/A, and no false-positive reactions were detected. The major advantage of this assay was that it measured the biological activity of the neurotoxin. The major problem was its being less sensitive than the mouse LD₅₀ assay. The major factor determining the sensitivity is the rate of cleavage of the synthetic peptide by the toxin, and many factors were shown to influence the rate of cleavage. However, the test method developers did note that “although not a complete substitute for the mouse test, this first-generation endopeptidase assay could be used during the manufacture of the clinical product to significantly reduce the level of testing in animals” (Hallis, et al., 1996).

The immunoassay for detecting BoNT/A activity toward recombinant SNAP-25 substrates that was developed by Ekong, et al. (1997) was shown to be more sensitive than the mouse assay, and to have good correlation with the *in vivo* results. “Using recombinant DNA methods, a segment of SNAP-25 spanning the toxin cleavage site was prepared as a fusion protein.... Targeted antibodies specific for the N and C termini of SNAP-25, as well as the toxin cleavage site, were prepared and used in an immunoassay to demonstrate BoNT/A endopeptidase activity towards recombinant SNAP-25 substrates.... The endopeptidase assay has proved to be more sensitive than the mouse bioassay for detection of toxin in therapeutic preparations. A good correlation with results obtained in the *in vivo* bioassay ($r = 0.95$, $n = 23$) was demonstrated.” “The assay has been shown to be particularly suitable for the detection of BoNT/A in clinical preparations where its use would eliminate the use of large numbers of animals currently required for testing purposes” (Ekong, et al., 1997). This endopeptidase assay is now in routine use at the NIBSC (Bottrill, 2003).

Schmidt, et al. (2001) developed high-throughput, solid-phase protease activity assays for BoNT serotypes A, B, D, and F to expedite toxin inhibitor discovery. Each assay consists of a fluorescein-labeled cleavable oligopeptide, based on the natural substrate sequence. The peptide is immobilized in multiwell plates where it is exposed to test compounds. The assays required relatively low concentrations of toxins (10–100 ng/ml) and short incubation times (1–3 h). The

fluorescence assays were selective, gave very low background readings, and were stable upon prolonged storage.

Wictome, et al. (1999) developed an endopeptidase assay for the detection of BoNT type B that is more sensitive than the mouse bioassay. “The assay is amplified by the enzymatic activity of the neurotoxin's light chain and includes the following three stages: first, a small, monoclonal antibody-based immunoaffinity column captures the toxin; second, a peptide substrate is cleaved by using the endopeptidase activity of the type B neurotoxin; and finally, a modified enzyme-linked immunoassay system detects the peptide cleavage products. The assay that was developed requires only a few specialized reagents, has greater sensitivity than that of the mouse bioassay, measures the biological activity of the neurotoxin, and is sufficiently robust to detect BoNT/B in a range of foodstuffs (Wictome, et al., 1999). This specific assay was developed for safety validation of foodstuffs, but some of its features should be useful in developing or refining a similar assay as a replacement to the mouse bioassay for BoNT potency.

Another assay developed for BoNT/B that could be adopted for BoNT/A was described by Ferracci, et al. (2005): “Western blotting and SPR (surface plasmon resonance) methods revealed that BoNT/B and F totally cleave their substrate (VAMP) on immunisolated SVs (synaptic vesicles). Real-time monitoring of the immunocapture of native SVs from crude lysates on SPR sensor chips enabled the detection of picogram amounts of different SV proteins. Pre-incubation of a membrane fraction containing SVs with BoNT specifically inhibited capture by anti-VAMP antibodies, and amounts as low as 0.1 pg of BoNT/B were detected. This automated SPR assay is approx. 200 times more sensitive, and 25 times more rapid, than the *in vivo* BoNT/B test currently used. Moreover, the method can be performed using a few thousand cultured neurons and constitutes a new screening assay for inhibitors.”

Dong, et al. (2004) developed sensor-based assays to detect BoNT toxin activity *in vitro*, as follows: “To detect toxin activity, fragments of the toxin substrate proteins, synaptobrevin (Syb) or synaptosome-associated protein of 25 kDa (SNAP-25), were used to link cyan fluorescent protein (CFP) to yellow fluorescent protein (YFP). Cleavage of these fusion proteins by BoNTs abolished fluorescence resonance energy transfer between the CFP and YFP, providing a sensitive means to detect toxin activity in real-time *in vitro*.” Some of the details needed to evaluate the potential of this method as an alternative to the mouse LD₅₀ assay for BoNT potency assessment were not provided; however, this looks like a promising alternative to the mouse assay.

Endopeptidase activity neutralisation has also been used as an *in vitro* approach for the detection of antibodies to BoNT (Sesardic, et al., 2004).

Gaines Das, et al. (1999) published a case study on validation efforts using the endopeptidase assay and immobilized BoNT/A synthetic substrates. They found positional effects (position of sample on the multiwell assay plate) to be a significant problem as well as other systematic effects that resulted in “a consistent bias in estimates of relative potency....” Their evaluation of possible technical problems in validating this type of endopeptidase assay could be useful in establishing future assay and validation study designs.

F. Endopep-MS assay

Scientists at the U.S. Centers for Disease Control and Prevention (CDC) have developed the Endopep-MS assay as a potential replacement for the mouse LD₅₀ assay (Barr, et al., 2005). The

in vitro Endopep-MS assay can be used to assess the potencies of and differentiate between several types of BoNTs (types A, B, E and F), “by linking the highly specific activities of the toxins with target peptide specific for each toxin serotype. The product peptides derived from the activities of the toxins are detected by matrix-assisted laser-desorption ionization of time-of-flight mass spectrometry (MS).” This method can detect toxin equivalents of as little as 0.01 of the mouse LD₅₀ dose. “An HPLC-tandem MS method for quantifying active toxin, where the amount of toxin can be correlated to the amount of product peptides”, is also available (Barr, et al., 2005).

