TECHNICAL REPORT

Modelling the impact of a change in Meat Inspection sensitivity on the surveillance of bovine tuberculosis (bTB) at the country level[1]

European Food Safety Authority[2, 3]

European Food Safety Authority (EFSA), Parma, Italy

SUMMARY

To assess the impact of proposed changes to the current Meat Inspection (MI) on the overall sensitivity of surveillance and control of animal diseases and welfare conditions, a quantitative assessment was performed based on expert opinion and modelling. An external Consortium (COMISURV), under the provision of an EFSA procurement, performed this work. Although the COMISURV report deals with different animal species, the assessment described in this report is focused on bovine tuberculosis only.

The assessment was done by estimating the overall system sensitivity under the two different MI options (as if they were two different tests with different test sensitivity values) using all the relevant parameters: the design herd prevalence; the number of herds in the country; the size of the herds; the within herd prevalence and the proportion of bovines sent to the slaughterhouse.

The final results confirm that, when considering the MI procedures as a test and given that all animals tested negative, the lower the sensitivity at individual level, the lower the confidence when stating that the prevalence is below the relevant threshold (foreseen in the relevant regulation).

A higher confidence level (e.g. 95 %) could be achieved by increasing the sample size, but, in practical terms, this is unfeasible as it is not possible to increase the slaughtering rate and/or the number of herds per country. Hence if the MI test sensitivity is lowered, a greater number of Officially Tuberculosis Free (OTF) countries will not be able to achieve an acceptable confidence of detecting at least one positive animal, should the prevalence be equal to or above the threshold set up by the law.

Looking at the problem from another perspective, a MI based on a lower test sensitivity, compared to the classical MI, can only perform at the same level of confidence (for the same Design Prevalence) if the Within-Herd Prevalence (WHP) is higher.

KEY WORDS

Tuberculosis, Freedom from disease, Hypergeometric, Herd sensitivity, Area sensitivity

Suggested citation: European Food Safety Authority; Modelling the impact of a change in MI sensitivity on the surveillance of bTB at the country level. Supporting Publications 2013:EN-450. [40 pp.]. Available online: www.efsa.europa.eu/publications

TABLE OF CONTENTS

BACKGROUND AS PROVIDED BY THE EUROPEAN COMMISSION

Regulation (EC) No 854/2004 of the European Parliament and of the Council lays down specific rules for the organisation of official controls on products of animal origin intended for human consumption[4]. Inspection tasks within this Regulation include:

- Checks and analysis of food chain information

- Ante-mortem inspection

- Animal welfare

- Post-mortem inspection

- Specified risk material and other by-products

- Laboratory testing

The scope of the inspection includes monitoring of zoonotic infections and the detection or confirmation of certain animal diseases without necessarily having consequences for the placing on the market of meat. The purpose of the inspection is to assess if the meat is fit for human consumption in general and to address a number of specific hazards: in particular the following issues: transmissible spongiform encephalopathies (only ruminants), cysticercosis, trichinosis, glanders (only solipeds), tuberculosis, brucellosis, contaminants (e.g. heavy metals), residues of veterinary drugs and unauthorised substances or products.

During their meeting on 6 November 2008, Chief Veterinary Officers (CVO) of the Member States agreed on conclusions on modernisation of sanitary inspection in slaughterhouses based on the recommendations issued during a seminar organised by the French Presidency from 7 to 11 July 2008. The CVO conclusions have been considered in the Commission Report on the experience gained from the application of the Hygiene Regulations, adopted on 28 July 2009. Council Conclusions on the Commission report were adopted on 20 November 2009 inviting the Commission to prepare concrete proposals allowing the effective implementation of modernised sanitary inspection in slaughterhouses while making full use of the principle of the 'risk-based approach'.

In accordance with Article 20 of Regulation (EC) No 854/2004, the Commission shall consult EFSA on certain matters falling within the scope of the Regulation whenever necessary.

EFSA and the Commission's former Scientific Committee on Veterinary Measures relating to Public Health have issued in the past a number of opinions on meat inspection considering specific hazards or production systems separately. In order to guarantee a more risk-based approach, an assessment of the risk caused by specific hazards is needed, taking into account the evolving epidemiological situation in Member States. In addition, methodologies may need to be reviewed taking into account risks of possible cross-contamination, trends in slaughter techniques and possible new inspection methods.

TERMS OF REFERENCE AS PROVIDED BY THE EUROPEAN COMMISSION

The scope of this mandate is to evaluate meat inspection in order to assess the fitness of the meat for human consumption and to monitor food-borne zoonotic infections (public health) without jeopardizing the detection of certain animal diseases nor the verification of compliance with rules on animal welfare at slaughter. If and when the current methodology for this purpose would be considered not to be the most satisfactory to monitor major hazards for public health, additional methods should be recommended as explained in detail under points 2 and 4 of the terms of reference.

---

[4] OJ L 226, 25.6.2004, p. 83.

The objectives of the current legal provisions aimed at carrying out meat inspection on a risk-based analysis should be maintained.

In order to ensure a risk-based approach, EFSA is requested to provide scientific opinions on meat inspection in slaughterhouses and, if considered appropriate, at any other stages of the production chain, taking into account implications for animal health and animal welfare in its risk analysis. In addition, relevant international guidance should be considered, such as the Codex Code of Hygienic Practice for Meat (CAC/RCP 58-2005), and Chapter 6.2 on Control of biological hazards of animal health and public health importance through ante- and post-mortem meat inspection, as well as Chapter 7.5 on slaughter of animals of the Terrestrial Animal Health Code of the World Organization for Animal Health (OIE).

The following species or groups of species should be considered, taking into account the following order of priority identified in consultation with the Member States: domestic swine, poultry, soliped animals over six weeks old, soliped animals under six weeks old, domestic sheep and goats, farmed game and domestic solipeds.

In particular, EFSA, in consultation with the European Centre for Disease Prevention and Control (ECDC), is requested within the scope described above to:

1. Identify and rank the main risks for public health that should be addressed by meat inspection at EU level. General (e.g. sepsis, abscesses) and specific biological risks as well as chemical risks (e.g. residues of veterinary drugs and contaminants) should be considered. Differentiation may be made according to production systems and age of animals (e.g. breeding compared to fattening animals).

2. Assess the strengths and weaknesses of the current meat inspection methodology and recommend possible alternative methods (at ante-mortem or post-mortem inspection, or validated laboratory testing within the frame of traditional meat inspection or elsewhere in the production chain) at EU level, providing an equivalent achievement of overall objectives; the implications for animal health and animal welfare of any changes suggested in the light of public health risks to current inspection methods should be considered.

3. If new hazards currently not covered by the meat inspection system (e.g. Salmonella, Campylobacter) are identified under terms of reference (TOR) 1, then recommend inspection methods fit for the purpose of meeting the overall objectives of meat inspection. When appropriate, food chain information should be taken into account.

4. Recommend adaptations of inspection methods and/or frequencies of inspections that provide an equivalent level of protection within the scope of meat inspection or elsewhere in the production chain that may be used by risk managers in case they consider the current methods disproportionate to the risk, e.g. based on the ranking as an outcome of terms of reference 1 or on data obtained using harmonised epidemiological criteria (see annex 2[5]). When appropriate, food chain information should be taken into account.

BACKGROUND AS PROVIDED BY THE EFSA AHAW UNIT

In this framework the SAS unit was asked to perform an assessment of the impact of a change in MI sensitivity on the surveillance of bTB at the country level as a contribution to the AHAW Scientific Opinion on the public health hazards to be covered by inspection of meat (Bovine) (EFSA BIOHAZ Panel, 2013).

---

[5] Annex 2 of the original European Commission mandate.

TERMS OF REFERENCE AS PROVIDED BY THE EFSA AHAW UNIT

To develop a model to assess the impact of the two meat inspection options (i.e. actual practices vs. Visual Only) on the probability of detecting at least one positive herd when the prevalence is greater than the design prevalence values (i.e. within herd prevalence and positive herd prevalence) laid down by the relevant regulation.

The same model should be able to provide an estimation of the sensitivity (confidence) of the surveillance system. An estimation of the Probability of freedom across time should also be investigated. Information on the sensitivity of the Meat Inspection test at individual level should be retrieved from the relevant scientific literature, expert opinion and the COMISURV report.

In detail the modelling work should therefore allow:
- to estimate the HSe of detecting an infected herd through slaughter surveillance for given values of herd size, herd sensitivity and with-in herd prevalence.
- to estimate the confidence of being "free" for an OTF country (estimated by the SSe at the herd DP of 0.001) given the number of cattle and herds examined with no positive cases found over 1 year.
- to provide estimation of parameters that can contribute in assessing if the actual meat inspection procedures may have an added value to the Visual Only option in substantiating freedom from BTB despite its low sensitivity at individual level.

