Usefulness of testing for *Clostridium botulinum* in powdered infant formula and dairy-based ingredients for infant formula

Dated : 17 January 2014 (revision 1)

Background

In early August 2013, a New Zealand dairy company announced that a whey protein concentrate it produces for infant formula, beverages and animal feed was potentially contaminated with *Clostridium botulinum*, leading to recalls of infant formula products. Although there have been no reports of any public health consequences from this incident, it has raised the level of concern for the safety of these products with respect to *C. botulinum*. The International Commission on the Microbiological Specifications for Foods (ICMSF) (www.icmsf.org) has received a number of enquiries from industry and regulatory authorities to provide advice on the appropriateness of testing for *C. botulinum* in these products. In response to these enquiries, ICMSF has prepared this document.

*Clostridium botulinum*

*C. botulinum* is an important foodborne pathogen that causes serious neuroparalytic disease. The organism comprises four physiologically distinct groups capable of producing protein neurotoxins that have similar toxicological activity. Human botulism is caused by types A, B, E, and rarely F and there are rare reports of some strains of non-botulinum clostridia (e.g. *C. butyricum* and *C. barati*) also causing botulism. In 2013, a new neurotoxin, designated type H, was reported in a strain of *C. botulinum* linked to infant botulism (Barash and Arnon, 2013). This strain was also capable of producing type B toxin. *Clostridium botulinum* is an anaerobic, Gram-positive, spore-forming bacterium that is commonly found in soils and aquatic sediments throughout the world. There are four clinically distinct forms of botulism in humans: foodborne botulism (resulting from ingestion of pre-formed toxin in a food); infant botulism (as a result of colonisation of the infant intestine accompanied by toxin production in vivo); adult infectious botulism (similar to infant botulism but occurring in immunocompromised adults); wound botulism (caused by the organism infecting deep wounds and producing toxin in vivo). Due to the serious nature of the neuroparalytic disease caused and unparalleled potency of the toxin, *C. botulinum* remains one of the major targets for control in foods today.

Product use and risks

Due to its potency, there is no tolerance for the presence of neurotoxin or for conditions allowing growth of *C. botulinum* in foods. For foods for which *C. botulinum* is a relevant hazard to consider, industry assures food safety by eliminating the hazard or by controlling its presence or possible outgrowth. The primary factors for control of *C. botulinum* in foods are prevention of growth through reduced pH (<4.6) or reduced water activity (<0.93), storage at chilled temperature (≤ 3°C) or frozen, through the presence of preservatives or through a combination of these or, alternatively, thermal inactivation (121°C for 3 min for the most heat resistant types in low acid foods) to effectively eliminate the organism from products. The main risks for foodborne botulism are associated with inadequate control of one or more of these main factors, or recontamination (e.g. through faulty seals on hermetic packaging) leading to the survival, germination and outgrowth of *C. botulinum* spores. In many countries, foods associated with foodborne botulism tend to be home-prepared, such as traditionally-prepared salted or fermented foods (e.g. fish, cured meats), and foods canned or bottled in the home. In these cases, foods are prepared and then stored under conditions that allow outgrowth of spores that are present in the raw materials. Occasionally, commercially manufactured products can cause foodborne botulism where, again, there is failure in applying the relevant controlling factors, such as inadequate heat treatment, inadequate...
Infant botulism and infant formula

Of the different forms of botulism, infant botulism is the most common form reported in the USA (Shapiro et al., 1998) and the European Union (ECDC, 2011). Infant botulism is caused when neurotoxicogenic *Clostridium* spores are ingested, germinate and colonise the immature gut of the infant and produce neurotoxin. Infants less than one year old are at risk of infant botulism and those less than 6 months old are particularly vulnerable, which likely reflects the inability of the infant’s immature intestinal microbiota to resist colonization by *C. botulinum*. The clinical presentation and management of the disease is discussed by Radsel et al (2013). The identification of foodstuffs associated with cases of infant botulism is difficult due to the generally long incubation time before manifestation of the disease, but the main documented evidence exists for honey as a vehicle (Midura, 1996) and spores of *C. botulinum* have been isolated from honey (Hauschild et al., 1988; Al-Waili et al., 2012). In the UK, a case of infant botulism was presumptively linked to infant formula (Brett et al., 2005) and *C. botulinum* type B spores were reported in both an opened and unopened can. However, the AFLP (Amplified Fragment Length Polymorphism) type of the isolate from the unopened can was different to that of the isolate from the infant. In addition, differences in the PCR profile of *C. botulinum* isolated from the unopened can and the isolate from the infant suffering from infant botulism led Johnson et al. (2005) to conclude that the infant formula powder was not the source of transmission of spores to the infant. With respect to hazardous microorganisms associated with infant formula, FAO/WHO categorized *C. botulinum* as “causality less plausible or not yet demonstrated” because, although having been identified in powdered infant formula, they had not been implicated as causing illness in infants (FAO/WHO, 2004). Consequently, *C. botulinum* is not considered a hazard in the Codex international hygiene standard for infant formula (CAC, 2008). In a recent report on infant botulism by the UK Food Standards Agency (ACMSF, 2006), experts carrying out a risk assessment on infant botulism agreed that, based on available data on the presence of spores of *C. botulinum* in infant food (i.e., broader than just powdered infant formula) the most likely mean value was 0.3 spores per kg, the minimum mean was 0.001 spores per kg, and the maximum mean value was 10 spores per kg. The experts considered that levels of 10 spores per pack of 121 g (hence approximately <1 spore per serving of 13 g of powdered infant formula) would represent only a low risk of disease.