Boyer, et al. (2005) again described the Endopep-MS assay “that detects the specific endoproteinase activities of all seven BoNT types by mass spectrometry (MS). Each BoNT type cleaves a unique site on proteins involved in neuronal transmission. Target peptide substrates based on these proteins identify a BoNT type by its enzymatic action on the substrate and the production of two peptide products, which are then detected by matrix-assisted laser desorption/ionization time-of-flight MS or liquid chromatography electrospray ionization MS/MS. The detection limits achieved range from 0.039 to 0.625 mouse LD(50)/mL for toxin types A, B, E, and F in a buffer system.” “The Endopep-MS assay meets the four requirements for BoNT detection. It detects toxin activity, is comparatively fast, and has excellent sensitivity and inherent specificity.” (Boyer, et al., 2005).

Kalb, et al., 2005 extended the previous work on Endopep-MS to include toxin subtype identification through mass spectrometric analysis on BoNT/A isolated from spiked milk.

G. *In vitro* cell-based assays

BoNTs inhibit neurotransmitter release in certain cell types and cell lines. To date, cell-based assays have not been sufficiently sensitive, and many promising cell models have yet to be developed into useful assays. The use of cell-based models for antitoxins (neutralizing antibodies) appears to be more developed than cell-based assays for detecting BoNT potency. Promising developments “are focused on cellular assays utilizing primary rat embryonic cord cells or more conveniently *in vitro* differentiated established cell lines such as human neuroblastoma cells” (Sesardic, et al., 2004).

Williamson, et al. (1996) was “the first to demonstrate in physiologically relevant cells, *i.e.* in intact functioning neurons, a direct correlation between the clostridial neurotoxin-induced block in neurotransmitter release and the cleavage of toxin-specific protein substrates, VAMP, SNAP-25, or syntaxin....” A number of cell types have been used to study BoNTs, however, “most require relatively large amounts of toxin or permeabilization to internalize the neurotoxin” (Welch, et al., 2000). Welch et al., (2000) described a primary culture of embryonic rat dorsal root ganglia (DRG) neurons that exhibits calcium-dependent substance P secretion when depolarised with elevated extracellular potassium and is naturally BoNT sensitive. The DRG neurons showed a different IC₅₀ for each of the toxins tested with a 1000 fold difference between the most and least potent neurotoxins.... BoNT/A cleavage of SNAP-25 was seen as early as 2 h...and the effects of BoNT/A were observed for as long as 15 days. This primary neuronal culture system represents a new and sensitive cellular model for the *in vitro* study of the botulinum neurotoxins.”

Hall, et al. (2004) described an *in vitro* technique for the detection and quantification of BoNT neutralizing antibodies, which would be an alternative to the MLA rather than the mouse LD₅₀ assay. The assay uses embryonic spinal cord neurons (eSCN), is more sensitive than the MLA,

and provides quantitative data. The results are based on neurotransmitter ([³H]-glycine) release, after 16 h exposure to BoNT/A, and a reduction in the level of inhibition of release when antibodies neutralize BoNT/A.

Using full-length SNAP-25 and Syb as the linkers, Dong et al. (2004) developed two fluorescent biosensors that can detect toxin activity within living cells. “Cleavage of the SNAP-25 fusion protein abolished fluorescence resonance energy transfer between CFP and YFP, and cleavage of Syb resulted in spatial redistribution of YFP fluorescence in cells. This approach provides a means to carry out cell-based screening of toxin inhibitors and to study toxin activity *in situ*. By using these biosensors, they found that the subcellular localizations of SNAP-25 and Syb are critical for efficient cleavage by BoNT/A and B, respectively.”

Chaddock, et al. (2004) used the concept that BoNTs potently and specifically inhibit neurotransmitter release in some cell types to retarget clostridial endopeptidases to cells *in vitro* to lead to inhibition of secretion of a range of transmitters for the purpose of demonstrating this technique as a potential analgesic. This is a relatively new area of investigation, and its success could result in the commercial production of even more BoNTs and the related increase in animal use for testing potency.

H. Human foot *Muscles extensor digitorum brevis*

We are including this human *in vivo* method in our discussions, not as a recommendation as an alternative to the mouse LD₅₀ assay, but because it could be useful in assessing the relevance and performance of the other alternative methods.

A summary of the method described by Bigalke, et al., (2001): Compound muscle action potentials (CMAPs) were recorded from the *M. extensor digitorum brevis* of the human foot by conventional surface techniques. The *N. peroneus* was stimulated at the ankle joint by applying supramaximal electric shocks at a distance of 8 cm from the recording electrodes. Recording took place at the site of the maximal circumference of the bulging muscle. Before injecting the toxin, a control CMAP was recorded. Varying volumes (run in quadruplicate) of a BoNT-containing solution were injected into the extensor muscles of volunteers. Each muscle was injected with one single dose only. Muscle weakness was assessed (CMAP recording) 24 h later and then at daily intervals for 6 days. To quantify the effects, the quotient of post-toxin treatment CMAP amplitude and the pretreatment CMAP amplitude was calculated (S_x/S_1). When both amplitudes were equal there was no toxin effect, and the quotient had the maximum value of 1. The amplitude of CMAPs declined over the 6 day study, however, substantial fluctuation in the responses was observed.

The human foot *Muscles extensor digitorum brevis* method appears to be a widely used clinical technique for evaluating BoNT products, and the BoNT effects in the test subjects are sometimes evaluated for up to 90 days (Bigalke, 2001; Bigalke, et al., 2001; Eleopra et al., 2002; Jost, et al., 2005).

I. Concluding comments

It is anticipated that one or more of the proposed *in vitro* alternative methods described in this section, once validated, could fully eliminate the need for *in vivo* testing of BoNT products. In the meantime, it may be appropriate to more rapidly validate and implement an *in vivo* or *ex vivo* method for use during a transition period.

IV. Test Substances

Only two types of the botulinum toxins, BoNT/A and BoNT/B, are approved for human therapeutic or cosmetic applications. Even within one toxin type, the different commercial preparations of the BoNTs are reported to have different potencies and clinical effects (Jankovic, 2004).

The first commercially available preparation, approved by the FDA in 1989, was the type A BoNT preparation, Botox, produced by Allergan, Inc. (Irvine, CA). Other BoNT/A commercial preparations that have become available include the BoNT/A products Botox Cosmetic (Allergan, Inc.), Dysport (Beaufour-Ipsen, Dreux, France; Ipsen Ltd., UK), and the Chinese product “Hengli” (Lanzhou Institute of Biological Products, Lanzhou, China) (Jankovic, 2004). The type B BoNT product known as Myobloc, which in some countries is also referred to as NeuroBloc® is produced by Elan Pharmaceuticals, Inc., Morristown, NJ. (Jankovic, 2004).