ASSESSMENT

1.   Introduction

To assess the impact of proposed changes to the current MI on the overall sensitivity for surveillance and control of animal diseases and welfare conditions, a quantitative assessment was performed based on expert opinion and modelling. An external Consortium (COMISURV), under the provision of an EFSA procurement, performed this work. The detailed methodology, as well as results and conclusions, together with assumptions and limitations of the modelling, can be found in the COMISURV report for Bovine MI (Dupuy et al., 2012). Although the COMISURV report deals with different animal species, the assessment described in this report is focused on bovine tuberculosis only.

1.1.   Description of the disease and prevalence and relevance in EU

Bovine tuberculosis (bTB) is an infectious disease of cattle caused by *Mycobacterium bovis and M. caprae*[6] and one of the biggest challenges facing the cattle farming industry in some EU Member States.

The risk of transmission of *M.bovis* to humans is currently considered as negligible due to the non-meat-borne nature of the agent. The role of slaughterhouse MI in bovine TB surveillance is, however, of great relevance for the surveillance programmes of the infection in herds and animals.

According to the most recent Zoonoses Report (EFSA, 2013a), 15 MS as well as Norway and Switzerland, were OTF in 2011 according to EU legislation, while some MSs have parts of the country classified as OTF. From the remaining 12 MSs, 9 had bTB in cattle herds, with a total prevalence of 1.12% in 2011. This prevalence has increased steadily from 2007, when it was of 0.46%.

During 2011, a total of 194 cattle herds were infected by *M. bovis* in 5 of the 15 OTF MSs, but the threshold level of 0.1% of positive cattle herds was not exceeded, and the MSs remained OTF according to Directive 98/46/EC[7].

1.2.   Surveillance system currently in place

Slaughterhouse surveillance with lesion detection during commercial slaughter is used as cost-efficient method for passive surveillance of bTB both in OTF and in non-OTF countries, in the latter to supplement live cattle farm testing. The finding of a tuberculous animal at slaughter initiates an investigation through skin testing of the herd of origin and any other potentially exposed animals (Schiller et al. 2010).

The detection and investigation of suspected bTB cases at meat inspection are generally important for the control of the infection anywhere within the EU:

- In non-OTF countries and regions as a supplement to regular and *ad hoc* bTB testing of animals and herds to control or eradicate the disease.

- In OTF countries and regions as a means for continuous monitoring, as required by international rules and regulations, to substantiate claims that bTB prevalence is below the required design prevalence. In these countries, MI may be the only surveillance component in place, or it may coexist with skin testing of herds spaced at longer intervals of time (e.g. 4 years).

---

[6] Both *M.bovis and M.caprae* cause tuberculosis in bovines an other species, including humans. Further in the text, only *M.bovis* is mentioned, but any reference to *M.bovis,* unless the contrary is specified, also includes *M.caprae*.
[7] Council Directive 98/46/EC of 24 June 1998 amending Annexes A, D (Chapter I) and F to Directive 64/432/EEC on health problems affecting intra-Community trade in bovine animals and swine OJ L 198, 15.7.1998, p. 22–39

1.2.1. Surveillance of bTB in OTF MSs or zones thereof

According to EU legislation (Council Directive 98/46/EC), when bTB prevalence falls to a certain threshold, a MS, or part thereof, may be declared OTF if some conditions are met (see Section 2).

With reduction in prevalence of bTB, there is a gradual transition from field-based surveillance on farms to surveillance by MI at the slaughterhouse, and in OTF countries/regions the detection of bTB by MI at slaughterhouse becomes the essential element of the bTB surveillance and substantiates the official bTB freedom status. Examples of this policy may be found in OTF countries such as Australia, USA, Canada or Germany (Corner et al., 1990; Kaneene et al., 2006; Probst et al., 2011; Schiller et al. 2010).

1.3. Impact of proposed changes on Surveillance and Control

The efficacy of MI procedures for detecting *M. bovis* infection may be influenced by many factors, related to the pathobiology of the infection, the intensity of inspection, the skills and dedication of the inspector, and other variables like speed of the chain, etc. (Corner, 1994). In current MI procedures applied in the EU, key tasks for the detection of suspected tuberculous lesions include visual inspection and palpation of the lungs, and palpation and incision of relevant lymph nodes. Detailed examination (including palpation and incision) of these sites may detect as much as 85.4% of animals with a single tuberculous lesion (primary sites of infection) (Corner, 1994). If palpation of lungs and lymph nodes and incision of lymph nodes are omitted, small suspect lesions in these organs may go undetected.

It is generally accepted that the sensitivity of the current MI system for detection of bTB is low. If a change of the current MI to a visual only system were to be introduced, it could further reduce the sensitivity of detection, making the system inefficient and unreliable for surveillance, especially in the case of OTF MS or zones, where MI is the main or only surveillance system in place. To assess the impact of changes on the overall MI sensitivity, the following approach was followed:

- Recent scientific information was reviewed to obtain estimates of sensitivity of detection of bTB by MI, and factors affecting it;

- The information and conclusions contained in two recent reports (FSA, 2013a; COMISURV report comparing the current MI procedures with a visual only system) were reviewed;

- Finally, the effect of a reduction of sensitivity of detection of individual animals by MI on global herd surveillance in OTF zones was modelled (see Section 2) to assess suitability of MI surveillance under criteria of EU legislation (Council Directive 98/46/EC).

1.3.1. Sensitivity of the current MI system for bTB surveillance.

Estimates of the sensitivity of detection of bTB through MI have been published in several reports (Corner et al., 1990; Corner, 1994; Asseged et al., 2004). Comparisons between the results, however, are difficult, partly because a common gold standard is not available and partly because inspection procedures may vary in efficiency between slaughterhouses and inspectors. An in depth discussion on this issue is included in the AHAW Section of the BIOHAZ Scientific Opinion (EFSA BIOHAZ Panel, 2013).

1.3.2. Impact of changes in MI system: the COMISURV report and the UK Food Standards Agency Report.

Two recent reports have dealt with the impact of changes on MI on surveillance and control of animal diseases.

The impact of proposed changes to the current MI on the overall sensitivity for surveillance and control of animal health and welfare conditions was recently assessed by expert opinion and modelling through an EFSA procurement.

A qualitative risk assessment was commissioned by the UK Food Standards Agency (FSA) to determine the implications of a change to a visual only inspection system on the surveillance of animal diseases (FSA, 2013a).

Similar conclusions were drawn in previous assessments of EFSA, where it was concluded that omission of palpation or incisions as performed under current MI legislation would reduce the detection rate of tuberculosis in bovines, and therefore would negatively affect animal disease control.

All assessments performed up to now conclude that the change to a visual only MI system will reduce the detection capability of bTB at the slaughterhouse. In some reports (as in the COMISURV report), it has been estimated by stochastic models that the change would cause a five-fold reduction of sensitivity of MI (test sensitivity) at the animal level. However, the overall impact of this test sensitivity reduction on the MI-based surveillance of bTB is uncertain (FSA, 2013a), and has not been previously quantitatively assessed.

As previously reviewed in this opinion, the bTB surveillance strategy in OTF countries/regions relies mainly on detection of residual *M. bovis* infection at the slaughterhouse by current MI procedures. Practical experience from several OTF countries shows that surveillance based on MI and subsequent trace-back of the infection to other farms provides sufficient data for sustaining the OTF status. In some non-EU countries, however, the detection of suspected lesions and submission of samples to the laboratory is actively encouraged by providing incentives.

The suitability of slaughterhouse surveillance for bTB has been compared recently with other surveillance components in Belgium, being an OTF country (Welby et al., 2012). These authors analysed four different active surveillance components for bTB in a stochastic scenario tree simulation model. From the simulations and the results of the external validation model, it was concluded that surveillance by MI provides the best sensitivity for Belgium. The large sampling coverage was considered an important factor in explaining the high sensitivity of this component (Welby et al., 2012).

2. Outline of the assessment

In collaboration with the Animal Health and Welfare (AHAW) Unit of EFSA, a model was implemented with the aim of assessing the impact of different meat inspection options on the overall sensitivity of a bTB surveillance system in OTF. Conditions for granting and sustaining the OTF status of MS territories or zones thereof are specified in EU legislation. Council Directive 98/46/EC lists down in Annex I the conditions for a Member State or region of a MS to be officially tuberculosis-free (OTF). Point (a) states "the percentage of bovine herds confirmed as infected with tuberculosis has not exceeded 0.1% per year of all herds for six consecutive years [*omissis*], the calculation of this latter percentage to take place on 31 December each calendar year". In addition, in point 5 it is said "The MS or part of a MS will retain officially tuberculosis-free status if the conditions [*omissis*] continue to be met".

In other terms, each country, or part thereof, has to implement a surveillance system in order to demonstrate on a yearly basis that the prevalence of positive herds in the area does not exceed the threshold set by legislation (i.e. 0.001). Many studies have provided principles and methods for the

modelling exercise presented in this report to assess the impact of the two meat inspection options under investigation (i.e. current and visual only) on the surveillance system sensitivity[8] at country level (or part of it) (Cameron et al., 1998; Cameron, 2012; Cannon, 2001 and 2002; Martin et al., 2007).

