Opinion Article

Botulinum Toxin Testing on Animals is Still a Europe-Wide Issue

Katy Taylor, Corina Gericke and Laura Rego Alvarez
European Coalition to End Animal Experiments, London, UK

Abstract
There have been significant developments in the use of animals to test Botulinum toxin products in Europe in recent years. This paper summarises and discusses these from the perspective of the animal protection organisation. A cell-based assay has been validated by Allergan and is now being used for the replacement of the mouse bioassay for the batch testing of their Botulinum toxin A products. Two further companies (Merz and Ipsen) have recently validated similar cell-based assays to replace animals in their batch testing.
However, the number of animals being used in batch tests across Europe remains at record levels; an estimated 400,000 animals per year, based on official statistics and non-technical summaries. There are concerns from animal protection organisations about the authorisation of animal testing for Botulinum toxin products that are to be used for aesthetic purposes. Furthermore, should testing for companies that have not yet implemented the alternative method continue to be permitted under EU Directive 2010/63 on the use of animals for scientific purposes? Whilst we are on the cusp of an era where the mouse bioassay has been replaced for the potency testing of Botulinum toxin A for injection, it is important that Europe sees a reduction of animal testing in real terms.

1 Introduction

Botulinum toxin (Bt) via injection has been used therapeutically for a variety of medical and cosmetic (hereafter "aesthetic") purposes for approximately 30 years (see Bottril, 2003).

The use of animals to test batches of Bt has been a concern to animal protection groups and others since the early 2000s when the use of Bt for aesthetic purposes began to dramatically increase worldwide. The tests are of concern for a number of reasons: a) the severity of the suffering experienced by the animals, b) the high numbers of animals used on an annual basis, c) the trivial purpose for which these products are often used, and d) the delay in implementing a complete replacement of the animal test for all Bt products. A number of authors have discussed these issues in the scientific literature (Balls, 2003, 2010; Bottril, 2003; Straughan, 2006; Adler et al., 2010; Bitz, 2010; Pickett, 2011, 2012). However, there have been some recent, significant developments in this area and a review of the current status of the use of animals and alternatives is now overdue.

We use "aesthetic" throughout this paper as some consider that it is misleading to use the term "cosmetic", since the Bt products are injectables and therefore not cosmetic products as defined by the EU Cosmetics Regulation 1223/2009 (EC, 2009; Pickett, 2011). However, as we discuss here, the general public would describe the use of these products for aesthetic purposes as "cosmetic" and indeed even the General Medical Council of the UK[1] refer to the use of Bt for aesthetic purposes as a "cosmetic" procedure.

The European Coalition to End Animal Experiments (ECEAE) has had a long running campaign against animal testing for Bt products and has been in regular contact with the manufacturers and regulatory authorities. In this article we summarise the progress with the implementation of alternatives to the animal test for batch testing Bt products and highlight some of the ongoing issues. Our evidence is drawn from official statistical reports, company reports, non-technical summaries of animal experiments published by member states, investigation reports, legal cases, and correspondence with the manufacturers.

2 The mouse bioassay

Botulinum toxin type A (BtA) produced for injection is a dried preparation of neurotoxin. It is obtained by purification of the

[1] General Medical Council. Guidance for doctors who offer cosmetic interventions (12.04.2016). https://bit.ly/2KESF5U

Received July 10, 2018; Accepted October 4, 2018;
Epub October 10, 2018;

ALTEX 36(1), 81-90. doi:10.14573/altex.1807101

Correspondence: Katy Taylor, PhD, European Coalition to End Animal Experiments, 16a Crane Grove, London, N7 8NN, UK
(katy.taylor@crueltyfreeinternational.org)

liquid from a broth-culture of *Clostridium botulinum* type A bacteria. The substance produced is a complex of proteins derived from this biotechnology process and therefore the composition and biological activity of each batch can vary. This, together with the fact that Bt is one of the most dangerous neurotoxins known to man, means that each batch produced needs to be quality-controlled to ensure it is of the same consistency and potency as previous batches (Adler et al., 2010).

The European Pharmacopoeia (PhEur) is the legally recognised list of acceptable methods of quality control for medicinal products that are to be marketed in Europe. Monograph 2113 of the PhEur (EDQM, 2012) requires the use of a mouse bioassay (MBA) to assess the consistency and potency of products containing botulinum toxin A for injection. Monograph 2584 (EDQM, 2011) has very similar requirements for injectable products derived from botulinum toxin type B. The MBA is essentially an LD50 test; graded doses of the product are injected into mice and the LD50 is calculated from the lethality in each dose group.

According to the PhEur, *bulk preparations* of purified BtA have to be assayed by the MBA as well as *every batch of the final product* that is derived from the bulk preparation. It is the potency testing of the batches of Bt products that are responsible for the high numbers of mice used (see below). This is because many batches are derived from the same bulk preparation. The same MBA tests are also periodically required by drug regulators to demonstrate the stability of the products according to ICHQ5C (ICH, 1995), typically on an annual basis or when changes to the production process or the products themselves are made (process validation, according to ICH Q5E (ICH, 2004)) (Adler et al., 2010).

3 Severe suffering

The MBA, as performed according to the PhEur, involves groups of mice being injected into their peritoneal cavity with differing dilutions of the Bt product. After being injected, the mice are placed back into their cages in small groups for the duration of the test (usually 72 or 96 hours). The numbers of mice that have died or are killed by "humane endpoint" (see below) by the end of the test period are counted. The LD50 is calculated based on the number of mice dying in the various dilution groups. Approximately 90% of the mice in the highest concentration group are expected to die, 10% in the lowest (Adler et al., 2010).

For those animals receiving a sufficient dose of toxin, signs of poisoning start to show within hours. The main effect is paralysis of the lower body; affected mice begin to stagger and those more severely affected cease to be able to walk. As the paralysis develops over the first 24 hours, it affects the muscles of respiration and the mice will begin to gasp for air and become cyanotic. Those who die are considered to have succumbed to asphyxiation. One particular concern with the assay, however, is that the more severely affected mice cannot reach food or water and may therefore be dying as a result of dehydration and weight loss and not the toxin *per se*. This concern applies to LD50 testing in general; "Mortality of animals is often the result of lack of food and water only, not the primary effects of the substance. If small rodents are not capable of feeding, they die within hours" (Hartung and Koëter, 2008). This uncontrolled variable could explain why variations in LD50 values have been observed between laboratories even for the same preparation of the same product (Sesardic et al., 2003).