Botox and Botox Cosmetic are produced in the Republic of Ireland and in the USA, and are identical formulations, with lower doses used for the cosmetic applications (Bottrill, 2003). “Botox is manufactured with the purified native 900-kDa type A neurotoxin complex from *Clostridium botulinum* type A-Hall (Allergan) strain. This complex is composed of the botulinum neurotoxin (BoNT) and several toxin associated proteins known as the hemagglutinins (HAs) and the non-toxic non-hemagglutinin protein (NTNH)” (Zhang, et al., 2003). The complete gene sequence of the BoNT complex of type A-Hall (Allergan) strain has been reported (Zhang, et al., 2003).

In July 2004, Elan Pharmaceuticals sold its worldwide rights for Myobloc /Neurobloc injectable solution to Solstice Neurosciences, Inc. “Myobloc is now solely available through Solstice Neurosciences, and is currently licensed in the US, Canada and Europe for the treatment of cervical dystonia” (Solstice Neurosciences, 2005).

“The Centre for Applied Microbiology & Research [UK] manufactures Dysport which is then marketed by Ipsen Limited both for use in certain medical treatments, and for ancillary non-medical use” (The United Kingdom Parliament, 2002).

Jost et al. (2005) described a novel BoNT/A, NT201, that is free of complexing proteins. They reported comparable “efficacy, time to onset of action, duration of action, and tolerability” in a human clinical trial comparing NT201 with Botox (Jost, et al., 2005).

A description of Allergan’s BoNT/A is provided in the draft product label (FDA, 2004), as follows:

DESCRIPTION: BOTOX® (Botulinum Toxin Type A) Purified Neurotoxin Complex is a sterile, vacuum-dried purified botulinum toxin type A, produced from fermentation of Hall strain *Clostridium botulinum* type A grown in a medium containing casein hydrolysate, glucose and yeast extract. It is purified from the culture solution by dialysis and a series of acid precipitations to a complex consisting of the neurotoxin, and several accessory proteins. The complex is dissolved in sterile sodium chloride solution containing Albumin (Human) and is sterile filtered (0.2 microns) prior to filling and vacuum-drying.

One Unit of BOTOX® corresponds to the calculated median intraperitoneal lethal dose (LD₅₀) in mice. The method utilized for performing the assay is specific to Allergan’s product, BOTOX®. Due to specific details of this assay such as the vehicle, dilution

scheme and laboratory protocols for the various mouse LD₅₀ assays, Units of biological activity of BOTOX® cannot be compared to nor converted into Units of any other botulinum toxin or any toxin assessed with any other specific assay method. Therefore, differences in species sensitivities to different botulinum neurotoxin serotypes precludes extrapolation of animal-dose activity relationships to human dose estimates. The specific activity of BOTOX® is approximately 20 Units/nanogram of neurotoxin protein complex.

Each vial of BOTOX® contains 100 Units (U) of *Clostridium botulinum* type A neurotoxin complex, 0.5 milligrams of Albumin (Human), and 0.9 milligrams of sodium chloride in a sterile, vacuum-dried form without a preservative.

A description of the preparation and purification process for BoNT type A provided in the EP monograph (EP, 2005) is summarized as follows:

BoNT type A is produced for commercial application in fermentation batches. A highly toxigenic strain of *C. botulinum* of known toxin type A and confirmed absence of genes encoding other botulinum toxins (particularly type B), with known origin and history, is grown using suitable media. Tests used to characterize the strain include morphological, cultural, biochemical, genetic and serological properties of the strain. During development, reference preparations are established for subsequent verification of batch consistency during production and for control of the bulk purified toxin and finished product. The toxin is purified by suitable methods to remove nucleic acids and components likely to cause adverse reactions. Limits of acceptance are established and each new purified toxin is subjected to an additional battery of tests. One of the required tests is the specific activity which is confirmed in a mouse model of toxicity or by *in vivo/ex vivo* methods validated with respect to the LD₅₀ assay and expressed in mouse LD₅₀ units per milligram of protein. BoNT type A for clinical applications is provided as a dried preparation, which is reconstituted before use.

According to the EP monograph (EP, 2005), the activity of the bacterial toxin from the fermentation process is tested at several phases of the production process, as well as the testing to confirm the potency of the final product. Thus, BoNT samples from various stages of the production processes of the different manufacturers could be considered for use as test substances. Toxin samples from different master and working seed lots could also be test substances, as well as the toxins from different lots and from different manufacturers. Negative controls would include toxin samples paired with the appropriate neutralizing antibodies. Since the molecular mechanism of action is slightly different for BoNT/A and BoNT/B, type A and type B test samples would have to be tested separately when using some of the *in vitro* methods.

V. Reference Method Data and Protocol(s)

A. Reference method data

There are no reference method data from prevalidation or validation studies that are published or otherwise publicly available at this time. Such data should be obtained directly from the test method users and regulators, prior to conducting an expert workshop.

The use of existing *in vivo* data would be preferable to the generation of new *in vivo* data for validation of the proposed alternative methods. However, given the responsiveness of the LD₅₀ test method to assay conditions that include differing constituents of the test material (McLellan, et al., 1996), the LD₅₀ test and the new test methods should ideally be conducted using identical

test materials in any validation study. Existing *in vivo* data might still be used if a sample of the test material was retained and stored appropriately to retain its activity, and if this test material could be made available for the new study. We recommend that any new study at least explore the possible use of existing LD₅₀ data. Possible sources of this data (and corresponding test materials) would include: NIBSC, CDC, and BoNT manufacturers.

A collaborative study to determine the extent of inter- and intra-laboratory and inter-assay variability in potency estimates was conducted using two *in vivo* mouse LD₅₀ methods [the NIBSC method (McLellan, et al., 1996) and an in-house method] by ten laboratories in five countries (Sesardic, et al., 2003). The activities of three preparations of type A BoNT were evaluated in this study. The results and materials used in this study could be useful, if available, to new validation studies on the proposed alternative methods. This study also confirmed that the use of a reference reagent (standard) is essential to minimize the between and within laboratory variability of the mouse LD₅₀ assay. In contrast to estimates of LD₅₀ units, estimates of relative potency are obtained when using a reference standard. The study confirmed that between laboratory variability can be reduced from greater than 30% to less than 15% by the use of relative potency (Sesardic, et al., 2003). Thus, to optimize the success of any new validation study, the *in vivo* reference method protocol should be constructed to obtain the relative potency data.