Basically, the assessment was done by estimating the overall system sensitivity under the two different MI options (as if they were two different tests with different test sensitivity values) using all the relevant parameters:

- The design herd prevalence, i.e. the threshold for infected herds (as included in current EU legislation to define the surveillance goal, i.e. for a country to be recognized as OTF);

- the number of herds in the country;

- the size of the herds;

- the within herd prevalence and

- the proportion of bovines sent to the slaughterhouse.

The surveillance system sensitivity values that were obtained could then be compared and evaluated in terms of the overall ability of the system to detect an actual prevalence of positive herds greater than the one requested by the legislation or, from a more appropriate perspective, the confidence that the actual prevalence is below the threshold, given that no animal tested positive at the slaughterhouse.

It must be clear that this evaluation does not deal with the interpretation of the outcomes of the MI procedures. In particular, the aim of this assessment is not to suggest decisions to be taken in case an animal is found to be positive at slaughterhouse. The consequences of the adoption of the visual only option will be evaluated in terms of confidence, i.e. the probability that the actual prevalence is below the threshold set by the relevant regulation, given that all animals tested negative.

2.1. Methodology and theory

In probabilistic terms, the probability of having exactly **k** positive animals out of the total **N** positive animals can be calculated using the Binomial probability mass function as follows:

$$P(x = k) = \binom{N}{k} \rho^k (1 - \rho)^{N-k} = \frac{N!}{(N - k)! \cdot k!} \rho^k (1 - \rho)^{N-k}$$

where **ρ** is the probability that a single animal is infected.

When **N** is equal to 1, the probability of that animal being positive is equal to the probability **ρ**:

$$P(D+) = \rho$$

When using a test and considering that this test is not perfect, the probability of getting a positive result has to take into account the probability of the animal being truly diseased **ρ** and the probability of the test detecting the positive animal, given that the animal is truly positive ($P(T+|D+)$). Commonly, in a population, the prevalence of the disease (**Prev**) is considered as a good proxy for this probability. Therefore, the probability of getting a positive test from a randomly selected animal in a population is equal to:

---

8 Surveillance System Sensitivity: the ability of the surveillance system of detecting at least one positive unit (herd) when the actual prevalence is greater than the design prevalence.

$$P(T+) = P(T+|D+) \cdot Prev$$

From the previous Equation it is possible to derive the probability of getting a negative test from a randomly selected animal in a population:

$$P(T-) = 1 - (P(T+|D+) \cdot Prev)$$

In general in a survey more than one test is performed. The probability of testing "**n**" animals from a population characterised by disease prevalence "**Prev**" and getting no positive response is equal to:

$$P(x = 0) = [1 - (P(T+|D+) \cdot Prev)]^n$$

Which means that the probability of getting at least one animal testing positive, given that all the animal tested come from a population where the disease prevalence is **Prev**, is equal to:

$$P(x \geq 1) = 1 - [1 - (P(T+|D+) \cdot Prev)]^n$$

1

The first step to calculate the System Sensitivity is then to calculate the Herd Sensitivity using Equation 1, as follows (see Equation 2):

$$HSe_i = 1 - [1 - (TestSe \cdot WHPrev)]^{n_i}$$

2

where **i** is the i-th herd, **HSe** is the Herd Sensitivity; **TestSe** is the sensitivity of the test; **WHPrev** is the prevalence of positive animals within a herd; **n** is the number of tested animals

The Herd Sensitivity is the ability of a round of tests (precisely, **n** tests) to detect at least one positive animal when the actual within herd prevalence is above a given threshold (**WHPrev)**.

This herd level sensitivity value can also be interpreted as the level of confidence when stating that a herd can be considered OTF after *n* negative tests. For example, if the calculated herd sensitivity is equal to 95%, it can be assumed with a 95% confidence level that, after a round of all negative tests, the prevalence in that herd is below the expected design prevalence (**WHPrev**). If this was not the case (i.e. the actual prevalence is above WHPrev), at least one animal should have tested positive (with a 95% probability).[9]

Once the herd sensitivity is calculated for all the herds within the area of interest, these values can be combined to calculate the System Sensitivity (see Equation 2):

$$SSe = 1 - \prod_{i=1}^{w}[1 - (BHPrev * HSe_i)]$$

3

Where SSe is the System Sensitivity; w is the number of herds in the area of interest; BHPrev is the prevalence of positive herds in the area, i.e. the prevalence of herds with an actual within prevalence above WHPrev; HSei is the herd sensitivity for herd i.

The **SSe** expresses how confident one can be when stating that the herd prevalence is below the threshold, given that all herds tested negative. If this was not the case (i.e. the prevalence of positive herds is above BHPrev), at least one herd would have tested positive (i.e. one or more animals in one or more herds would have tested positive with a 95% probability). This would be the parameter of choice when comparing the ability of different MI systems to substantiate the continued OTF status of a country/region.

---

9 Note: in a Hypothesis Testing framework as this one, the reverse interpretation is not true, i.e. when a positive animal is detected it is not possible to state that the actual prevalence is above the threshold. It is just possible to assess that the prevalence is above 0 (zero). Further investigations, based on different approaches, are needed to estimate the actual prevalence.

The overall scenario is schematically illustrated in Figure 1. Each country (squares) has a given number of bovine herds (circles). Considering only the OTF Member States, the prevalence of positive herds against which the surveillance activity has to be designed (Design Prevalence) is 0.001, as specified in the EU regulation. This allows us to evaluate the ability of the MI system to detect, with an acceptable level of confidence[10], at least one positive herd when the prevalence of bTB-positive herds is greater than 0.1%. It is chosen as our standard procedure for comparing the detection ability of different MI systems at the proper level of aggregation, i.e. at the herd level, which is determined by the OTF regulation. From each herd, including the positive ones, a proportion of animals is sent to the slaughterhouse each year (slh). With Equation 2 the HSe for each herd is calculated. Combining the different herd sensitivity values (Equation 3), it is then possible to calculate the SSe (referred to as Area Sensitivity, as it is the sensitivity of a MI surveillance system implemented in an area of interest, i.e. at country level or part of it). Clearly, different values for the carcass level test sensitivity will have an impact on the HSe and on the ASe.
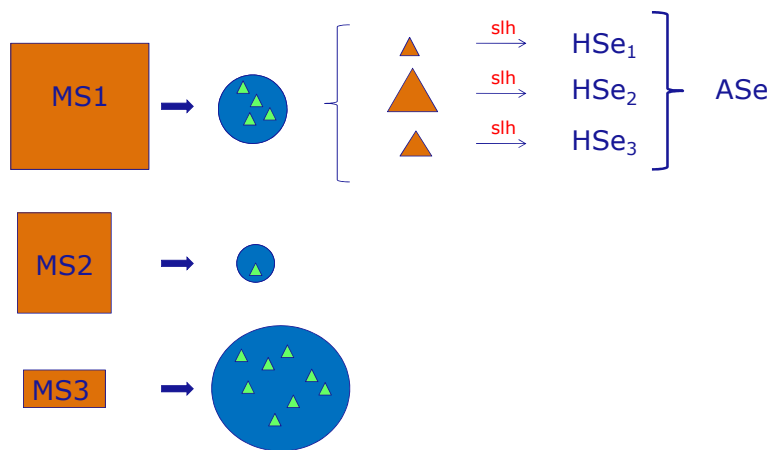

Figure 1: overall scenario. MS=Country; Circles = total number of herds; Triangles in the circles= negative herds; Triangles = positive herds (0.001 of the total number of herds); slh = yearly average proportion of animals sent to the slaughterhouse; HSe = Herd Sensitivity; ASe = Area Sensitivity (country, or part of it).