No pain relief is given to the animals as the effects of the toxin are in themselves not expected to be painful . Nevertheless, tests with death as endpoint or when death can be expected are generally recognised to warrant a recording of severe suffering, the highest suffering permitted under Directive 2010/63 governing animal experiments in the EU (EC, 2010). The authorities in Great Britain, Ireland, and Germany, where it is public knowledge that batch testing of Bt products is conducted, all prospectively categorise these tests under the severe category.

In 2009, the British Union for the Abolition of Vivisection (BUAV, now Cruelty Free International) released footage of the mice undergoing these tests during an undercover investigation of a UK laboratory (Wickham Laboratories) that was conducting Bt tests on behalf of Ipsen, a manufacturer of BtA products, Dysport® and Azzalure®[3]. A subsequent inquiry by the Home Office, who are responsible for regulating animal experiments in Great Britain, found Wickham to be in breach of their project licence by causing unnecessary animal suffering. They had previously asked Wickham to adopt a procedure of checking the animals in hourly intervals during the peak period of likely severe suffering. Technicians were trained to kill mice if they looked like they would be unlikely to survive until the next observation period in an attempt to reduce suffering.[2] BUAV claimed that Wickham's records showed that out of the approximately 50% of the animals dying as a consequence of the test, only about 25% of the mice were "humanely" killed, 75% were being found dead.[3] There were other breaches such as mice being killed using inappropriate methods. The Home Office admitted that, despite the laboratory's efforts, the proportion of mice that were "humanely killed" as opposed to being found dead was typically around 20%. In a follow-up after the inquiry the Home Office found that this had only moderately improved to 32%.[2]

It seems that there is reluctance to kill the mice too early in the test for fear of affecting the results (Adler et al., 2010;[2]). In the ZEBET workshop on Bt testing, one company admitted that only between 11-21% of the dead mice had been humanely killed (Adler et al., 2010). The participants considered that further refinement of the test, by killing mice not *in extremis*, would require separate validation (Adler et al., 2010). Indeed, in recent correspondence the Irish Health Product Regulatory Authority say that there is an ongoing validation of a humane

[2] Animals Scientific Procedures Inspectorate. A review on the issues and concerns raised in the report, "The Ugly Truth – A BUAV Investigation at Wickham Laboratories". Nov 2010.

[3] British Union for the Abolition of Vivisection. The Ugly Truth: A BUAV Investigation at Wickham Laboratories, Nov 2009 (available upon request).

endpoint protocol that is currently requiring the use of additional animals.[4]

4 High numbers

The number of animals used per batch test is not described in the PhEur and it appears that it may vary between manufacturers (Pickett, 2011). The manufacturers do not comment publicly about the numbers they use per batch (Bitz, 2010; Pickett, 2011). Bottrill (2003) estimated that each MBA may use "at least 100 mice". The BUAV investigation found that a typical MBA used "hundreds" of animals but were requested by Ipsen to keep the exact figure confidential. Since the MBA is a quality-control test and must be performed for every batch released onto the market, the total number of animals used for the testing of Bt products can quickly escalate as production of Bt products rises. Indeed, as the use of Bt products for both medical and aesthetic purposes has increased over the last 15 years, this has undoubtedly had an effect on the numbers of animals used.

Contract testing laboratories in Great Britain, Ireland, and Germany are known to be testing Bt products on behalf of the manufacturers. Figure 1 shows the trends in animal use in Great Britain, Ireland, and Germany from 2006 to 2013 for acute lethal tests in mice, which constitute mainly of the MBA for Bt products based on our knowledge of the licensing practices at that time. Since 2014, under the new rules for reporting the use of animals in scientific procedures across the EU (EC, 2012), the category where the MBA for Bt products should be recorded changed to "Quality control: batch potency". Figure 1 includes the numbers given by Great Britain, Ireland, and Germany for 2015 and 2016 in this category. The numbers for 2014 were discounted as it appears incorrect reporting fields were used by Great Britain and Ireland at least as they settled into the new system.

The numbers in Figure 1 appear to suggest that Great Britain was already conducting many MBAs in 2006. Indeed, a previous BUAV investigation of Wickham Laboratories had noted the testing of Dysport® on animals in 1992.[3] During the 2000s the numbers of acute tests in Great Britain was consistently around 80,000 animals per year until 2010 when there was a small decline to less than 70,000, from when it then began to increase and by 2013 was in excess of 100,000 mice per annum. In 2009, Wickham laboratories was licensed to conduct the MBA on up to 70,000 mice per year,[3] however in 2014 the licence was renewed and the number had increased to 100,000 mice per year.[5] Wickham currently test for three (undisclosed) Bt companies.[6] The total number of mice used in batch potency tests was 144,957 in 2015 and 130,973 in 2016 according to the national statistics.

Testing appears to have started in Germany in 2006/7, when the number of acute tests rose from approximately 5,000 per year to over 20,000. The testing was started at the Laboratory for Pharmacology and Toxicology (LPT) in Hamburg at the request of Merz for their Bt products.[7] In 2014, LPT began testing on behalf of Eisai as well[8] and the number of mice used in batch potency tests escalated to 112,139 in 2015 according to the new categorisation. We suspect that the acute lethal category in the statistics up to 2013 was not capturing all the mice actually used in Bt batch testing as Merz reported in 2015 that 35,000 mice per annum had been their normal level of testing.[9] A total of five non-technical summaries for 2014 permitted testing on 150,000 mice in total for Bt products that year. In 2015 the summaries predicted a decrease in testing to just two projects totalling 48,000 animals. The total number of mice used in batch potency tests however was greater than this; 78,932 in 2016, according to the German statistics.

The scale of acute tests in Ireland is particularly significant. Testing appears to have started in Ireland in 2007/8 when the numbers of animals used rose from 1,000 to 42,000. By 2013 the total animals used had steadily risen and approached 200,000. In 2015, 150,030 animals were reported to have been used in batch potency tests and at 167,549 batch potency tests in 2016 still represents the majority (74%) of all animal procedures conducted in Ireland.