B. Reference method protocol(s)

“The biological activity of therapeutic preparations of botulinum type A toxin (BoNT/A) is assessed by mouse LD₅₀ assay and the labelled unit of activity is expressed in mouse LD₅₀ units (defined as median lethal intraperitoneal dose in mice at 72 h)” (Sesardic, et al., 2003).

The following summary of the *in vivo* method was provided by Stephens and Balls (submitted): “The standard method for assessing the potency of BoNT batches is a mouse LD₅₀ test (Bottrill, 2003). In this procedure, mice are sorted into dose groups, given a single intraperitoneal injection of toxin, and monitored over 3-4 days to determine the amount required to kill 50% of a test group of mice. Death from suffocation is the endpoint, which results from paralysis of the diaphragm musculature. Although the precise details are not available, over 100 mice are used per test, and the mouse testing is carried out up to three times prior to batch release. Calculations from the test data yield an LD₅₀ value (the dose which would kill half the number of animals in a test group), which is then standardized as one ‘mouse unit’.”

Pearce, et al. (1994) evaluated the numbers of animals needed for adequate precision of the LD₅₀ assay. They concluded that using more than a 5-dose, 50-animal assay did not increase the precision of the LD₅₀ assay, and that 72 h was a sufficient time for obtaining precise data. However, for optimal evaluation of BoNT preparations for clinical use, they recommended multiple estimates of potency using the LD₅₀ assay, using 40-50 mice per estimate.

Several published methods for conducting the mouse LD₅₀ assay are provided in Appendix B.

VI. Test Method Data and Results

A. Test method data

Test method data would be in the possession of the test method developers and users, and should be obtained directly from them, prior to conducting an expert workshop.

B. Test method results

Each of the proposed test methods is summarized in Section III, and copies of the major published articles for each proposed test method are provided in Appendix A.

VII. Accuracy of the Proposed Test Method

There are no accuracy data publicly available at this time. This data would be generated during the prevalidation and validation studies for the proposed test methods.

Some of the proposed alternative methods have been compared in limited studies to the *in vivo* LD₅₀ assay, including the following:

Sesardic, et al., (2004) reported the relative sensitivities of the various potency assays when used to detect neutralizing antibodies to BoNT as follows: the mouse LD₅₀ bioassay can often detect 10 mIU/ml of antitoxin; an intramuscular paralysis assay could detect 1 mIU/ml; and a mouse *ex vivo* diaphragm assay was the most sensitive and could detect 0.5 mIU/ml.

Sesardic, et al., (1996) conducted a validation study on the abdominal ptosis assay and reported good agreement between and the LD₅₀ assay in estimating the relative potencies of 12 BoNT preparations, with a correlation coefficient (R_s) of 0.98.

VIII. Reliability (Repeatability/Reproducibility) of the Proposed Test Method

Some reliability results are discussed in the published methods of the proposed alternatives. For example, Pearce, et al., (1994, 1995a) reported results for the precision of the mouse hind-limb paralysis assay. The precision of estimates using the mouse hind-limb assay was better than the that for estimates of the LD₅₀ in mice (Pearce et al., 1994). Pearce, et al. (1995a) reported a two-fold lower coefficient of variation for the hind-limb assay results when using the same numbers of doses or animals. Sesardic, et al., (1996) provided some intra- and inter-assay statistics for the abdominal ptosis assay.

IX. Test Method Data Quality

There are no data quality data publicly available at this time. This data would be generated during prevalidation and validation studies for the proposed test methods.

X. Other Scientific Reports and Reviews

We have not conducted an exhaustive literature search, but the information and references provided in this document along with the relevant manuscripts and documents provided in Appendices, provide a thorough overview of current alternative methods to the mouse LD₅₀ botulinum potency assay.

XI. Animal Welfare Considerations (Refinement, Reduction, and Replacement)

A. History of FRAME and HSUS activities

The following section is a composite of text from two publications by Martin Stephens and Michael Balls (Stephens and Balls, submitted; Balls and Stephens, submitted):

The LD₅₀ testing of botulinum toxin products runs counter to three trends in the application of the Three Rs of replacement, reduction, and refinement, and in animal welfare generally. First, the use of the LD₅₀ test is being phased out worldwide. This was symbolized most dramatically in the field of industrial chemicals when, in 2002, the OECD deleted the LD₅₀ Test (its Test Guideline 401) from its Health Effects Test Guidelines (OECD, 2002). Second, the use of death as an endpoint is the *bête noire* of the growing field of humane endpoints (Olfert, et al., 1998; OECD, 2000; ILAR, 2000). The third trend, applicable to LD₅₀ testing of Botox Cosmetic, is the phasing out of animal testing of products with a cosmetic use. For example, in 2004, the European Union banned all forms of animal testing of cosmetic products (Europa, 2003).

In 2003, the Fund for the Replacement of Animals in Medical Experiments (FRAME) drew attention to the issue of LD₅₀ testing of BoNT products with the publication of an exposé entitled “Growing Old Disgracefully ...” (Bottrill, 2003; Balls, 2003). Bottrill’s article pointed out the paradox that “the general public had enthusiastically adopted the use of a BoNT product as a wrinkle treatment against a background of moves to stop the animal testing of cosmetic ingredients and cosmetic products – and in particular, Ipsen Limited UK was producing and testing their product, Dysport, in the UK, although the Government claimed to have banned all testing of cosmetics. The Government’s response was that Dysport was only licensed for therapeutic use, and that individual clinicians and their patients had themselves to accept the responsibility for using the product for other purposes” (Balls and Stephens, submitted). It’s also been proposed that manufacturers and governments may find it convenient not to be able to define whether the batches of BoNT being tested will be used for therapeutic or cosmetic purposes (Balls and Stephens, submitted).