2.2. Model steps and assumptions

The simulation model is rather simple and the steps are briefly listed below:

STEPS	DESCRIPTION	ASSUMPTIONS
1	Simulation of different scenarios where the area of interest (MS or part of it) is characterised by a certain number of bovine herds	The number of herds per country was obtained from EUROSTAT data and dated from 2007. It was assumed that the figures did not change significantly (see Section 2.3.1).
2	Simulation and assignment of the herd size to each of the herds included in the area of interest	The probabilistic distribution of the herd size used in the simulation was derived from a modelling exercise on data from Belgium, Denmark and United Kingdom. It was assumed that the fitted distribution was representative of the European scenario (see Section 2.3.2).
3	Random selection of a proportion of	It is assumed that all herds have the same

---

10 Note: a confidence of 95% was used in the present modelling exercise though no value is indicated in the relevant EU legislation . As is common knowledge in the statistical framework, 95% is broadly accepted as a reference value for significance.

	positive herds within the area of interest (0.1%). The random process ensures that the herd size distribution will be representative for the whole country.	probability of being infected, independently of their size. Although in reality the probability of infection may vary according to the herd size, many other factors may have an impact on the within herd probability of infection (e.g. production type, age of the animals). Including these elements in the modelling exercise would have meant making a series of necessary assumptions, as well as information on potential factors (which was not available in general), and for this reason, a simple assumption (homogeneous probability of infection across herds) was considered. As the pool of positive herds is the outcome of a random sampling from the total number of herds, the sample will be representative of the area (i.e. more herds with small size). It could be argued that, as the prevalence is adjusted for small herds (see step 4), this means that the majority of the slaughtered animals would come from those small herds with a high prevalence, introducing an important bias (consider that positive herds with 1 animal have a within herd prevalence of 100%). In reality, this bias is mitigated by step 5 of the simulation model: as each animal in each herd has the same probability of being selected for the slaughter, then the lower the number of animals in a herd the less frequent this herd will be included in the calculation of the system sensitivity.
4	Adjustment of the within herd design prevalence as follows: • For herds with more than 1000 animals: DP = 0.001 • For herds with less than 1000 animals: DP = 1/herd size	This adjustment was included since for herds with less than 1000 animals the number of positive animals when the prevalence is 0.001 would have been less than 1, which is of course impossible. For those herds, an alternative design prevalence was calculated as 1 (the minimum number of positive animals) out of the total number of animals in that herd.
5	Random selection from the positive herds of the proportion of animals to be sent to the slaughterhouse and tested.	
6	Based on the adjusted design prevalence, calculation of the HSe for each herd (See Equation 2)	
7	Calculation of the ASe (confidence; see Equation 3)	

It has to be pointed out that the Equation used in step 7 to calculate the ASe had to be adapted to the simulated scenario as follows:

$$SSe = 1 - \prod_{i=1}^{w}[1 - (HSe_i)]$$

<div align="right">4</div>

It can be seen in Equation 4 that **BHPrev** does not appear anymore because the product includes only the HSe values from the **w** positive herds (which were selected randomly in step 3).

It is also essential to highlight that the equations used assume a diagnostic test with 100% specificity (Sp=1). About this assumption in Cannon et al. it is stated that in general, the design of any survey to demonstrate freedom from / absence of infection should specify a sequence of further testing that would be done to clarify the true status when a positive reaction is detected and questioned (Cannon, 2002). Such a sequence would effectively result in a 100%-specific test. In addition, the assumption of perfect specificity has an important consequence: if a positive test result is returned by a system having 100% specificity, freedom from infection can no longer be claimed, as all positive results are true. Each surveillance system should be seen to encompass all necessary follow-up testing to resolve potential false positive results (Cannon, 2001; Dufour et al, 2001; Martin et al., 2007).

The simulation model was run under different scenarios (see Section 2.4)

2.3. Data input and descriptive statistics

In order to make the model work in a situation as close as possible to the actual scenario in Europe, it was decided to retrieve information from official European institutions and / or from official national institutions.

Data needed consisted of:
- The number of herds per country
- The size of the herds and the number of herds of each particular herd size
- The proportion of animals sent to the slaughterhouse per year per herd.

2.3.1. Number of herds per country

The data were retrieved from EUROSTAT, with experts' consent. The most updated information came from year 2007 and for 29 countries (27 Member States and 2 EAA; see Table 1).

Table 1: List of the available data on the number of herds per country (EUROSTAT 2007)

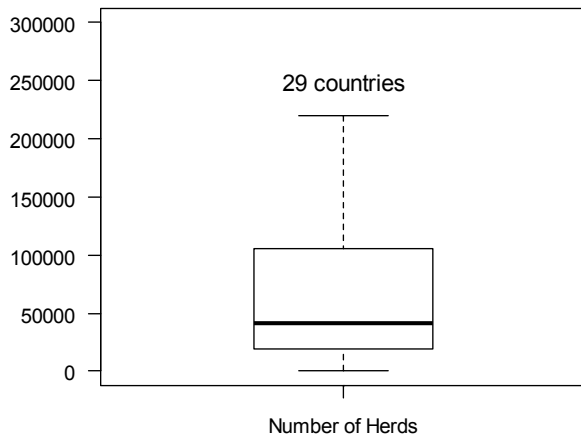| Country	N of herds
MT	230
CY	290
LU	1480
EE	7420
CZ	13950
SK	15450
DK	15590
FI	18630
NO	19640
HU	19800
EL	21540
SE	23870
BE	28470
NL	35260
SI	40840
CH	43720
LV	47350
PT	52140
HR	54370
AT	76740
UK	94650
IE	104930
ES	124030
LT	132600
BG	133330
IT	147020
FR	219970
PL	718250
RO	1067730

Figure 2: Boxplot on the distribution of the number of herds per country. Data = EUROSTAT 2007

Summary statistics are shown in Figure 2 and in Table 2.

Table 2: Summary statistics of the data on number of herds per country (source: EUROSTAT)

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	Records
230	18,630	40,840	113,100	104,900	1,068,000	29

It can be seen that the frequency distribution of the number of herds per country is skewed to the right: although the maximum number of herds is considerably large (1,068,000), the median and the mean show much lower values (40,840 and 113,000 respectively), i.e. there are many countries with a low number of herds.

As the interest was not on the specific situation of a given country, the data were used to make the evaluation in 29 different scenarios, with the purpose of reflecting the current European situation.

2.3.2. Herd size

Data on the number and size of herds from year 2011 were retrieved from three countries (Belgium, Denmark and United Kingdom).

The data from the three countries were then merged. Three different distributions were fitted to the data (see from Figure 3 to Figure 5) and the related Akaike Information Criterion (AIC, Burnham et al, 1998) calculated. The results suggested that a Weibull probabilistic distribution does allow for a reduced loss of information.
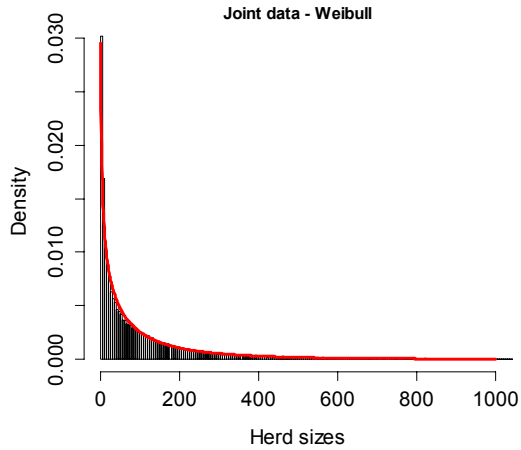
Figure 3: density plot of the herd size; joint data (Belgium, Denmark, United Kingdom), 2011. The fitted Weibull distribution is shown.
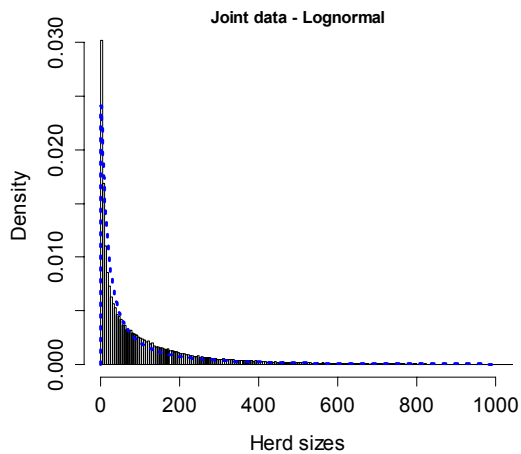

Figure 4: density plot of the herd size; joint data (Belgium, Denmark, United Kingdom), 2011. The fitted Lognormal distribution is shown.
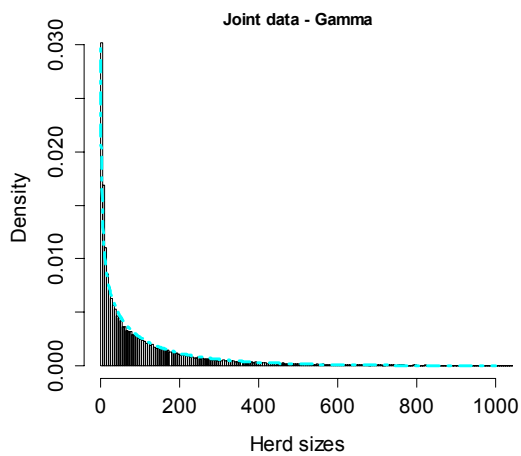

Figure 5: density plot of the herd size; joint data (Belgium, Denmark, United Kingdom), 2011. The fitted Gamma distribution is shown.

Again, the best fitting probabilistic distribution was the Weibull (see Table 3), which was then adopted for the simulation modelling exercise.

Table 3: Akaike Information Criteria values calculated for the 3 fitted distributions

	AIC
Weibull	1369202
Lognormal	1371561
Gamma	1373833

2.4. Parameters and scenarios

2.4.1. Simulation of a realistic scenario

Based on the available data, i.e. the number of herds per country (see Section 2.3.1) and the estimated number of animals per herd within a country (see Section 2.3.2), it was possible to simulate a realistic scenario with 29 different hypothetical countries. For each country it was possible to have details on the number of herds and on the number of animals per each herd.