Limitations of testing for *Clostridium botulinum* in food or food ingredients

Detection and enumeration of *C. botulinum* in foods is difficult. This is partly because of the large differences in the microbial ecology of strains of the ‘species’. In fact, *C. botulinum*, as currently defined, includes organisms that are sufficiently genetically different that they could be classified as distinct species (Peck et al., 2010; Fischer et al., 2012), and the species is mostly characterised by the ability to produce botulinum neurotoxins. Other *Clostridium* species that produce botulinum toxins are now also recognised and multiple methods are required to detect all strains of relevance (Fischer et al., 2012). Methods based on enrichment, followed by confirmation by PCR, are generally recommended. An ISO standard (ISO, 2013)
for the molecular detection of clostridia carrying botulinum neurotoxin A, B, E and F genes by PCR has recently been published.

Confirmation of toxin production, using the mouse bioassay, is required in some situations (Solomon and Lilly, 2001), but this is not suited to routine food microbiology laboratories since special security and biosecurity precautions are required (EFSA, 2005). Therefore, only a few, specialised, laboratories internationally are able to do this work. Even though the mouse bioassay is the ‘gold standard’ method for detection of botulinum toxin, it has several drawbacks including non-specific mortality during the course of the test (often due to the response of the animals to endotoxins from Gram-negative bacteria), availability of appropriate controls and materials, long testing time (4 days), and animal ethics issues (AOAC International, 2001). The full mouse bioassay involves 3 stages: toxin screening; assessment of toxin titre; and neutralization with monovalent antitoxins but some procedures can be shortened, or done in parallel, if the main interest is only to prove the presence of botulinum toxin rather than to understand and resolve the epidemiological aspects of cases or outbreaks. Toxicity testing should include heat treated samples (botulinum toxin is inactivated by heat and mice injected with these samples should not die) and following injection of samples, symptoms should be closely monitored over a 48 h period. The sequence and nature of typical symptoms are well described and death of mice without these clinical symptoms is not sufficient evidence that injected material contained botulinum toxin (US FDA, 2001a). If typical symptoms are not observed and mice die after 24 h, or mice die at only very low dilutions, the results should be treated with suspicion. Following death of all mice except those that received the heated preparation, dilutions of the original preparations are used to determine the minimum lethal dose (toxin titre). In the final stage, mice are injected with monovalent antitoxins (types A, B, E and F) prior to challenge with the toxic preparation. If results indicate that toxin has not been neutralised, the test is repeated using monovalent antitoxins to types C and D, together with polyvalent antitoxin pool of types A-F. If results still indicate no neutralisation, this should raise suspicions.

While testing for \textit{C. botulinum} may be required in some circumstances, e.g., to assess a suspect lot, as part of a survey for a risk assessment, etc., routine testing of foods for presence of \textit{C. botulinum} is not usually recommended. This consensus view is due both to the difficulty of testing for \textit{C. botulinum} and also the very low levels of contamination of foods with this microorganism. For example, honey involved in cases of infant botulism has been found to contain as few as 5 - 80 spores per gram (Arnon et al., 1979; Midura et al., 1979). In other studies, levels of \textit{C. botulinum} in honey that were not associated with infant botulism were not higher than 7 per 25 g (ACMSF, 2006). Notably, an oft-cited "infectious dose" for infant botulism is 10 - 100 spores (Arnon, 1992).

\textbf{Testing for Clostridium botulinum in powdered infant formula}

Whilst \textit{C. botulinum} generally presents a low level of risk in powdered dairy products, spores of the organism have incidentally been found.