Since the suffering of mice is a matter of great concern, whatever the eventual use of the product, Bottrill also raised the question of why an LD₅₀ test was still being used for any BoNT products, given that this kind of procedure had been replaced by more-humane methods in the OECD Health Effects Test Guidelines for oral systemic acute toxicity testing, and listed a number of proposals for reducing the numbers of animals used, refining the procedure in order to minimize the suffering caused to them, and eventually replacing the need for animal testing altogether.

The NIBSC, the UK national control agency, has made efforts to develop and use refinement and replacement alternatives. The NIBSC uses *in vitro* methods on a routine basis, and only uses a non-lethal *in vivo* test when, rarely, the results of an *in vitro* test are inconclusive or close to pass/fail specifications. The NIBSC has also put some effort into the development of methods which could totally obviate the need for animal testing (eg, Ekong et al, 1997). Ipsen Limited UK is working with the NIBSC and others to develop suitable batch release tests.

Ipsen Limited UK continues to use the mouse LD₅₀ test to measure the potency of Dysport. FRAME considers that the Government (the UK Home Office) should close the loophole which permits BoNT destined to be used for cosmetic purposes to be tested in animals, despite the ban on testing cosmetic products in the UK. The claim by the producer is that BoNT is only officially licensed (and therefore tested) for clinical use, and its use for cosmetic purposes involves a private contract between a physician and a patient, at their own risk.

At the European level, the European Centre for the Validation of Alternative Methods (ECVAM) and the European Directorate for the Quality of Medicines (EDQM) are working together and with others to review what progress is being made in applying the Three Rs to BoNT testing and to assist in moving forward.

The HSUS campaign to replace the mouse LD₅₀ assay for BoNT potency testing has focused exclusively on the testing of Botox Cosmetic by its manufacturer, Allergan, Inc. The HSUS regards the LD₅₀ testing of products for cosmetic use as unacceptable. Beginning in January 2004, the HSUS first attempted to work *with* Allergan. Three things were sought from Allergan: (1) public disclosure of the details of its current potency testing of Botox Cosmetic; (2) public disclosure of the details of its current efforts to develop alternatives to the mouse LD₅₀ testing of Botox Cosmetic; and (3) adoption of a well-funded and publicly available plan to rapidly end the LD₅₀ testing of Botox Cosmetic. The attempted dialog with Allergan was largely unproductive (Stephens and Balls, submitted).

Allergan did confirm that the company uses the mouse LD₅₀ assay to assess the potency of Botox Cosmetic, and claimed to have an active alternatives program to replace this testing. However, the company provided few details either of its current testing practices or of its alternatives efforts. Allergan claimed that Botox Cosmetic and its sister product, Botox, share the same active ingredient, BoNT type A, so LD₅₀ testing for the two products is inextricably linked and testing for cosmetic purposes cannot be cleanly separated from testing for therapeutic purposes. Allergan also noted the international regulations calling for the LD₅₀ testing of BoNT products.

Beginning in October 2004, the HSUS appealed to its members and constituents to urge Allergan to work with the HSUS on rapidly replacing LD₅₀ testing for Botox Cosmetic. Since Allergan refused to work with the HSUS or to disclose information about its testing and alternatives practices, the HSUS also turned to the U.S. Food and Drug Administration (FDA), which had approved Botox Cosmetic and Botox. The FDA regulates these products as pharmaceuticals, and now oversees their manufacture and sale. The HSUS was specifically interested in information about the potency testing currently required or encouraged for these products. It was hoped that the agency could help answer several key questions, including the following: (1) What are the current testing practices? (2) Does the FDA require or encourage these practices? (3) How have these practices changed over the years? and (4) What is the FDA itself doing to promote LD₅₀ alternatives? In 2004, The HSUS filed two *Freedom of Information Act* requests with the FDA, in order to obtain the sought-after information, but the agency was largely unresponsive. The HSUS then took legal action against the agency, which is still pending. The HSUS anticipates that the FDA could provide critical information about BoNT testing, including the numbers of animals used per test and the number of tests conducted prior to release of a given batch of product.

The LD₅₀ testing of BoNT products in general, and of products for cosmetic use in particular, is particularly out of step with the times, and is therefore particularly in need of scrutiny and action. Significant progress, however, has been hampered by the lack of publicly available details about current testing practices and alternatives efforts. The technical challenges to developing a non-animal alternative for BoNT product testing will be best met with collaborative efforts open to scrutiny and to constructive criticism. FRAME and The HSUS are now seeking to work collaboratively with institutions such as ECVAM and the ICCVAM to accelerate the pace of progress, in the confident belief that we all share the same interest in making available products which are made as safe as possible for human use, but by using modern methods and

progressively reducing reliance on the traditional application of painful test procedures to laboratory animals.

B. *Animal welfare considerations specific to the proposed test methods*

The local paralysis assays in mice, either *in vivo* or *ex vivo* (see Table 1), represent refinements to the LD₅₀ assay, which causes greater distress and suffering due to death from systemic paralysis and suffocation. The *in vivo* mouse hind limb assay takes two days to complete compared to less than one day for the abdominal ptosis assay, possibly making the latter less distressful to the animal. The *ex vivo* mouse phrenic nerve-hemidiaphragm system represents another step toward refining the LD₅₀ assay by replacing the use of a live animal with the use of a mouse muscle tissue.

The *in vitro* methods, SNAP- 25/endopeptidase assay, Endopep-MS, and possibly others, are proposed replacements for the mouse LD₅₀ assay for botulinum potency testing. The validation, use, and regulatory acceptance of one or more of these *in vitro* methods would address all of the animal welfare concerns related to the potency testing of BoNT products.

These same alternative methods that are used to test for potency can be used to replace the mouse assay in additional applications such as detecting BoNT in diagnostic samples, and can be slightly modified to measure antitoxin (neutralizing antibodies), which is another area of increasing animal use where the current standard method is the MLA. Therefore, validation and acceptance of one or more of the proposed alternative methods would also potentially contribute to additional reductions in animal use by replacing the LD₅₀ assay for these applications.