As an example, country 10 has 19800 herds and the empirical density function of the herd size is shown in Figure 7. Note that the shape of the empirical density function reflects the Weibull probabilistic distribution fitted to the joint data (Belgium, Denmark, United Kingdom).
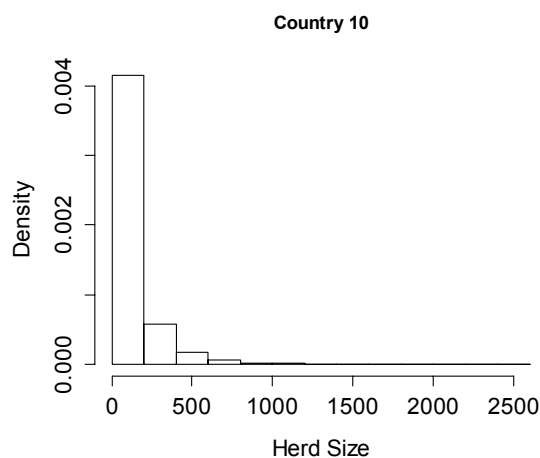

Figure 6: Empirical density plot for hypothetical country 10

2.4.2. Slaughter (replacement) rate

The number of animals sent by each farm to the slaughterhouse is equal to the number of tested animals per farm, as all of them undergo the meat inspection.

The model was run under three plausible values for the replacement rate, i.e. the proportion of animals slaughtered per year per herd:
- 20% (minimum)
- 35% (best guess)
- 40% (maximum)

The option of using a stochastic model was considered, but rejected. Although it is common to use a (Beta) Pert distribution in such a situation, this still represents a strong assumption. Indeed, the probability density between the best guess and the extreme values is arbitrarily imposed, while, in fact, no knowledge is available. If the true underlying distribution was bimodal, for example, the imposed Pert distribution would introduce an important bias which would be reflected in the final outcome. The option of comparing the outcome under three (in this case) scenarios as listed in the bullet points above (according to the so called "what if approach") was considered to be more appropriate than relying on an extra assumption in relation to the probabilistic distribution to be used, especially if exact information on the values between the extremes are not of interest.

2.4.3. Test (Meat Inspection) sensitivity

The problem of estimating the values to be used in the model for the Meat Inspection sensitivity (hereinafter referred to as "test sensitivity") was approached in the same way as for the replacement rate (see section 2.4.1).

In this case, it was decided to employ the sensitivity values derived from a meta-analysis by VLA (2011). The detailed analysis of this information is published in a recent EFSA Opinion (EFSA, 2012) for the current, classical (CL) meat inspection (Table 4).

Table 4: Test sensitivity values (for classical meat inspection)

Q 0.025.	Mean	Q 0.975
0.38	0.71	0.92

In addition to these values, it was agreed to use the lowest value available across the relevant scientific literature. This value comes from a study conducted in Ethiopia and is equal to 0.286.

The values to be used for the Visual Only (hereinafter referred to as "VO") option were estimated from the results of the COMISURV quantitative assessment (COMISURV report - Dupuy et al., 2012). In particular, two scenarios were explored where the VO test sensitivity was 3 and 5 folds lower than the current system (see Table 5).

Table 5: VO test sensitivity (for the Visual Only meat inspection option)

	Min	Q 0.025.	Mean	Q 0.975
3 fold less	0.095	0.127	0.237	0.307
5 fold less	0.057	0.076	0.142	0.184

2.5.   Within Herd prevalence (WHP) and Herd Prevalence (HP)

Two different main scenarios were considered when running the simulation model: OTF countries (or regions) and non-OTF countries (or regions).

In the first case, the design prevalence at the herd level (i.e. the threshold prevalence of positive herds or HP) was fixed at 0.1% as described in the EU regulation for countries or part thereof that need to demonstrate freedom from bTB or keep the free-status. The WHP was considered to be very low, i.e. the threshold prevalence of infected animals within a positive herd was assumed to be 0.1% (i.e. equal to the HP). The values used for OTF countries can be seen as a "worst case scenario" for a surveillance system as the two design prevalence values that have to be detected are extremely low. From an analytical point of view this is, actually, the best option as it evaluates the surveillance systems under investigation in an extreme and critical situation. This allows highlighting the potential weaknesses occurring when modifying the system, as in this case when the sensitivity at animal level is modified (i.e. lowered). The results are presented in Section 2.6.1.

In the case of non-OTF countries it is, of course, not possible to give a fixed and absolute (i.e. valid at European level) estimation of the HP and even less of the WHP in infected herds. For this reason, the performance of the surveillance system was explored across different values of WHP and HP (see Section 2.6.2).

2.6. Results

2.6.1. OTF countries / regions

As described in the previous section, the performance of the surveillance system based on meat inspection was firstly explored when implemented in OTF countries (or part thereof). In this case, the relevant prevalence values were set as follows:

- WHP: 0.001 (for herds with more than 1000 animals) and *1/herd size* (for herds with less than 1000 animals)

- HP: 0.001 (as prescribed in the legislation)

Note that the results presented in the main body of this report are the ones originated by a simulation based on the parameter values considered by the experts as the most likely, i.e.:
- a proportion of slaughtered animals per area (country or part of it) per year equal to 35%;
- a CL-Test Se equal to 0.71, and
- a VO-Test Se 5 fold and 3 fold lower, respectively.

Additional results can be found in the appendix of this report (see Appendix A. and B.).

Figure 7, where the Area Sensitivity is a function of the total number of herds in that area, shows how the probability of detecting at least one positive farm, i.e. the detection ability of the system, decreases importantly when the VO option is implemented.
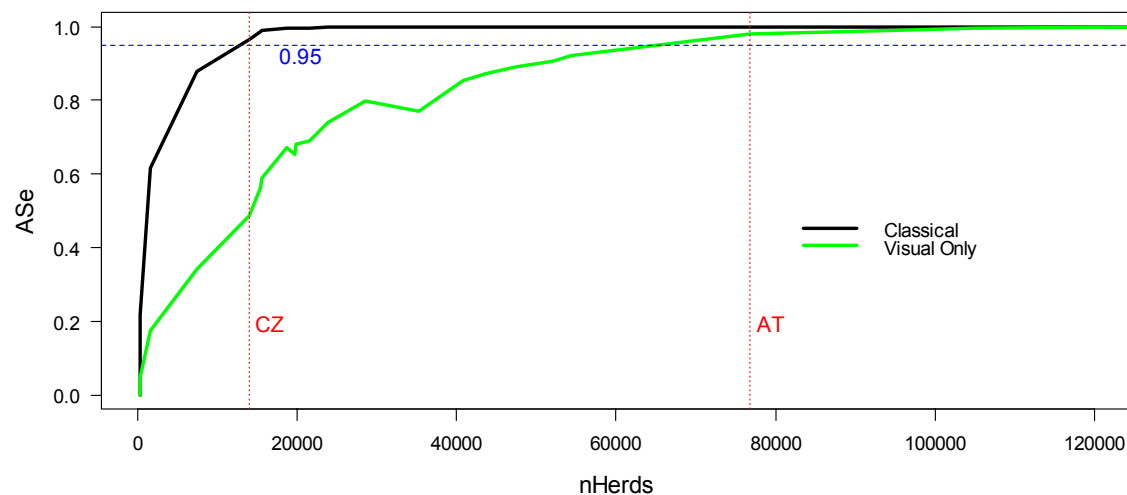
Legend: ASe=Area (country or part of it) Sensitivity, i.e. probability of detecting at least one positive herd when the prevalence of positive herds is above 0.001; nHerds = total number of herds in the area of interest
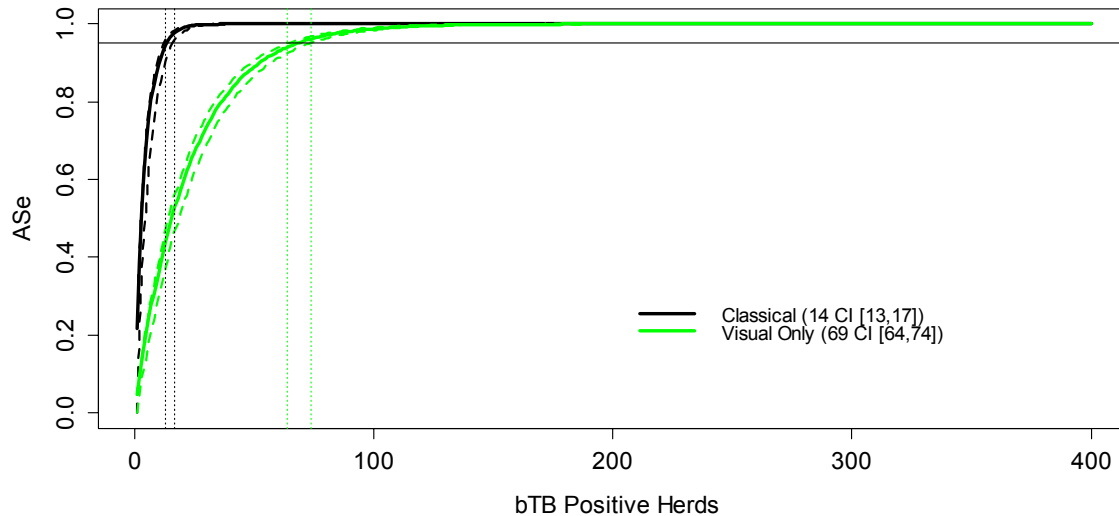
Figure 7: Probability of detection expressed in terms of Area Sensitivity as a function of the number of herds per area. The VO test sensitivity is 5 fold less than the CL test sensitivity.