The total number of mice[10] used for quality control batch purposes in 2015 in the three countries was 407,126 and 377, 454 in 2016.

It must be noted, however, that the "quality control: batch potency" category can include vaccine batch testing of veterinary vaccines, for example vaccines for foot and mouth disease, tetanus (Clostridium), rabies, leptospirosis, and various fish and chicken diseases as well as batch testing of human vaccines, for example, vaccines for diphtheria, polio, and pertussis. Some of these tests used other species but some are conducted on mice, so even just looking at the numbers of mice used (possible only for Germany and Great Britain) still does not enable an accurate figure for Bt testing to be obtained.

To date, only 13 EU countries have published data on their use of animals for quality control purposes and it appears that the number of animals reported to have been used under this category can vary quite widely, see Table 1. In 2016 it ranged from 0 animals (Sweden and Finland) to 66,345 (Belgium). Austria (17,604 animals), Belgium (66,345 animals), Hungary (13,252), Poland (10,585) and Spain (39,956) have a relatively high proportion of batch potency tests; for what products these are conducted is not yet known.

It is therefore not currently possible to determine the proportion of batch potency tests that are performed for Bt products

[4] Letter to Cruelty Free International from HPRA (30.03.2018).

[5] Non-technical summary from 2014, volume 13, project 8. https://bit.ly/2Sk7TQx

[6] Cruelty Free International v Secretary of State for the Home Department [2017] EWHC 3295 (Admin) (13.12.2017).

[7] Answer by the Federal Government to the parliamentary question of MP Nicole Maisch, reference number 18/2189 (22.07.2014). https://bit.ly/2FLzicL

[8] Answer by the Federal Government to the parliamentary question of MP Nicole Maisch, reference number 18/4593 (10.04.2015). https://bit.ly/2QqMKa2

[9] e-mail from Merz to Doctors Against Animal Experiments (20.11.2015).

[10] Technically "procedures" are recorded but since in these tests the animals are only used once, animals and procedures can be used interchangeably here.

alone. Clearly, a proportion of the numbers in Great Britain, Ireland, and Germany will be for other types of products. However, the predicted numbers of animals to be used in Bt tests from the non-technical summaries available from Germany and Great Britain suggest the majority (approx. 75%) of the category is indeed made up of this testing. Ireland acknowledge that 95% of the batch potency tests use mice but have not published the non-technical summaries for the Bt testing to enable us to see the exact predicted numbers. However, we recently obtained the project applications for testing to be done in 2018 and it is clear from these that the majority of the category will be the MBA (six licences were obtained that totalled approximately 146,500 mice to be used during 2018).

A conservative estimate of approx. 400,000 animals used in Bt batch testing in the EU each year is probably the best estimate to date. This does not include the number of animals used in the research and development of new Bt products, only the batch testing. It is possible that we are starting to see a decline in numbers of animals used but due to the recent change in reporting requirements from 2014, which have introduced some uncertainty, and an apparent increase in numbers in Ireland in 2016, it is too soon to tell. What is clear, however, is that despite

Tab. 1: Batch potency tests in 2016 (of all EU countries providing the information)
Countries in grey are known to test Bt products under this category.

Country	Total batch potency procedures in 2016 (*mice, if given*)	Total procedures in 2016	% batch potency tests of total procedures
Austria	(*16,720*) 17,604	236,459	7
Belgium	66,345	535,829	12
Croatia	194	21,901	1
Denmark	(*3,491*) 5,486	273,224	2
Finland	0	105,615	0
Germany	(*78,932*) 155,904	2,189,261	7
Great Britain	(*130,973*) 145,190	3,936,723	4
Hungary	13,252	170,075	8
Ireland	167,549	226,934	74
Poland	(*6,758*) 10,585	185,699	6
Slovakia	235	12,855	2
Spain	39,956	917,986	4
Sweden	0	350,664	0

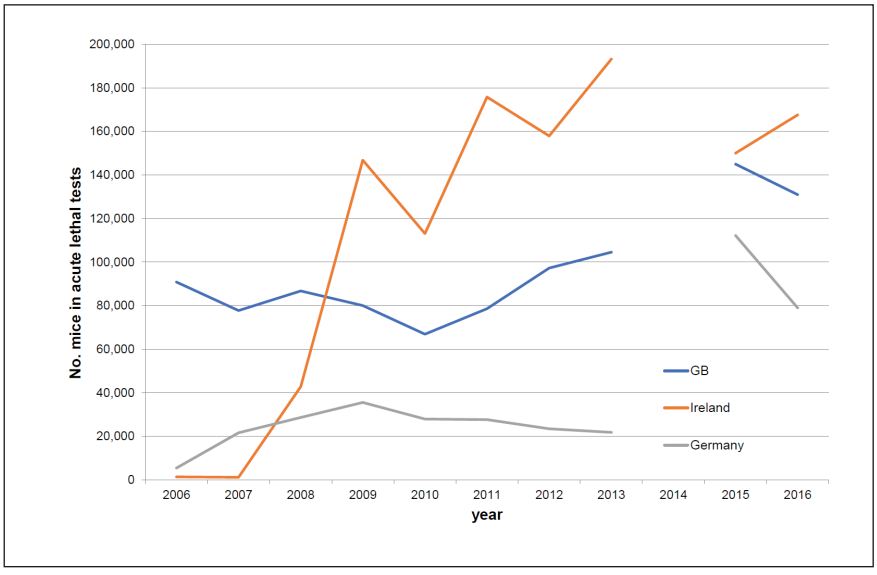

Fig. 1: Number of mice used in acute lethal toxicity tests in Great Britain, Ireland and Germany between 2006 and 2013, and in batch potency tests in 2015 and 2016
Data from official statistical reports from the Home Office, Health Products Regulatory Agency, and The Federal Ministry of Food and Agriculture, respectively; data from 2014 unreliable due to change in reporting; numbers for Ireland in 2015/16 are not given at the species level.

the presence of alternatives to the MBA (see below) testing is still at significant levels.