XII. Practical Considerations

The alternatives to the mouse LD₅₀ assay provide other advantages compared to current methods. Many of the proposed alternative assays would reduce the overall costs of potency testing, and all would require less time than the current method to complete and obtain results. Some of the proposed methods, however, would require more laboratory technician time to perform the actual assay compared to the mouse bioassay, although the end result would typically be obtained sooner. The costs of the different methods are not known at this time; however, a few of the publications on the alternative methods have mentioned that they require no special reagents or equipment. The endopep-MS, which requires a mass spectrometer, could be prohibitively expensive for institutions where this instrumentation is not currently available. However, this instrumentation is widely available at the present time. All of the assays should be transferable to other labs providing that appropriate training and a good protocol are provided.

XIII. References

Allergan, Inc. (2005). Allergan Reports Fourth Quarter Operating Results; Pharmaceutical Sales Increased 16 Percent for the Fourth Quarter; Board of Directors Declares Increase of Quarterly Dividend. Available at:

<http://www.shareholder.com/AGN/ReleaseDetail.cfm?ReleaseID=155017>.

American Society for Aesthetic Plastic Surgery (ASAPS). (2005). Cosmetic Surgery National Data Bank 2004. Available at: <http://www.surgery.org/download/2004-stats.pdf>.

Aoki, K.R. (2001). A comparison of the safety margins of botulinum neurotoxin serotypes A, B, and F in mice. *Toxicon* 39(12), 1815-1820.

Arnon, S.S., Schechter, R., Inglesby, T.V., Henderson, D.A., Bartlett, J.G., et al. (2001). Botulinum toxin as a biological weapon. Medical and public health management. Botulinum toxin as a biological weapon. *JAMA* 285, 1059-1070.

Balls, M. (2003). Botulinum toxin testing in animals: the questions remain unanswered. *ATLA* 31, 611-615.

Balls, M. and Stephens, M. (submitted). What price vanity? The urgent need to replace painful animal potency tests for botulinum toxin products. Submitted.

Barr, J.R., Moura, H., Boyer, A.E., et al. (2005). Botulinum neurotoxin detection and differentiation by mass spectrometry. *Emerging Infectious Diseases* 11, 1578-1583.

Bigalke, H., (2001). Pharmacology of botulinum neurotoxin type A. *Wien Klin. Wochenschr.* 113 Suppl 4, 2-5.

Bigalke, H., Wohlfarth, K., Irmer, A. and Dengler, R. (2001). Botulinum A toxin: Dysport improvement of biological availability. *Experimental Neurology* 168, 162-170.

Bottrill, K. (2003). Growing old disgracefully: the cosmetic use of botulinum toxin. *ATLA* 31, 381-391.

Boyer, A.E., Moura, H., Woolfitt, A.R., Kalb, S.R., McWilliams, L.G., Pavlopoulos, A., Schmidt, J.G., Ashley, D.L. and Barr, J.R. (2005). From the mouse to the mass spectrometer: detection and differentiation of the endoprotease activities of botulinum neurotoxins A-G by mass spectrometry. *Anal. Chem.* 77, 3916-3924.

Centers for Disease Control and Prevention (CDC) (1998). Botulism in the United States, 1899-1996. Handbook for Epidemiologists, Clinicians, and Laboratory Workers, Atlanta, GA. Available at: <http://www.cdc.gov/ncidod/dbmd/diseaseinfo/botulism.pdf>.

CenterWatch. (2004). Drugs approved by the FDA. Available at: <http://www.centerwatch.com/patient/drugs/dru722.html>.

Chaddock, J.A., Purkiss, J.R., Alexander, F.C., et al. (2004). Retargeted clostridial endopeptidases: inhibition of nociceptive neurotransmitter release in vitro, and antinociceptive activity in *in vivo* models of pain. *Mov. Disord.* 19 Suppl 8, S42-S47.

Dong, M., Tepp, W.H., Johnson, E.A. and Chapman, E.R. (2004). Using fluorescent sensors to detect botulinum neurotoxin activity in vitro and in living cells. *Proc. Natl. Acad. Sci. USA* 101(41), 14701–14706.

Dressler, D., Lange, M., Bigalke, H. (2005). Mouse diaphragm assay for detection of antibodies against botulinum toxin type B. *Mov Disord.* 2005 Aug 2; [Epub ahead of print].

Ispen. (2004). About Dysport. Available at: <http://www.dysport.co.nz/about.php>.

Ekong, T.A.N., Feavers, I.M, and Sesardic, D. (1997). Recombinant SNAP-25 is an effective substrate for *Clostridium botulinum* type A toxin endopeptidase activity *in vitro*. *Microbiology* 143, 3337-3347.

Ekong, T.A., McLellan, K., Sesardic, D. (1995). Immunological detection of *Clostridium botulinum* toxin type A in therapeutic preparations. *J. Immunol. Methods* 180(2), 181-191.

Eleopra, R., Tugnoli, V., Quatralo, R., Gastaldo, E., Rossetto, O., De Grandis, D. and Montecucco, C. (2002). Botulinum neurotoxin serotypes A and C do not affect motor units survival in humans: an electrophysiological study by motor units counting. *Clin. Neurophysiol.* 113(8), 1258-1264.

Environmental Protection Agency (EPA). (1997). Rule on microbial products of biotechnology: summary of the public's comments and the agency's response. Available at: http://www.epa.gov/biotech_rule/c005.htm.

Environmental Protection Agency (EPA). (2002). Develop routine and rapid methods to analyze biological agents. National Exposure Research Laboratory. Available at: <http://www.epa.gov/nerl/research/2002/pdf/g8-1.pdf>.

Europa. (2003). Directive 2003/15/EC of the European Parliament and of the Council of 27 February 2003 amending Council Directive 76/768/EEC on the approximation of the laws of the Member States relating to cosmetic products (Text with EEA relevance). Available at: <http://europa.eu.int/eur-lex/lex/LexUriServ/LexUriServ.do?uri=CELEX:32003L0015:EN:HTML>.

European Pharmacopeia (EP). (2005). Botulinum toxin type A for injection. *European Pharmacopeia Monograph 5.0*, 1117-1119.

Ferracci, G., Miquelis, R., Kozaki, S., Seagar, M., Leveque, C. (2005). Synaptic vesicle chips to assay botulinum neurotoxins. *Biochem. J.* 391(Pt 3), 659-666.