In other words, following a year where no slaughtered animal tested positive, under the present CL meat inspection 25 countries out of 29 would be 95% confident that the prevalence of positive herds is below 0.1%. On the contrary, if implementing the VO option, only 10 countries out of 29 would be able to reach the same conclusion with the same degree of confidence.

From a quantitative point of view, Table 6 shows that the ASe (i.e. the detection ability) for actual (CL) meat inspection reaches 95% if the total number of herds is at least equal to 13,950. When implementing the VO meat inspection option, the surveillance system needs at least a number of herds equal to 76,740.

Table 6: ASe calculated for 29 different situations where the total number of herds increases. The 29 values correspond to data of 29 countries (Source EUROSTAT). In bold: ASe values above 95%.

| COUNTRY	nHerds	ASeCL	ASeVO
MT	230	0	0
CY	290	0.217482	0.047721
LU	1480	0.616178	0.173206
EE	7420	0.878711	0.340372
CZ	13950	**0.96537**	0.486835
SK	15450	**0.985261**	0.563507
DK	15590	**0.988591**	0.587077
FI	18630	**0.996489**	0.672669
NO	19640	**0.995201**	0.652827
HU	19800	**0.99682**	0.680899
EL	21540	**0.997342**	0.690684
SE	23870	**0.998926**	0.738072
BE	28470	**0.99968**	0.797255
NL	35260	**0.999424**	0.769348
SI	40840	**0.99994**	0.852985
CH	43720	**0.999971**	0.871967
LV	47350	**0.999987**	0.890855
PT	52140	**0.999994**	0.905014
HR	54370	**0.999997**	0.92021
AT	76740	**1**	**0.981376**
UK	94650	**1**	**0.988239**
IE	104930	**1**	**0.995328**
ES	124030	**1**	**0.997649**
LT	132600	**1**	**0.998772**
BG	133330	**1**	**0.998537**
IT	147020	**1**	**0.999254**
FR	219970	**1**	**0.999905**
PL	718250	**1**	**1**
RO	1067730	**1**	**1**

Legend: ASe=Area (country or part of it) Sensitivity, i.e. probability of detecting at least one positive herd when the prevalence of positive herds is above 0.001; bTB Positive Herds = total number of POSITIVE herds in the area of interest (i.e. 0.1% of the total number herds in that area/country)

Figure 8: Probability of detection expressed in terms of Area Sensitivity as a function of the number of bTB positive herds per area. The VO test sensitivity is 5 fold less than the CL test sensitivity.

The curves in Figure 7 are irregular as only 29 data points are available for estimation of the ASe. In order to smooth the curves, the ASe values were displayed as a function of positive herds (instead of the total number of herds). A bit more in detail: an increasing number of positive herds was simulated with a size following the fitted Weibull probability distribution used for the previous simulation. As only the herds with more than 1 positive animal contribute to the calculation, the attention can be focused on these ones, only. The total number of herds can be easily calculated as the number of positive herds is just a proportion (i.e. 0.1% of the total number of herds). Figure 8 provides, indeed, a more detailed picture.

It can be seen that in order to be 95% confident that the prevalence of positive herds is below 0.1%, under the current meat inspection procedures, a median of 14 positive herds (CI [13,17]) need to be in the area (meaning that the total number of herds must be at least 14000, CI [13000, 17000]). On the contrary, 69 positive herds are needed, under the VO meat inspection, to reach the same conclusion with the same level of confidence, meaning that the total number of herds must be at least equal to 69000 (CI [64000, 74000]). From a more realistic point of view, it can be stated that, given the prevalence threshold set in the regulation and the amount of animals sent to the slaughterhouse each year, an area with less than 64000 herds implementing the visual only option will never be able to be 95% confident that the actual prevalence is below 0.001, even if all slaughtered animals tested negative. This number of herds in the area is much lower using the current MI procedures: in this case, the areas that will not be able to reach a 95% confidence will be the ones with less than 13000 herds.
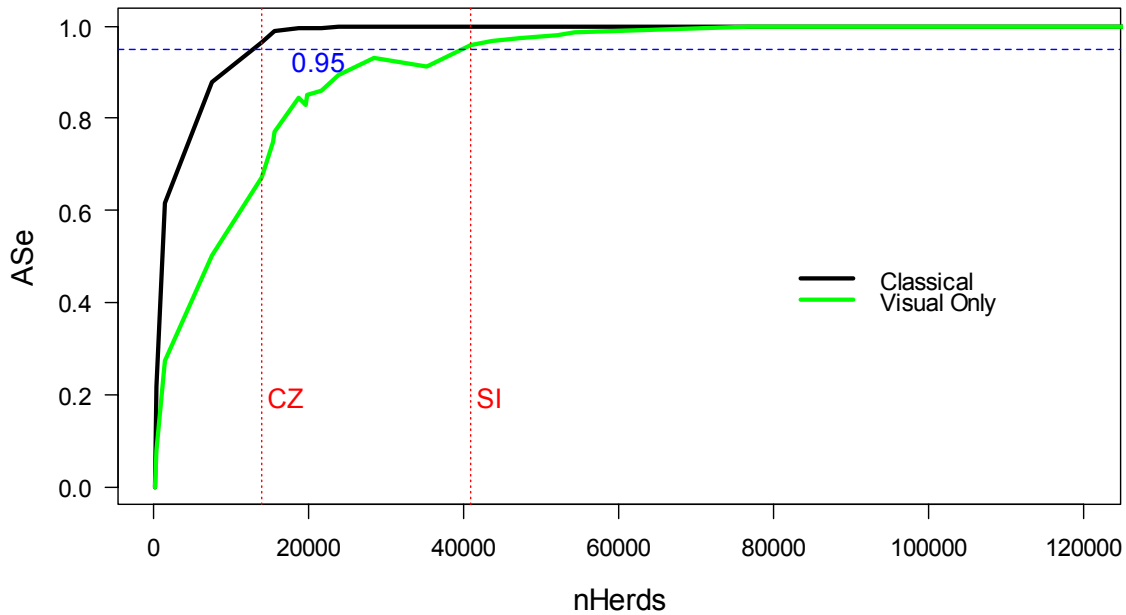
Legend: ASe=Area (country or part of it) Sensitivity, i.e. probability of detecting at least one positive herd when the prevalence of positive herds is above 0.001; nHerds = total number of herds in the area of interest.
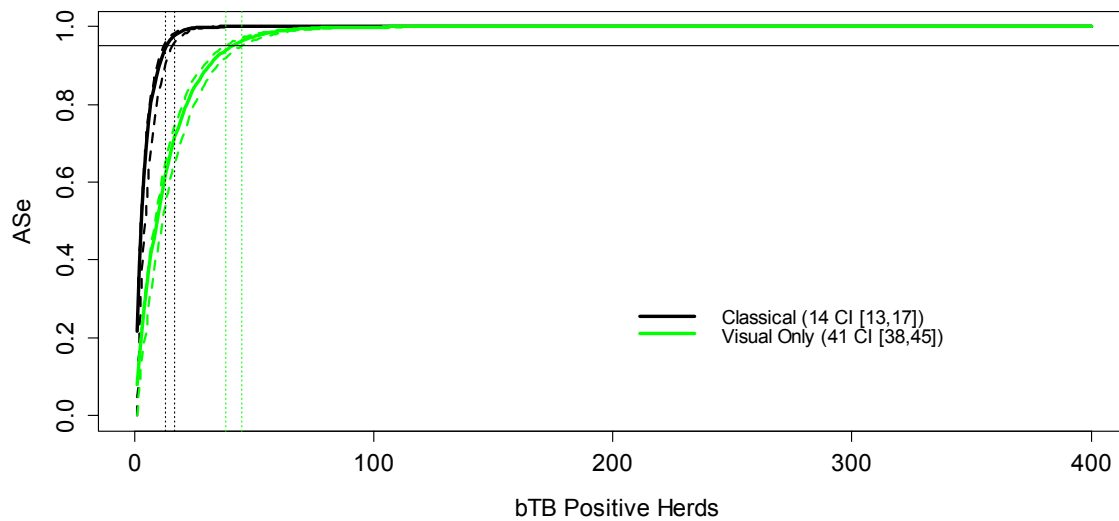
Figure 9: Probability of detection expressed in terms of Area Sensitivity as a function of the number of herds per area. The VO test sensitivity is 3 fold less than the CL test sensitivity

As expected, when the difference between the sensitivities of the two meat inspection options is smaller (that is, 3 fold less, rather than 5 fold less), so is the difference in terms of ASe (see Figure 9). However, Table 7 shows that under current meat inspection conditions 25 countries would be able to detect at least one positive herd (when the prevalence of positive herds is above 0.1%) with a 95% confidence, while under Visual Only conditions, only 15 countries would be able to reach the same conclusion with the same degree of confidence.