The scale of the use of animals used for batch testing of Bt products could be due to increased production of the existing BtA products on the EU market as well as new products being authorised (see Section 5). Table 2 lists the Bt products available on the EU market; others such as Botulax™ made by Hugel Inc. from South Korea, BTXA™ made by Hughs from China and Neuronox™ made by Medytox from South Korea may soon become available. One growth area is in the production of BtB products, which have the same MBA requirements as BtA products. BtB is currently licensed in the EU for the treatment of torticollis (cervical dystonia) only. Its effects appear to wear off more quickly than those of BtA. However, according to Bottrill (2003), patients may develop neutralizing antibodies against BtA after several applications and the aesthetic use of BtB may therefore increase as more people become tolerant to BtA.

Until now there has never been an accurate assessment of the total number of animals used in Bt batch testing across Europe. Bitz (2010) estimated 600,000 mice world-wide per year based on an extrapolation of the number of mice used on behalf of Ipsen in the UK, Merz in Germany, and Allergan's annual sales.

Tab. 2: Botulinum toxin products for sale in the EU and the status of replacement of the mouse bioassay

Trade name*	Bt type	EU market authorisation uses	Marketing authorization holder in EU	Potentially testing	Progress with replacement of batch test
Botox	A	Limb spasticity, facial and eye spasms, overactive bladder, excessive sweating, vertical and horizontal frown lines, crow's feet[a]; First authorised in UK 1994	Allergan (Ireland)	Ireland	CBA first approved in USA, Canada, and Switzerland (2011), then all (CBA) licensed in 27 EU (2013)[b] (70% of their sales)[c]. Also approved in Australia[d]. Goal was to reduce testing by 95% in 3 years from 2012[e].
Vistabel	A	Vertical and horizontal frown lines, crow's feet[f]; First authorised in UK 2006	Allergan (Ireland)	Ireland	As Botox
Dysport	A	Limb spasticity, spasms, cervical dystonia, excessive sweating; First authorised in UK 1990	Ipsen (UK)	UK, possibly elsewhere	Began development of CBA in 2011[g], positive approval in EU and Switzerland in 2018 (Ipsen, 2018).
Azzalure	A	Vertical frown lines, crow's feet[h]; First authorised in UK 2009	Galderma/ Nestle Skin Health, Switzerland (distributer, Ipsen product)	Was UK, now unknown	As Dysport
Xeomin	A	Limb spasticity, eye spasms, cervical dystonia; First authorised in UK 2007	Merz (Germany)	Germany	Began development of CBA in 2011. Approved of CBA by FDA in early 2015[i]/late 2015 Europe. To date CBA has been approved in 36 countries[j].
Bocouture	A	Vertical and horizontal frown lines, crow's feet[k]; First authorised in UK 2010	Merz (Germany)	Germany	As Xeomin
NeuroBloc	B	Cervical dystonia; First authorised 2001 EMA	Eisai (Japan) (distributer, Elan/ Perrigo, Ireland product)	Germany	In the process of validating the CBA[l].
Botulax	A	Expected to enter 2018[m]	Croma Pharma, Austria (distributer, Hugel Inc., South Korea, producer	unknown	Working on the CBA, EMA have informed them they must use it to gain regulatory approval[n].

*Different strengths of these products are licensed individually

[a]When the severity of these lines has an important psychological impact on the patient; [b]Letter to GAIA (23.06.2014); [c]Letter to ECEAE (24.05.2013); [d]e-mail to ECEAE 17 April 2014; [e]Allergan receives positive opinions for first-of-its-kind, fully *in vitro*, cell-based assay for Botox® and Vistabel®. Press release, Allergan Inc., Marlow, UK (23.02.2012); [f]When the severity of these lines has an important psychological impact on the patient; [g]Ipsen communication Feb 2013; [h]In adult patients under 65 years, when the severity of these lines has an important psychological impact on the patient; [i]e-mail to ECEAE (13.03.2015); [j]e-mail to ECEAE (16.02.2017); [k]in adult patients under 65 years, when the severity of these lines has an important psychological impact on the patient; [l]e-mail to ECEAE (02.09.2015); [m]http://www.theinvestor.co.kr/view.php?ud=20170518000934; [n]e-mail to ECEAE (26.04.2016)

Pickett (2011), an Ipsen employee, criticised the estimate but did not offer a better one. Given our estimate of 400,000 mice per year in Europe, the global figure, which includes the large markets in America and Asia, is likely to be many times higher and indeed 600,000 may be an underestimate.

5 Aesthetic purposes

BtA products are licensed for the medical treatment of neurological disorders (e.g., focal spasticity, including limb spasticity caused by cerebral palsy or stroke, symptomatic relief of blepharospasm (eyelid spasm), hemifacial spasm and idiopathic cervical dystonia or torticollis (painful neck contractions), prevention of chronic migraine, bladder disorders (overactive bladder), and skin disorders (e.g., excessive sweating). The Bt is usually injected near the affected area and helps relax the muscle, providing temporary symptomatic relief.

The temporary relaxing of frown lines and wrinkles is also listed as a licensed use under skin disorders for some BtA products (see Tab. 2), but only if the lines have an "important psychological impact on the patient". For some products this is also only if the patient is under 65. It is widely considered that use of the product to change appearance, irrespective of any psychological impact, is responsible for a significant proportion of the sales of all BtA products (Bottrill, 2003). Medical practitioners routinely prescribe the products "off-label" (Bottrill, 2003;[11];[12]). This includes where the customer simply wishes to retard manifestation of the ageing process. Crucially, medical practitioners are permitted to administer Bt "off-label" to address the "patient's [subjective] needs"[11], whatever these are: psychological impact from frown lines or another perceived undesirable physiological trait does not have to be shown. Beauty clinics are able to offer botox products to their customers provided they have been prescribed by a medical practitioner[1] and indeed this practice is now widespread across Europe.

Whilst it is widely considered that this "aesthetic" use of all BtA products is significant (Bottrill, 2003), most of the manufacturers are secretive about the split of their sales between aesthetic and medical purposes. Allergan are more transparent; they reported that total global sales of their aesthetic product, Botox® Cosmetics ($1,369.2 million) was 43% of all Botox® products in 2017.[13] Given that, in the EU at least, Botox® (and not just Botox® Cosmetics (or Vistabel® as it is called in the EU)) can and is widely prescribed off-label for aesthetic purposes, then this does suggest that for Allergan's Bt products, aesthetic use constitutes more than 50% of its sales. According to Ipsen's 2016 annual report, sales of Dysport® were €286.7 million[14]. Merz reported that their Speciality Neurology area ("driven by" the botulinum neurotoxin Xeomin®) and Aesthetics (which includes Bocouture®) contributed 63% of a total product revenue of €1,023.2 million in the fiscal year 2016/17[15]. The total annual, worldwide sales of Bt products, therefore, for both medical and aesthetic purposes is in excess of €3,629 million (3.6 billion) for just the three manufacturers selling into the EU (Ipsen, Merz, and Allergan).