Food and Drug Administration (FDA). (2002). Botox Cosmetic: A look at looking good. *FDA Consumer Magazine*. Available at: http://www.fda.gov/fdac/features/2002/402_botox.html.

Food and Drug Administration (FDA). (2003a). FDA Product Approval Information for BOTOX. Available at: <http://www.fda.gov/cder/biologics/products/botaller122100.htm>; September 16, 2003.

Food and Drug Administration (FDA). (2003b). FDA Product Approval Information for BOTOX COSMETIC: Available at: <http://www.fda.gov/cder/biologics/products/botuall041202.htm>; September 24, 2003.

Food and Drug Administration (FDA). (2004). BOTOX® (Botulinum Toxin Type A) Purified Neurotoxin complex. Available at: http://www.fda.gov/medwatch/SAFETY/2004/jul_PI/Botox_PI.pdf.

Gaines Das, R.E., Heath, A.B., Martin, H. and Sesardic, D. (1999). Validation of *in vitro* assays for botulinum toxin: A case study. In: *Alternatives to Animals in the Development and Control of Biological Products for Human and Veterinary Use*. (Eds. Brown, F., Hendriksen, C., Sesardic, D.) *Dev. Biol. Stand.* 101, 267–276.

Hall, Y.H., Chaddock, J.A., Mouldsdale, H.J., Kirby, E.R., Alexander, F.C., Marks, J.D. and Foster, K.A. (2004). Novel application of an *in vitro* technique to the detection and quantification of botulinum neurotoxin antibodies. *J. Immunol. Methods* 288(1-2), 55-60.

Institute for Laboratory Animal Research (ILAR). (2000). Humane Endpoints for Animals Used in Biomedical Research and Testing. *ILAR Journal* 41, No. 2.

Jankovic, J. (2004). Botulinum toxin in clinical practice. *Journal of Neurology, Neurosurgery, and Psychiatry* 75, 951-957.

Jost, W.H. and Kohl, A. (2001). Botulinum toxin: evidence-based medicine criteria in rare indications. *J. Neurol.* 248, Suppl. 1, 39-44.

Jost, W.H., Kohl, A., Brinkmann, S., Comes, G. (2005). Efficacy and tolerability of a botulinum toxin type A free of complexing proteins (NT 201) compared with commercially available botulinum toxin type A (BOTOX degrees) in healthy volunteers. *J. Neural Transm.* 112(7), 905-913.

Kadkhodayan, S., Knapp, M. S., Schmidt, J. J., Fabes, S. E., Rupp, B. & Balhorn, R. (2000). Cloning, expression, and one-step purification of the minimal essential domain of the light chain of botulinum neurotoxin type A. *Protein Expression Purif.* 19(1), 125-130.

Kalandakanond, S. and Coffield, J.A. (2001a). Cleavage of intracellular substrates of botulinum toxins A, C, and D in a mammalian target tissue. *J. Pharmacol. Exp. Ther.* 296(3), 749-755.

Kalandakanond, S. and Coffield, J.A. (2001b). Cleavage of SNAP-25 by botulinum toxin type A requires receptor-mediated endocytosis, pH-dependent translocation, and zinc. *J. Pharmacol. Exp. Ther.* 296(3), 980-986.

Kalb, S.R., Goodnough, M.C., Malizio, C.J., Pirkle, J.L. and Barr, J.R. (2005). Detection of botulinum neurotoxin A in a spiked milk sample with subtype identification through toxin proteomics. *Anal. Chem.* 77(19), 6140-6146.

Keller, J.E., Cai, F., Neale, E.A. (2004). Uptake of botulinum neurotoxin into cultured neurons. *Biochemistry* 43(2), 526-532.

Lalli, G., Bohnert, S., Deinhardt, K., Verastegui, C. & Schiavo, G. (2003). The journey of tetanus and botulinum neurotoxins in neurons. *Trends Microbiol.* 11(9), 431-437.

McLellan, K., Das, R.E., Ekong, T.A., Sesardic, D. (1996). Therapeutic botulinum type A toxin: factors affecting potency. *Toxicon* 34(9), 975-985.

National Toxicology Program (NTP) Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM), ICCVAM. (2003). *Guidelines for the Nomination and Submission of New, Revised, and Alternative Test Methods*, NIH Publication No. 03-4508, National Institute of Environmental Health Sciences, National Institutes of Health, U.S. Public Health Service, Department of Health and Human, September 2003.

Naumann, M., and Jankovic, J. (2004). Safety of botulinum toxin type A: a systematic review and meta-analysis. *Curr. Med. Res. Opin.* 20(7), 981-990.

Olfert, E., Bhasin, J., Latt, R., Macallum, E., McCutcheon, K., Rainnie, D. and Schunk, M. (1998). Guidelines on: choosing an appropriate endpoint in experiments using animals for research, teaching and testing. Ottawa: Canadian Council on Animal Care.

Organisation for Economic Cooperation and Development (OECD). (2000). Guidance Document on the Recognition, Assessment, and Use of Clinical Signs as Humane Endpoints for Experimental Animals Used in Safety Evaluation. No.19. Paris: OECD.

Organisation for Economic Cooperation and Development (OECD). (2002). OECD Test Guideline 401 will be deleted: A major step in animal welfare. Available at: http://www.oecd.org/document/52/0,2340,en_2649_34377_2752116_1_1_1_1,00.html.

Pearce, L.B., Borodic, G.E., First, E.R. and MacCallum, R.D. (1994). Measurement of botulinum toxin activity: evaluation of the lethality assay. *Toxicol. Appl. Pharmacol.* 128(1), 69-77.

Pearce, L.B., Borodic, G.E., Johnson, E.A., First, E.R. and MacCallum, R. (1995a). The median paralysis unit: a more pharmacologically relevant unit of biologic activity for botulinum toxin. *Toxicon* 33(2), 217-227.

Pearce, L.B., First, E.R. and Borodic, G.E. (1995b). Botulinum toxin: death versus localized denervation. *J. R. Soc. Med.* 88(4), 239-240.

Pearce, L.B., First, E.R., MacCallum, R.D. and Gupta, A. (1997). Pharmacologic characterization of botulinum toxin for basic science and medicine. *Toxicon* 35(9), 1373-1412.