In a survey of dairy powders in France, levels in the range 2 - 5 per kg were reported (Carlin et al., 2004) but were not known to be associated with illness. Barash et al. (2010) found very low levels of clostridial spores (1.1 to > 23/100g) in powdered formulae but \textit{C. botulinum} was not detected in the samples analysed. As noted earlier, ACMSF (2006) considered that typical maximum levels in milk powders are in the order of 1 cell/spore per 100g.

To design a suitable sampling plan for assessing product contamination, it is necessary to decide on the required performance of the sampling plan, i.e. for the microbiological criterion, what level of spores will distinguish ‘acceptable’ from ‘unacceptable’ product lots. For typical infant formula, 12 - 14 g is made up in 100 ml of water, which is a common serving size of ‘reconstituted milk’. From the reports cited above, and for illustrative purposes only, it could be assumed that 1 spore per 100 g of powdered formula or (for instance) its main protein
component whey powder, has a low probability of causing infant botulism. A suitable sampling plan should provide confidence that if this level is exceeded it will be reliably detected.

The performance of sampling plans can be approximated by the binomial distribution function (Jarvis, 2008). A more sophisticated analysis has been presented by van Schothorst et al. (2009).

The binomial function permits a first estimate of the number of samples that must be tested and shown not to be contaminated, to provide assurance that the batch is not more contaminated than the microbiological criterion allows, assuming that the contamination is homogenous throughout the batch. As there is some element of chance involved in any sampling scheme, the number of samples required will increase when higher levels of confidence are required. The number of samples will also depend on the size of the samples analysed.

Based on these considerations and assumptions, Table 1 exemplifies the number of samples that should be tested (at selected levels of confidence, and for different samples sizes), and shown not to be contaminated, to be confident that the level of contamination in the lot is less than one spore per 100 grams of infant formula powder.

Table 1: Numbers of samples (of specified size) that must be tested from a lot (with homogenous contamination) and shown not to be contaminated to ensure that the average contamination with C. botulinum in the lot is less than 1 CFU per 100 g. The table shows the effect of sample size and rigour of testing (expressed as confidence that the result is accurate).

<table>
<thead>
<tr>
<th>sample size</th>
<th>confidence in result of testing</th>
<th>number of samples required</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 gram</td>
<td>95%</td>
<td>299</td>
</tr>
<tr>
<td>10 gram</td>
<td>95%</td>
<td>29</td>
</tr>
<tr>
<td>25 gram</td>
<td>95%</td>
<td>11</td>
</tr>
<tr>
<td>25 gram</td>
<td>99%</td>
<td>17</td>
</tr>
<tr>
<td>25 gram</td>
<td>90%</td>
<td>9</td>
</tr>
</tbody>
</table>

However, the level of contamination is unlikely to be homogenous throughout the batch, and it is more likely to be log-normally distributed. If it is assumed that the variability in logCFU contamination levels within a batch has a standard deviation of 0.8 (log[CFU/g]), a batch that satisfies the criterion of ≤1 spore/100g in 99% of cases, would have an average arithmetic concentration throughout the batch of < 1 spore per ~1.3 kg of powdered infant formula or average log concentration of ≤-3.86 logCFU/g. (see Appendix 1). To demonstrate (with 95% confidence) that the batch satisfies the criterion requires > 180 samples of 25 g to be tested, and none to be found positive. This is an impractically large number of samples.

**Current practice and testing for sulphite reducing clostridial spores**

Given the methodological limitations described above, routine microbiological testing for C. botulinum to ensure food safety is not recommended. Instead, tests for sulphite-reducing clostridia (Fischer et al., 2012) have been proposed to identify product lots produced under conditions that could be associated with increased risk of C. botulinum contamination.

The detection of sulphite-reducing clostridia (SRC) has been used extensively in both water and food microbiology, but its uses are based on different assumptions and protocols. In water microbiology, SRC are largely used as either a faecal and/or process control indicator. The
presence of SRC in the absence of other faecal indicators is indicative of remote or intermittent contamination, usually *Clostridium perfringens*, in the distant past (Sartory et al., 1993; AS/NZS, 2000; Robles et al., 2000; Adcock and Saint, 2001; Marcheggiani et al., 2008). SRC use as a process control indicator is based on the greater resistance of *Clostridium* species to water treatments than viruses and protozoa (Adcock and Saint, 2001).