The scale of use of BtA products for aesthetic purposes is of concern to animal protection groups. It seems to us to be counter-intuitive to have a European ban on testing on animals for cosmetics purposes and yet turn a blind eye to the testing of Bt products used for the same purpose. Irrespective of a specific ban on testing cosmetics, if the harm:benefit assessment required under Directive 2010/63 is done properly, in our view, testing batches intended for aesthetic purposes should not be permitted. According to Directive 2010/63, all projects must undergo a harm:benefit assessment, whereby the harms to the animals are weighed against the benefits to humans (Article 38). Testing of batches destined to end up in beauty clinics to be used for aesthetic purposes should not pass the harm:benefit assessment in our view. This is because the purpose of testing is vanity and each test causes severe suffering to hundreds of mice. In our opinion, this is probably one of the strongest examples of where the harm:benefit assessment, if it is to mean anything in practice, should be preventing the animal testing of Bt products for aesthetic purposes.

Cruelty Free International has brought two Judicial Review proceedings against the UK Government regarding their authorisation of the MBA in the UK. The cases have focused on what the condition placed on the Wickham licence that the testing must only be "for medicinal products" (acknowledged to mean "purposes") actually means and how the government is enforcing it. The outcome of both cases has been unsatisfactory in the opinion of Cruelty Free International. Initially the government claimed that it was up to the medicines regulator to ensure that the products were not being used for cosmetic purposes. Later, it conceded that it had a duty to take reasonable steps to satisfy itself that batches of Bt to be tested under licence carried a marketing authorisation as a medicinal product and were to be used for medicinal purposes.[16] However, following the case, it admitted that it only checked that there was a marketing authorisation in place for each batch of Bt tested on animals and did not ask for more details about actual end-use. The Irish Health Products Regulatory Agency (HPRA) has also recently confirmed that it simply asks if there is a market authorisation for each product.[4]

In the second case, Cruelty Free International argued that if the government only asked if the product had a market authorisation then the condition on the licence meant nothing in practice (since all Bt products in practice have an authorisation because of their

11 General Medical Council. Information to help you if you want a cosmetic procedure. https://bit.ly/2TMLhcO

12 British Broadcasting Corporation. Botox Bust: How Some Doctors Are Breaking the Rules (08.03.2018). https://bbc.in/2KKNeTf

13 Allergan. Annual Report Pursuant to Section 13 or 15(d) of the Securities Exchange Act of 1934. For the fiscal year ended December 31, 2017. https://bit.ly/2FUBij0

14 Ipsen (2016). Building Bridges for Patient Care. https://bit.ly/2rb4JCW

15 Merz. Press Release: Corporate Fiscal Year with Ups and Downs. Frankfurt, Germany, 24.10.2017. https://www.merz.com/blog/news/fiscal-year-ups-downs/

16 Letter from the Home Office to BUAV (17.08.2011).

genuine medical indications) and the general public was being misled. The government reiterated the difficulties in distinguishing genuine vanity uses of the product from medical ones. Part of the difficulty, in its opinion, was that it does not know which batches are going to end up in beauty clinics or hospitals, nor the specific reasons why a practitioner might prescribe the product to a patient. The judge found in favour of Cruelty Free International with regard to the interpretation of the licence limitation – no testing could be done if any part of a batch was intended to be used for vanity purposes – but was content that the government was doing as much as they could to enforce the limitation.[17] In their "Response To Claimant's Amended Grounds" the Home Office did point out that from January 2017 Azzalure® (the aesthetic version of Dysport®) was no longer being tested in the UK.

Indeed, following analysis of a series of written parliamentary questions in 2003, Balls (2003) had identified that the UK government were permitting testing of Bt products as if it were all going to be used for medical purposes. It seems therefore that this has been governmental practice for many years and that the British High Court is defeatist about being able to prevent animal-tested Bt being used in beauty clinics.

In fact, the British Government actually receives royalties from the sale of Dysport®. Dysport® used to be made at Porton Down in Wiltshire (the government defence establishment recently in the news following the poisoning of Sergei and Iulia Skripal in Salisbury) by the Centre for Advanced Microbiology and Research (CAMR) and then marketed by Ipsen. Indeed, the word "Dysport" is an amalgam of "Porton Down" and "dystonia", the main medical condition for which the product is used. CAMR later changed its name to the Health Protection Agency (HPA). The HPA's role was to safeguard public health. The HPA in turn sold the right to make Dysport® to Ipsen but was entitled to royalties from sales of the product. The HPA has now been subsumed into Public Health England (PHE). PHE's accounts for 2016/17 show that it received nearly £34 million from receipts on royalties, mostly from Dysport®.[18] In 2009 the HPA merged with National Institute for Biological Standards and Control (NIBSC), another government agency and the laboratory which at the behest of the UK pharmaceutical regulator test batches of Dysport® that have already been tested at Wickham. In short, the British Government profits from the sale of Dysport®. This could appear to contribute to why the Home Office is reluctant to do anything that would prejudice sales.

6 Replacement of the animal test

The first alternatives to the MBA were developed by NIBSC, which is tasked with double-checking the LD50 results for Bt batch tests done in the UK. Initially the laboratory was repeating the LD50, but it quickly developed two alternative methods. Firstly in 1996 it developed a refinement method, the limb paralysis method (Sesardic et al., 1996), in which the mice are injected in their hind leg (rather than abdomen) and the extent of bulging of the leg is assessed after 24 or 48 hours. Then in 1997 they developed an *in vitro* test, the endopeptidase method (Ekong et al., 1997). This assay is an *in chemico* test that relies on mouse derived antibodies to measure the degradation of the SNAP-25 protein required for the release of neurotransmitters from the axon endings. Since 1997, the NIBSC have used the endopeptidase method, rather than the limb paralysis method or the original MBA, to check the veracity of the manufacturers' claims about the potency of the Bt batches released in the UK (Sesardic and Des Gaines, 2007).