Poewe, W. (2002). Respective potencies of Botox and Dysport: a double blind, randomised, crossover study in cervical dystonia. Editorial Commentary. *Journal of Neurology Neurosurgery and Psychiatry* 72, 430.

Roger Aoki, K. (2002). Botulinum neurotoxin serotypes A and B preparations have different safety margins in preclinical models of muscle weakening efficacy and systemic safety. *Toxicon* 40(7), 923-928.

Schmidt, J.J., Stafford, R.G. and Millard, C.B. (2001). High-throughput assays for botulinum neurotoxin proteolytic activity: serotypes A, B, D, and F. *Anal. Biochem.* 296(1), 130-137.

Segelke, B., Knapp, M., Kadkhodyan, S., Balhorn, R. and Rupp, B. (2004). Crystal structure of *Clostridium botulinum* neurotoxin protease in a product-bound state: Evidence for noncanonical zinc protease activity. *Proc. Natl. Acad. Sci. U SA* 101(18), 6888-6893.

Sesardic, D., Jones, R.G.A., Leung, T., Alsop, T. and Tierney, R. (2004). Detection of Antibodies Against Botulinum Toxins. *Movement Disorders* Suppl. 8, S85-S91.

Sesardic, D. (2002). Alternatives in testing of bacterial toxins and antitoxins. *Dev. Biol. (Basel)* 111, 101-108.

Sesardic, D., Leung, T. and Gaines Das, R. (2003). Role for standards in assays of botulinum toxins: international collaborative study of three preparations of botulinum type A toxin. *Biologicals* 31, 265-276.

Sesardic, D., McLellan, K., Ekong, T.A. and Das, R.G. (1996). Refinement and validation of an alternative bioassay for potency testing of therapeutic botulinum type A toxin. *Pharmacol. Toxicol.* 78(5), 283-288.

Solstice Neurosciences. (2005). About Myobloc®. Available at:
http://www.myobloc.com/hp_about/about.html.

Stephens, M.L. and Balls, M. (submitted) LD₅₀ Testing of botulinum toxin for use as a cosmetic, *ALTEX*, Submitted.

Takahashi, M., Kameyama, S. and Sakaguchi, G. (1990a). Assay in mice for low levels of *Clostridium botulinum* toxin. *Int. J. Food Microbiol.* 11(3-4), 271-277.

Takahashi, M., Noda, H., Takeshita, S., Fujiwara, T., Nakanoin, H., Mizunoya, T. and Sakaguchi, G. (1990b). Attempts to quantify *Clostridium botulinum* type A toxin and antitoxin in serum of two cases of infant botulism in Japan. *Jpn. J. Med. Sci. Biol.* 43(6), 233-237.

Tang-Liu, D.D., Aoki, K.R., Dolly, J.O., de Paiva, A., Houchen, T.L., Chasseaud, L.F. and Webber, C. (2003). Intramuscular injection of 125I-botulinum neurotoxin-complex versus 125I-botulinum-free neurotoxin: time course of tissue distribution. *Toxicon* 42(5), 461-469.

The United Kingdom Parliament. (2002). National Institute for Clinical Excellence. Available at: <http://www.parliament.the-stationery-office.co.uk/pa/cm200102/cmhansrd/vo020613/text/20613w24.htm>.

United States Congress. (2000). ICCVAM Authorization Act of 2000. Public Law 106-545. Available at: <http://iccvam.niehs.nih.gov/about/PL106545.pdf>.

United States Department of Agriculture (USDA). (2005). APHIS Agricultural Select Agent Program. Available at: http://www.aphis.usda.gov/programs/ag_selectagent/ag_bioterr_toxinslist.html.

Welch, M.J., Purkiss, J.R., Foster, K.A. (2000). Sensitivity of embryonic rat dorsal root ganglia neurons to *Clostridium botulinum* neurotoxins. *Toxicon* 38(2), 245-258.

Wictome, M., Newton, K., Jameson, K., et al. (1999). Development of an *in vitro* bioassay for *Clostridium botulinum* type B neurotoxin in foods that is more sensitive than the mouse bioassay. *Appl Environ Microbiol.* 65, 3787-3792.

Williamson, L.C., Halpern, J.L., Montecucco, C., Brown, J.E. and Neale, E.A. (1996). Clostridial neurotoxins and substrate proteolysis in intact neurons. *JBC* 271(13), 7694-7699.

Zhang, L., Lin, W.J., Li, S., and Aoki, K.R. (2003). Complete DNA sequences of the botulinum neurotoxin complex of *Clostridium botulinum* type A-Hall (Allergan) strain. *Gene* 315, 21-32.

APPENDIX A

Methods and/or protocols for proposed test method(s)

Published articles containing methods and/or protocols for the following proposed methods are provided in this appendix:

1. Mouse hind-limb assay

Pearce, et al., 1994

Pearce, et al., 1995a

Pearce, et al., 1995b

Roger Aoki K., 2002

Aoki, 2001

2. Abdominal ptosis assay

Takahashi, et al., 1990a

Sesardic, et al., 1996

3. Mouse phrenic nerve-hemidiaphragm assay

Bigalke, et al., 2001

Bigalke, 2001

Kalandakanond and Coffield, 2001a

Kalandakanond and Coffield, 2001b

4. SNAP-25/endopeptidase assay

Ekong, et al., 1997

Wictome, et al., 1999

Ferracci, et al., 2005

Dong, et al., 2004

5. Endopep-MS assay

Barr, et al., 2005

Boyer, et al., 2005

6. Neural cell-based assays

Dong, et al., 2004

APPENDIX B

Methods and/or protocol(s) for generation of reference method data

Published articles containing methods and/or protocols for the reference method are provided in this appendix.

Pearce, et al., 1994

McLellan, et al., 1996

Sesardic, et al., 2003

APPENDIX C

Copies of other relevant manuscripts and documents

Comparative reviews of alternatives to the mouse LD₅₀ for BoNT potency testing:

Pearce, et al. (1997)

Bottrill (2003)

Balls (2003)

Sesardic, et al., 2004 [alternatives for potency and neutralizing antibody testing]

Other relevant documents:

Bottrill, 2003

Balls, 2003

Balls and Stephens, 2006

Gaines Das, et al., 1999

Stephens and Balls, 2006