In food microbiology, use of SRC as a microbial indicator has been discussed for more than 50 years, but its routine use is limited to a relatively small number of foods (Mossel, 1959; Romagnoli and Brezzi, 1960; Prevost et al., 2013). While originally proposed as an index test for pathogenic *Clostridium* (Mossel, 1959; Weenk et al., 1991), this goal has been elusive due to different sources of *Clostridium* in foods (e.g., *C. perfringens* related to faecal contamination; *C. botulinum* related to soil contamination). Its current application in foods is either as an indication of faecal contamination (*C. perfringens*) and/or as an indicator of sanitation/process control related to potential growth and survival of anaerobic spore-forming bacteria. Standard methods for SRC are available, for instance, through ISO 15213 (ISO, 2003) and US FDA (US FDA, 2001). There are differences in ‘official’ SRC methods currently used (e.g., FDA BAM vs. ISO) on whether to heat food samples to eliminate vegetative cells. With respect to using SRC as a hygiene indicator, it is important to consider both vegetative cells and spores which may be present simultaneously in the sample: the ratio will depend on the length of process run times and conditions in the processing lines. It is therefore recommended not to apply a heat treatment. In this case, simultaneous plating of the sample on microbiological media suitable for the detection of Enterobacteriaceae may be warranted to avoid false-positive results due to the presence of non-spore forming bacteria (e.g., *Proteus*) that can reduce sulphite. However, even with a heat treatment to eliminate vegetative cells, confirmation may be required since there are facultative anaerobic spore-formers (e.g., *Bacillus licheniformis*) that can also grow and reduce sulphite, depending on the medium employed (Fischer et al., 2012).

While SRC are not widely used in foods as a hygiene indicator, a review of their key properties in relation to those expected of an effective indicator organism (Cordier and ICMSF, 2013) shows that SRC would fulfil most of them, though there is contention about points 1 and 9:

1. History of concomitant presence of indicator(s) and associated pathogen or its toxin;
2. Presence usually at higher levels than associated pathogen;
3. Presence indicative of an increased risk of faulty practices or faulty processes;
4. Survival or stability similar to or greater than that of the target pathogen;
5. Growth behaviour similar to or greater than that of the target pathogen;
6. Easily detectable and/or quantifiable;
7. Identifiable characteristics need to be stable;
8. Methods for indicator organisms need to fulfil the same requirements as the one for the pathogen, i.e., they need to be reliable and validated; in addition, they should be more rapid and less expensive;
9. Quantitative results should show a correlation between indicator concentration and level of the pathogen;
10. Results need to be applicable to process control;
11. Analyst health is not at risk.

Nonetheless, testing for SRC is, on the whole, deemed an appropriate indicator of process hygiene with respect to the control of anaerobic sporeformers. For example, limits for SRC currently used in trade range from a regulatory limit in the Russian Federation of 25 – 100 cfu/g for import of dairy derivatives, including milk proteins (Russian Federation, 2008) and advisory maximum levels of 10-25 cfu/g in the USA (US Dairy Export Council, 2013).
Use of tightened sampling plans following loss of process control

There appear to be no standardized sampling plans for the analysis of foods, with the exception of bottled natural mineral waters (CAC, 1985). However, Table 2 compares the relationship between the number of samples that must be analyzed via a 2-class attribute plan, using a plating method to ensure different levels of compliance and assuming a high degree of confidence (e.g. 95%) that the sampling plan would detect (and reject) such a lot.

Table 2 : Relationship between the numbers of samples from a batch that are analysed and found to contain <100 CFU/g and the highest proportion of samples from that lot that could be expected (95% confidence) to exceed 100 CFU/g.

<table>
<thead>
<tr>
<th>Number of samples analysed and shown not to contain &gt;100 CFU/g</th>
<th>Maximum proportion (%) of samples in the lot that could be expected to exceed 100 CFU/g</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>95</td>
</tr>
<tr>
<td>3</td>
<td>63</td>
</tr>
<tr>
<td>5</td>
<td>45</td>
</tr>
<tr>
<td>10</td>
<td>26</td>
</tr>
<tr>
<td>15</td>
<td>18</td>
</tr>
<tr>
<td>30</td>
<td>10</td>
</tr>
<tr>
<td>60</td>
<td>5</td>
</tr>
</tbody>
</table>

While statistical analysis of the two class sampling plan results enable us to estimate the proportion of samples that could exceed the microbiological criterion, the proportion that could exceed even higher levels will depend on the variability in the lot. For larger standard deviations, the proportion of samples at very high contamination levels would be expected to be higher (see van Schothorst et al, 2009). When using an enumeration method, as in the example given, the use of variables sampling plans (ICMSF, 2002; Jarvis, 2008) can reduce the number of samples required to evaluate the mean contamination levels, and acceptability, of the batch.