In 2005, PhEur monograph 2113 was updated to state 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 1 of the following:
1. Endopeptidase assay in vitro
2. Ex vivo *assay using the mouse phrenic nerve diaphragm*
3. Mouse bioassay using paralysis as the endpoint
For these other methods, the potency is calculated with respect to a suitable reference preparation calibrated in mouse LD50 units.

This added a third alternative to the mix (the *ex vivo* assay) and removed, in part, a regulatory block to their use. According to the European Commission, under Article 13(1) of the Directive 2010/63, PhEur methods are "recognised under EU legislation" and therefore any alternative methods listed in these monographs must be used instead of the animal test they replace. For methods such as the ones given in monograph 2113, where product specific validation is still needed, according to the Commission, companies should undertake this in "reasonable time"[19]. Although the methods listed in monograph 2113 were not, and are still not, given monographs of their own outlining how they should be conducted, since 2005 manufacturers have been able to use them as an alternative to the MBA provided that they validate them for their product. Companies could do this by submitting an update to their marketing authorisation to the relevant medicines regulator who would verify if the validation is acceptable and that the batches can be released using the new method instead of the MBA.

Publicly, however, it does not appear that much progress had been made by the Bt manufacturers to take up these methods since they were published in 1996/7. A number of workshops (two in 2006 hosted by EDQM and ICCVAM/ECVAM (reported by NIH, 2008) and one hosted by ZEBET in 2009 (Adler et al.,

17 Cruelty Free International v Secretary of State for the Home Department [2017] EWHC 3295 (Admin) (13.12.2017).
18 Public Health England (2017). Public Health England Annual Report and Accounts: 2016 to 2017. (18.07.2017).
19 EC (2012). Q and A document. http://ec.europa.eu/environment/chemicals/lab_animals/pdf/qa.pdf, p. 16-17.

2010) attempted to bring the manufacturers together to progress the alternatives, but it is not clear how influential they were.

Ipsen (and presumably the other manufacturers) did not consider the endopeptidase assay an adequate full replacement to the MBA because it only measures one property of botulinum toxin activity.[20] They did however state in 2016 that they were using the assay to reduce their animal use by 25%.[21] Adler et al. (2010) reported that EU regulators might be willing to accept the endopeptidase assay for demonstration of stability for at least every second control test per year, which could account for this reduction. Wickham apparently also attempted to validate the mouse limb paralysis assay but presumably unsuccessfully (Sesardic and Des Gaines, 2007;[2]). The *ex vivo* assay using the mouse phrenic nerve-diaphragm was developed at the German Medical School in Hannover, commissioned by manufacturer Merz, but it appears that the validation that began in 2009 (Adler et al., 2010) may also not have been successful.

The situation dramatically changed however with the surprise announcement by Allergan in 2011 that they had developed a cell-based assay (CBA) as a complete replacement for the batch potency test.[22] Allergan's assay was based on a human neuroblastoma cell line and a sandwich ELISA that measures the levels of cleaved SNAP-25. Allergan claimed that the assay had come at a cost of $65 million and a decade of research.[23] Allergan obtained approval to use the CBA for their Bt products sold in the US in 2011 and in the EU in 2012.[24] In correspondence with us, they have confirmed that all of their product released onto the EU market is now tested using the CBA, although they still need to test the reference standard using the MBA (against which the CBA is compared for each batch) and also for stability purposes.[23]

The number of animals used for batch potency testing in Ireland, however, has not decreased and indeed has risen since 2011, see Figure 1. In correspondence, the HPRA indicated that other manufacturers were using the Irish laboratory(ies).[4] Perhaps Ireland has become a hub for testing of other manufacturers and the decrease in animal numbers as a result of Allergan's efforts is being masked.

Short of donating or selling the assay to their competitors, Allergan's scientists did describe the assay in a paper published in 2012 (Fernandez-Salas et al., 2012). According to the authors the problems they had to overcome were generating a sensitive enough cell line (primary neuronal cells were not a viable option) as well as a detection assay that was robust, sensitive, and amenable to validation, such as an ELISA. Furthermore, a specific monoclonal antibody to SNAP-25 needed to be created. Details of this were published in a subsequent paper (Rhéaume et al., 2015). Whilst scientists working for their competitors complained that Allergan had not made the assay freely available (Pickett, 2012), in correspondence with us in 2013 Allergan said that "any competent cell biologist would be able to work out the assay".[24]

Nonetheless implementation of the CBA by other companies was not immediate. However, the delay seems to have been mostly during the validation for their specific products rather than the development of a CBA itself. Ipsen and Merz established an agreement to work together on the development of a CBA in 2011[21] and then separated a year or so later to validate it for their own products. Merz successfully applied to use the CBA in batch testing for products sold in the EU in 2015,[25] but Ipsen only in August 2018.[26] Eisai, who distribute BtB product NeuroBloc™, are also validating the CBA but have not yet applied to obtain approval to use it.[27]

The pace by which the manufacturers after Allergan have developed and, in particular, validated the CBA for their products is disappointing. Is six years a "reasonable time" according to the Commission's guidance? It is difficult for outsiders to establish whether the pace is slow because of genuine scientific issues or whether companies have simply not invested enough in the process. The malaise of both the UK regulator and the contract testing facility in the UK is clearly evident in the Home Office's own report,[2] in which one inspector noted in 2008 that Wickham had not yet validated the refinement limb paralysis method some 12 years after it was developed by NIBSC:

> *"Mr Z is keen to progress these and the move towards the lower severity flaccid paralysis assay but has simply been too busy to do anything about it." (April 2008)*

It is possible that there are scientific issues with the validation of the CBA, although it is not clear if that is to do with the CBA or the variability of the MBA, which is widely acknowledged. The Irish HPRA have indicated that the regulators may be requiring additional MBA tests "as a back-up" due to a lack of confidence in the alternative.[4] It is particularly worrying to us that in Ireland more mice may be being used to validate the same MBA with "humane" endpoints, when the use of a complete replacement is so close to being achieved.