Table 7: ASe calculated for 29 different situations where the total number of herds increases. The 29 values correspond to 29 countries (source EUROSTAT). In bold: ASe values above 95%. VO TestSe = CLTestSe/3

COUNTRY	nHerds	ASeCL	ASeVO
MT	230	0	0
CY	290	0.217482	0.07830184
LU	1480	0.6161783	0.27193147
EE	7420	0.8787106	0.50093838
CZ	13950	**0.9653698**	0.67156077
SK	15450	**0.9852608**	0.7497772
DK	15590	**0.9885914**	0.77164899
FI	18630	**0.9964885**	0.84509191
NO	19640	**0.9952007**	0.82896757
HU	19800	**0.9968203**	0.85131101
EL	21540	**0.9973416**	0.85900679
SE	23870	**0.9989259**	0.89353622
BE	28470	**0.9996795**	0.93027401
NL	35260	**0.9994244**	0.91382041
SI	40840	**0.9999401**	**0.95934352**
CH	43720	**0.9999711**	**0.96777097**
LV	47350	**0.9999869**	**0.97529122**
PT	52140	**0.9999938**	**0.98045438**
HR	54370	**0.9999972**	**0.98533042**
AT	76740	**1**	**0.99870971**
UK	94650	**1**	**0.9994006**
IE	104930	**1**	**0.99987191**
ES	124030	**1**	**0.999959**
LT	132600	**1**	**0.99998621**
BG	133330	**1**	**0.99998149**
IT	147020	**1**	**0.99999401**
FR	219970	**1**	**0.99999981**
PL	718250	**1**	**1**
RO	1067730	**1**	**1**

In addition, Figure 10 shows that in order to be 95% confident of a prevalence of positive herds below 0.1% under the current meat inspection procedures, a median of 14 positive herds (CI [13, 17]) need to be in the area (meaning that the total number of herds must be at least 14000, CI [13000, 14000]). Under the VO meat inspection, 41 positive herds would be needed, to reach the same conclusion with the same level of confidence, meaning that the total number of herds must be at least equal to 41000 (CI [38000, 45000]). Again, from a more practical point of view, it can be stated that, given the prevalence threshold set in the regulation and the amount of animals sent to the slaughterhouse each year, an area with less than 38000 herds implementing the visual only option will never be able to be 95% confident that the actual prevalence of positive herds is below 0.001, even if all animals tested negative. This number of herds in the area is much lower if the current MI procedures are implemented: in this case, the areas that won't be able to reach a 95% confidence will be the ones with less than 13000 herds.

Legend: ASe=Area (country or part of it) Sensitivity, i.e. probability of detecting at least one positive herd when the prevalence of positive herds is above 0.001; n_bTB_Herds = total number of POSITIVE herds in the area of interest (i.e. 0.1% of the total herds). VO TestSe = CLTestSe/3

Figure 10: Probability of detection expressed in terms of Area Sensitivity as a function of the number of bTB positive herds per area. The VO test Sensitivity is 3 folds less than the CL test sensitivity.

In order to be able to show all relevant information at a glance, the outputs of the simulations under the different scenarios are presented below (see Figure 11) in order to highlight the differences, in terms of test sensitivity, between the two meat inspection options.

Figure 11 displays on the y-axis the absolute number of positive herds (the proportion is always the same, i.e. 0.1% of the total number of herds in the area of interest) below which it is not possible to state, with a 95% confidence, that the area under investigation has a prevalence of positive herds below 0.1% (given that all animals and herds tested negative). As done also elsewhere in this document, the results can be interpreted also as follows: the output on the y-axis represents the absolute number of positive herds (the proportion is always the same, i.e. 0.1% of the total number of herds in the area of interest) needed by the monitoring system to detect at least one of them, with a 95% confidence, when the prevalence is above the threshold.

Going from the left to the right, it can be seen that the number of positive herds that are "needed" (and therefore the total number of herds in that area) decreases as the sensitivity estimate of the current meat inspection (TSeCL) and the slaughtering rate increase. It is also clear that the smaller the difference in sensitivity between the current meat inspection and the visual only option, the smaller the difference between the numbers of positive herds needed.
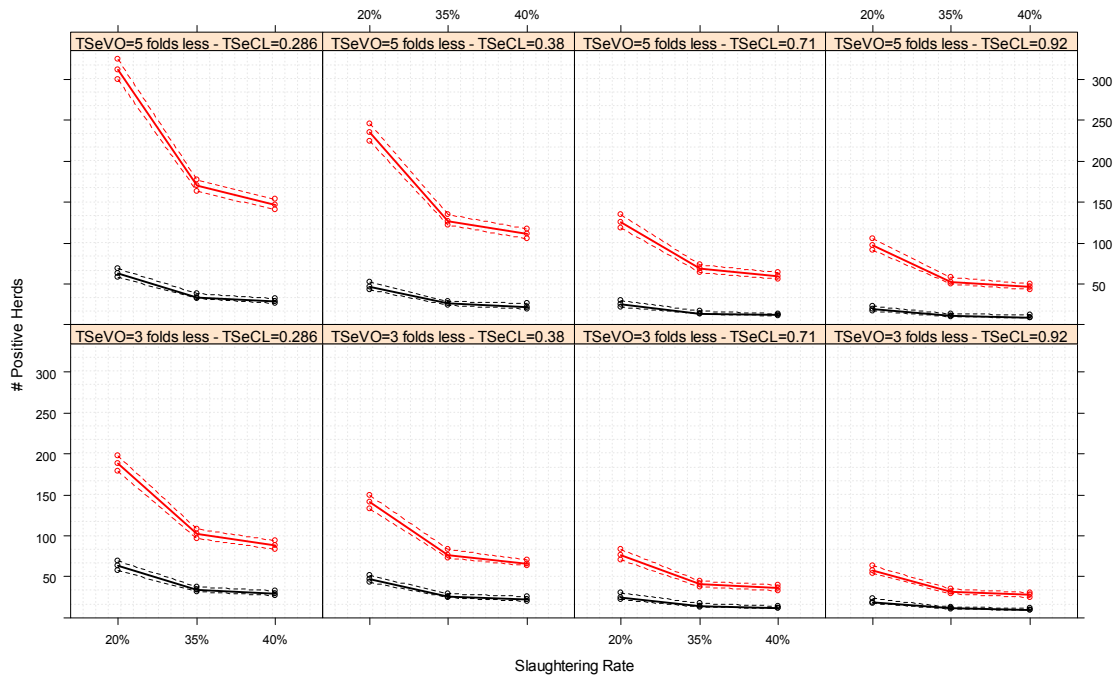
Legend: Solid black line=Surveillance System (SS) based on current meat inspection; Black dashed lines=95% confidence interval; Solid red line=Surveillance System (SS) based on Visual Only meat inspection; Red dashed lines=95% confidence interval; Top line: TSeVO is 5 folds less the TSeCL; Bottom line: TSeVO is 3 folds less the TseCL

Figure 11: Minimum number of positive herds needed in order to reach a System (Area) Sensitivity of 95% as a function of the slaughtering/replacement rate and of the sensitivity of the current meat inspection (TSeCL).

2.6.2. Non-OTF countries / regions

For those countries that are not OTF and therefore might have bTB at varying herd prevalences, the approach cannot be the same as above, since by definition, the methodology given in detail in Section 2.6.1 fits the purpose of substantiating OTF status at a fixed low design prevalence. Therefore, theoretically, another approach should be considered instead of the one used above (i.e. calculation of the sample size needed to detect a disease when this is present at a specific prevalence value).

Nevertheless, it is proposed to use a similar approach in order to estimate what would be the consequences, in terms of confidence at the herd level, of using meat inspection procedures with different sensitivity values. The idea is to answer the following question: provided that the meat inspection procedures gave only negative results, what is the level of confidence that the within herd prevalence is below the design prevalence (where the design prevalence, i.e. the target within-herd prevalence, is calculated as in Equation 5)?

$$DP_{ij} = {}^i/_j$$

5

where *i* is the number of positive animals (going from 1 to 50) and **j** is the herd size (going from *i* to 1000).