If a production lot exceeds the established limit selected, increased sampling of batches is advisable. The level of increased sampling is a management decision but could consider, for example, an increase in sampling by a factor of 5 – 10 until SRC values return to the normal baseline for the product and/or the source of the contamination has been determined and rectified. Control chart plotting of SRC may permit corrective actions to be taken before the rejection criterion is exceeded (ICMSF, 2002).

Recommendations

Given that C. botulinum is not considered a hazard in infant formula (CAC, 2008), the important constraints related to the analysis of C. botulinum, i.e. probably being heterogeneously distributed and at very low levels in powdered dairy products, the ICMSF does not recommend routine testing for this pathogen. End-product testing should be limited to samples investigated during an outbreak to determine its source, thus allowing regulatory authorities to take appropriate measures to protect consumers, such as the recall of the incriminated product.

It is recognized that there is no direct mathematical correlation between levels of spores of sulphite-reducing clostridia (SRC) and those of C. botulinum. The ICMSF nevertheless concludes that testing for SRC is appropriate as an indicator of process hygiene with respect to
the control of anaerobic sporeformers. Therefore, the level of SRC can be used to determine adherence to Good Hygiene and Manufacturing Practices during the production of dehydrated dairy ingredients or products such as powdered infant formulae. SRC levels exceeding a limit of \( m = 100 \text{ cfu/g} \), for example, would point to conditions potentially conducive to the multiplication of anaerobic clostridia in the processing lines, or some source of external contamination. Considering the characteristics and behaviour of various types of clostridia, such levels would then also point to conditions possibly supporting multiplication of \( C. \text{botulinum} \), when present, beyond normal low levels. In contrast, SRC levels below this limit would confirm the effectiveness of the established hygiene control measures, and hence a negligible risk for \( C. \text{botulinum} \).

Lots within the established limit could be used as intended. If a lot exceeded this limit, an investigation into possible causes would be warranted together with actions to prevent further occurrence of unacceptable levels. With regard to the affected lot, this should be subject to appropriate actions to mitigate the potential risk. Such actions could range from the rejection of the lot in question, its use in products targeted towards non-sensitive consumers or for products manufactured with a kill-step designed to eliminate clostridial spores.
Appendix 1. Statistical treatments of CFU vs LogCFU

Microorganisms grow (or decline) exponentially over time, and levels in foods can vary over many orders of magnitude. Therefore, microbial numbers are often expressed as their log-transformed values. This has the advantage that numbers are more ‘manageable’, e.g., 7.4 log CFU is easier to express and interpret than 25118864 CFU. Additionally, errors in determination of microbial counts are often relative to their numerical (“absolute”) concentration due to the enumeration methods used, which usually involve serial dilutions. A log transformation makes the errors more consistent in scale across very large differences in absolute numbers. This also has advantages for data analyses because most currently used statistics are only appropriate if the data conform to a ‘Normal’ distribution. The best available data suggest that the logarithm of microbial counts/concentrations in foods are consistent with the Normal distribution.

Accordingly, log transformation of numbers/concentrations of microorganisms in foods is often used and is usually effective and appropriate in describing, interpreting and analysing experimental results. This includes the design and interpretation of sampling plans that require that microorganisms in foods are “log-Normally” distributed (i.e. that the logCFU values are Normally distributed).

Conversely, the impact of organisms, including the public health risk from pathogens in foods, is most often related to the absolute number of organisms ingested. For example, at doses where probabilities of infection/illness are below 10% (i.e., below the ID10), most dose-response models for infection/illness predict that the probability of illness/infection is directly proportional to the dose ingested. Therefore, in the interpretation of the public health impact the absolute number of organisms is most relevant.

A consequence of the logarithmic transformation of data is that the average of log transformed values differs from the average of the untransformed values. For example, the average contamination level of three samples with 100, 1000 and 10000 CFU/g present is 11100 CFU/g /3 = 3700 CFU/g. The mean log contamination level is (2+3+4 logCFU)/3 = 3 logCFU/g. When converted to absolute numbers, this suggests a mean contamination level of only 1000 CFU/g. Importantly, in most situations, the risk from 3700 CFU/g is nearly four-fold higher than the risk from 1000 CFU/g.

As indicated above, it is the mean of the untransformed value that is most relevant to prediction of the effect of microbial loads. As such, while it is appropriate to calculate statistical summaries of microbial counts using log-transformed data (to facilitate the use of common statistical methods) it is the average of absolute values that is most relevant to estimation of risk.

Accordingly, it is important to specify clearly, and understand, the methods used in calculations and pay attention to the units used to express microbial loads, so that estimated risks can be compared without confusion.
Bibliography


*******