Unfortunately, the batch release requirements for each Bt product are not published by the drug regulators, so it is difficult for patients to see which products are now potentially "cruelty free". Instead the public are reliant on statements from the manufacturers or from animal protection groups who are asking these questions. What is clear from our correspondence with the manufacturers, however, is that no manufacturer has been able to replace the MBA in full and all are still conducting it to test the

[20] Ipsen Statement to the BUAV on the Use of the Snap-25 Endopeptidase Assay as a Replacement for the Mouse LD50 Potency Method (23.11.2009).

[21] Ipsen communication. IPSEN'S Commitment to Developing an Alternative to Animal Testing for its Botulinum Toxin. Sept 2016. http://www.ipsen.com/wp-content/uploads/2016/09/Animal-testing-2016-09-08.pdf

[22] Allergan. Press Release. Allergan Receives FDA Approval for First-of-Its-Kind, Fully In Vitro, Cell-Based Assay for Botox® and Botox® Cosmetic (OnabotulinumtoxinA) (24.06.2011). https://bit.ly/2RpqiLP

[23] Letter from Allergan to Cruelty Free International (07.01.2016).

[24] Letter from Allergan to ECEAE (23.05.2013).

[25] e-mail to ECEAE (13.03.2015).

[26] Ipsen. Press Statement. Ipsen's Cell-Based Assay Receives Approvals in the E.U. and Switzerland for its Botulinum Toxin. (27.08.2018). https://bit.ly/2KDX5Kf

[27] e-mails from Eisai to ECEAE (02.09.2015 and 15.11.2017).

bulk Bt (and possibly for stability purposes also). This is because the MBA was never standardised between the products. Each manufacturer measures the strength of their Bt product in mouse LD50 units that are not interchangeable between them (Sesardic et al., 1994). Every time a batch is tested, the MBA result is compared to that of a reference batch for that product that was tested previously. This is also the case for the testing needed for stability and the bulk preparation. It is the intention of all manufacturers to replace the MBA in full using the CBA. Allergan may be closest to achieving this,[24] but have not yet done so.

In 2012, PhEur monograph 2113 was updated to more strongly encourage the validation of alternatives:

> "*The LD50 is associated with severe suffering of animals and manufacturers are strongly encouraged to develop and validate assays that will reduce the number of animals used or refine or replace the test procedure with the goal of promoting animal welfare".*

The list of potential alternative assays however became vaguer:

> *"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, for example mouse bioassays using paralysis as the end-point,* ex vivo *assays using mouse phrenic nerve diaphragm, endopeptidase assays* in vitro *and cell-based assays. For alternative replacement methods the potency is calculated with respect to a suitable reference preparation calibrated in mouse LD50 units".*

According to Allergan, until all manufacturers have validated their own CBAs the PhEur will continue to list the mouse assay as the reference standard.[23] We are not aware of any impending update to the PhEur. This is unfortunate, since new Bt products produced by other companies are coming into the EU market (see Tab. 2) and until they have also validated the CBA for their products there may be no legal requirement to use it. At the very least a more detailed description of the CBA should now go into the PhEur. Another step that can help move the EU towards eradication of the MBA is for the drug regulators to refuse to licence any new Bt products unless their batch testing is performed using the CBA.

It is of interest to note that academics also appear to be developing cell-based assays for the detection of both BtA or BtB (e.g., Tegenge et al., 2012; Eckle et al., 2014; Pathe-Neuschäfer-Rube et al., 2015; Weingart and Loessner, 2016), but some of these appear to be one-off, uncompleted projects done in relative isolation from the manufacturers. NIBSC scientists, funded by the NC3Rs, have developed an assay using mouse embryonic stem cells for the detection of BtA (Yadirgi et al., 2017) and one using a neuronal cell-line for BtB (Rust et al., 2017). The German governmental Paul Ehrlich Institute has also developed a binding and cleavage assay (the BoNT/B BINACLE assay) for the detection of BtB (Wild et al., 2016). Whether these government agency projects will have more success at achieving complete replacement, particularly if they are not being developed in conjunction with the manufacturers, is not yet known. All this activity does still suggest there may be some duplication of effort caused by lack of communication and the absence of an organised validation effort.

7 Conclusion

We are now at the cusp of a new era where the three main Botulinum product manufacturers in Europe have validated a cell-based method for their batch testing that traditionally used live mice in a test that causes severe suffering and death. However, our analysis shows that, whilst replacement of the MBA has been progressing since 2011, the use of animals for batch testing has in fact increased across Europe. Approximately 400,000 animals are still being used annually in batch testing of Botulinum products across the EU. The reasons for this are not fully transparent but appear to be a combination of difficulties with validating the alternative methods, the need for full replacement of the reference standard and new manufacturers who are not using the alternative entering the EU market.

Many *in vitro* assays have been developed and several have been validated and received regulatory approval for specific products. However, the actual replacement of this severe assay in practice has been far too slow, in our opinion. National regulators need to increase their pressure on the manufacturers and contract testing facilities performing these tests to ensure that the implementation of the alternative is being expedited.

In our opinion, testing batches of Bt products on animals should not be permitted if the harm:benefit assessment required under Directive 2010/63 is to mean anything in practice – at least those that are intended to be distributed to beauty clinics. Manufacturers must know where their products are being sold and regulators must not allow them or their contract testing facilities to use complex supply chains as an excuse for circumventing the requirements of EU law.

Action is needed on an EU wide regulatory level to ensure that new products are not permitted onto the EU market unless they use the replacement method. Finally, those manufacturers that have developed an alternative method need to work with the European Pharmacopeia to ensure that final replacement is achieved.

References

Adler, S., Bicker, G., Bigalke, H. et al. (2010). The current scientific and legal status of alternative methods to the LD50 test for botulinum neurotoxin potency testing: The report and recommendations of a ZEBET expert meeting. *Altern Lab Anim 38*, 315-330.

Balls, M. (2003). Botulinum toxin testing animals: The questions remain unanswered. *Altern Lab Anim 31*, 611-615.

Balls, M. (2010). Editorial: Replacing the animal testing of botulinum toxin: Time to smooth out the wrinkles. *Altern Lab Anim 38*, 1-2.

Bitz, S. (2010). The botulinum neurotoxin LD50 test – Problems and solutions. *ALTEX 27*, 114-116. doi:10.14573/altex.2010.2.114

Bottrill, K. (2003). Growing old disgracefully: The cosmetic use of botulinum toxin. *Altern Lab Anim 31*, 381-391.

EC – European Commission (2009). Regulation (EC) No 1223/2009 of the European Parliament and of the Council of 30 November 2009 on cosmetic products. *Off J Eur Union L 342*, 59-209 (22.12.2009).

EC (2010). Directive 2010/63/EU of the European Parliament and the Council of 22 Sept. 2010 on the protection of animals used for scientific purposes. *Off J Eur Union L 276*, 33-79 (20.10.2010).

EC (2012). Commission Implementing Decision of 14 November 2012 establishing a common format for the submission of the information pursuant to Directive 2010/63/EU of the European Parliament and of the Council on the protection of animals used for scientific purposes (2012/707/EU).

Eckle, V. S., Drexler, B., Grasshoff, C. et al. (2014). Spinal cord – Skeletal muscle cocultures detect muscle-relaxant action of botulinum neurotoxin A. *ALTEX 31*, 433-440. doi:10.14573/altex.1304291

EDQM – European Directorate for the Quality of Medicines (2011). Monograph 2584 for botulinum toxin type B for injection (last updated July 2011).

EDQM (2012). Botulinum toxin type A for injection. European pharmacopeia monograph number 2113 (last update Jan 2012), current edition volume 9 (in force from July 2017).

Ekong, T. A., Feavers, I. M. and Sesardic, D. (1997). Recombinant SNAP-25 is an effective substrate for Clostridium botulinum type A toxin endopeptidase activity in vitro. *Microbiol 143*, 3337-3347. doi:10.1099/00221287-143-10-3337

Fernández-Salas, E., Wang, J., Molina, Y. et al. (2012). Botulinum neurotoxin serotype A specific cell-based potency assay to replace the mouse bioassay. *PLoS One 7*, e49516. doi:10.1371/journal.pone.0049516

Hartung, T. and Koëter, H. (2008). Food for thought ... on food safety testing. *ALTEX 25*, 3-9. doi:10.14573/altex.2008.1.3

ICH – International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use (1995). Guideline ICH Harmonised Tripartite Guideline Quality of Biotechnological Products: Stability Testing of Biotechnological/Biological Products Q5C (30.11.1995).

ICH (2004). ICH Harmonised Tripartite Guideline Comparability of Biotechnological/Biological Products Subject to Changes in Their Manufacturing Process Q5E (18.11.2004).

NIH – National Institutes of Health (2008). NICEATM-ICCVAM/ECVAM Scientific Workshop on Alternative Methods to Refine, Reduce and Replace the Mouse LD50 Assay for Botulinum Toxin Testing. NIH Publication No 08-6416.

Pathe-Neuschäfer-Rube, A., Neuschäfer-Rube, F., Genz, L. and Püschel, G. P. (2015). Botulinum neurotoxin dose-dependently inhibits release of neurosecretory vesicle-targeted luciferase from neuronal cells. *ALTEX 32*, 297-306. doi:10.14573/altex.1503061

Pickett, A. (2011). Comment on "The botulinum neurotoxin LD50 test – Problems and solutions". *ALTEX 28*, 60-61. doi:10.14573/altex.2011.1.60

Pickett, A. (2012). Comment: The botulinum toxin LD50 potency assay – Another chapter, another mystery. *Altern Lab Anim 40*, 221-223.

Rhéaume, C., Cai, B. B., Wang, J. et al. (2015). A highly specific monoclonal antibody for botulinum neurotoxin type A-cleaved SNAP25. *Toxins 7*, 2354-2370. doi:10.3390/toxins7072354

Rust, A., Doran, C., Hart, R. et al. (2017). A cell line for detection of botulinum neurotoxin type B. *Front Pharmacol 8*, 796. doi:10.3389/fphar.2017.00796

Sesardic, D., Gaines Das, R. E. and Corbel, M. J. (1994). Botulinum toxin how to define biological activity. *J Roy Soc Med 87*, 307.

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. *Phamacol Toxicol 78*, 283-288. doi:10.1111/j.1600-0773.1996.tb01376.x

Sesardic, D., Leung, T. and Gaines Das, R. (2003). Role of standards in assays of botulinum toxins: International collaborative study of three preparations of botulinum type A toxin. *Biologicals 31*, 265-276. doi:10.1016/j.biologicals.2003.08.001

Sesardic, D. and Gaines Das, R. (2007). Alternatives to the LD50 assay for botulinum toxin potency testing: Strategies and progress towards refinement, reduction and replacement. *AATEX 14, Spec Issue*, 581-585.

Straughan, D. W. (2006). Progress in applying the Three Rs to the potency testing of botulinum toxin type A. *Altern Lab Anim 34*, 305-313.

Tegenge, M. A., Böhnel, H., Gessler, F. and Bicker, G. (2012). Neurotransmitter vesicle release from human model neurons (NT2) is sensitive to botulinum toxin A. *Cell Mol Neurobiol 32*, 1021-1029. doi:10.1007/s10571-012-9818-2

Weingart, O. G. and Loessner, M. J. (2016). Nerve cell-mimicking liposomes as biosensor for botulinum neurotoxin complete physiological activity. *Toxicol Appl Pharmacol 313*, 16-23. doi:10.1016/j.taap.2016.10.010

Wild, E., Bonifas, U., Klimek, J. et al. (2016). In vitro potency determination of botulinum neurotoxin B based on its receptor-binding and proteolytic characteristics. *Toxicol In Vitro 34*, 97-104. doi:10.1016/j.tiv.2016.03.011

Yadirgi, G., Stickings, P., Rajagopal, S. et al. (2017). Immuno-detection of cleaved SNAP-25 from differentiated mouse embryonic stem cells provides a sensitive assay for determination of botulinum A toxin and antitoxin potency. *J Immunol Methods 451*, 90-99. doi:10.1016/j.jim.2017.09.007

Conflict of interest

The authors declare that they have no conflicts of interest.

Acknowledgements

The authors would like to thank Cruelty Free International Trust and Ärzte gegen Tierversuche e.V. (Doctors Against Animal Experiments) for supporting the authors during the drafting of this paper.