Briefly, the approach was performed as follows:

STEPS	DESCRIPTION	ASSUMPTIONS
1	Creation of a scenario with a number of positive animals (*i*) increasing from 1 to 50 and a herd size increasing from *i* to 1000.	NA
2	Calculation of the within herd Design Prevalence (see Equation 5)	Each animal has the same probability of being infected.
6	Calculation of the HSe for each herd (See Equation 2)	Each animal has the same probability of being selected for the slaughterhouse.
7	Assessment of the performance of the meat inspection options under investigation by comparison of the number of positive herds needed by the system to reach a 95% confidence level,	

The values used for the sensitivity of the relevant meat inspection options and the slaughtering/replacement rates are the same as given in Section 2.4.
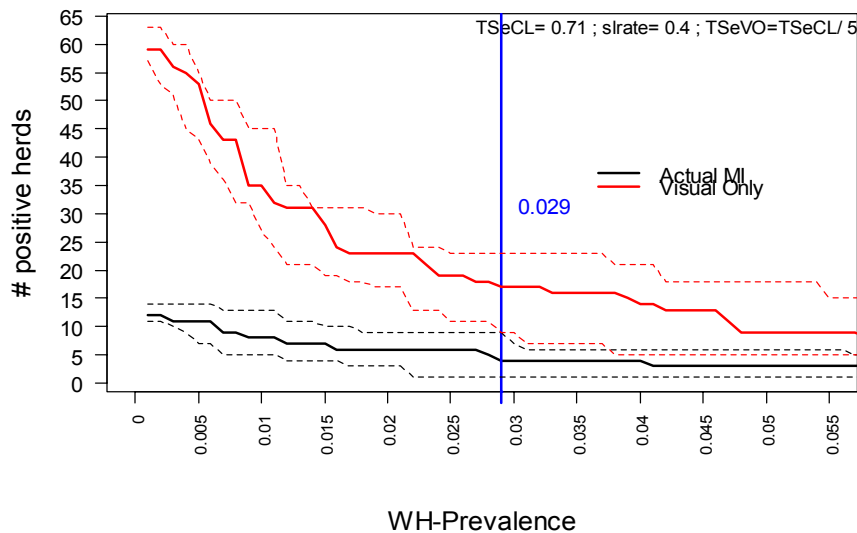
Legend: Black line=Number of positive herds needed by the monitoring system under the current MI (dashed black line=confidence values); Red line= Number of positive herds needed by the monitoring system under the VO option (dashed red line=confidence values); blue line=Within Herd prevalence value above which the 2 meat inspection procedures do not differ in terms of positive herds needed to reach the 95% confidence.

Figure 12: Number of positive herds needed by the monitoring system as a function of the within herd prevalence. Actual meat inspection procedures (black) and Visual Only option (red) are shown. TSeCL = 0.71; slaughtering rate = 0.4; TSeVO = TSeCL/5

Table 8: Number of positive herds (and related Confidence Interval) needed by the system to detect at least one of them when the prevalence is above the specified threshold (WHP)

WHP	N° pos Herds (CL=95%) Actual MI	N° pos Herds (CL=95%) VO option
0.001	12 (11-14)	59 (57-63)
0.002	12 (11-14)	59 (53-63)
0.003	11 (10-14)	56 (51-60)
0.004	11 (9-14)	55 (45-60)
0.005	11 (7-14)	53 (43-55)
0.006	11 (7-14)	46 (39-50)
0.007	9 (5-13)	43 (36-50)
0.008	9 (5-13)	43 (32-50)
0.009	8 (5-13)	35 (32-45)
0.01	8 (5-13)	35 (27-45)
0.011	8 (5-13)	32 (24-45)
0.012	7 (4-11)	31 (21-35)
0.013	7 (4-11)	31 (21-35)
0.014	7 (4-11)	31 (21-31)
0.015	7 (4-10)	28 (19-31)
0.016	6 (4-10)	24 (19-31)
0.017	6 (3-10)	23 (18-31)
0.018	6 (3-9)	23 (18-31)
0.019	6 (3-9)	23 (17-30)
0.02	6 (3-9)	23 (17-30)
0.021	6 (3-9)	23 (17-30)
0.022	6 (1-9)	23 (13-24)
0.023	6 (1-9)	21 (13-24)
0.024	6 (1-9)	19 (13-24)
0.025	6 (1-9)	19 (11-23)
0.026	6 (1-9)	19 (11-23)
0.027	6 (1-9)	18 (11-23)
0.028	5 (1-9)	18 (11-23)
0.029	4 (1-9)	17 (9-23)
0.03	4 (1-7)	17 (9-23)
0.031	4 (1-6)	17 (7-23)
0.032	4 (1-6)	17 (7-23)
0.033	4 (1-6)	16 (7-23)
0.034	4 (1-6)	16 (7-23)
0.035	4 (1-6)	16 (7-23)
0.036	4 (1-6)	16 (7-23)
0.037	4 (1-6)	16 (7-23)
0.038	4 (1-6)	16 (5-21)
0.039	4 (1-6)	15 (5-21)
0.04	4 (1-6)	14 (5-21)
0.041	3 (1-6)	14 (5-21)
0.042	3 (1-6)	13 (5-18)
0.043	3 (1-6)	13 (5-18)
0.044	3 (1-6)	13 (5-18)
0.045	3 (1-6)	13 (5-18)
0.046	3 (1-6)	13 (5-18)
0.047	3 (1-6)	11 (5-18)
0.048	3 (1-6)	9 (5-18)
0.049	3 (1-6)	9 (5-18)
0.05	3 (1-6)	9 (5-18)
0.051	3 (1-6)	9 (5-18)
0.052	3 (1-6)	9 (5-18)
0.053	3 (1-6)	9 (5-18)
0.054	3 (1-6)	9 (5-18)
0.055	3 (1-6)	9 (5-15)

Figure 12 shows an example of the results from this simulation exercise. It can be seen that the two monitoring systems perform at the same level of confidence (95%) only when the within-herd prevalence is above 0.029 (see also Table 8 from where it can be seen that the 0.029 within-herd prevalence is very optimistic as the real threshold value above which the two systems overlap is 0.038). For within-herd prevalence values below 0.029 the two monitoring systems need a significantly different number of positive herds to reach the same level of confidence: in particular, the lower the within-herd prevalence the greater the difference.

For illustration purpose, it is possible to calculate the number of infected animals that the two monitoring system need before showing up a positive test. As said, in an optimistic scenario, the two monitoring systems could perform at the same 95% confidence level once the WHP is equal to 0.029. Note that the within herd prevalence is assumed to be the same in all herds. If we consider an average herd size across Europe (112), the actual MI would show a positive test only when the number of infected animals will be between 3 $(0.029 * 112 * 1)$ and 30 $(0.029 * 112 * 9)$, with herds affected between 1 and 9. On the contrary, under the VO option, the number of animals that the system would require need to be between 30 $(0.029 * 112 * 9)$ and 75 $(0.029 * 112 * 23)$, with affected herds being between 9 and 23.
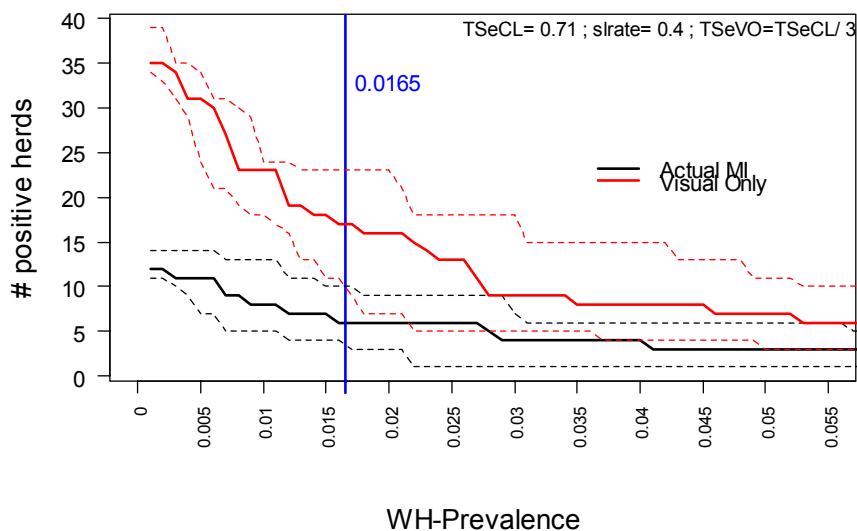
Legend: Black line=Number of positive herds needed by the monitoring system under the current MI (dashed black line=confidence values); Red line= Number of positive herds needed by the monitoring system under the VO option (dashed red line=confidence values); blue line=Within Herd prevalence value above which the 2 meat inspection procedures do not differ in terms of positive herds needed to reach the 95% confidence.

Figure 13: Number of positive herds needed by the monitoring system as a function of the within herd prevalence. Actual meat inspection procedures (black) and Visual Only option (red) are shown. TSeCL = 0.71; slaughtering rate = 0.4; TSeVO = TSeCL/3

The same considerations can be done when the difference between the sensitivity of the two monitoring systems is smaller (3 fold less).

It can be seen from Figure 13 that the two monitoring systems perform at the same level of confidence (95%) only when the within-herd prevalence is above 0.017 (see also

). For within-herd prevalence values below 0.017 the two monitoring systems need a significantly different number of positive herds to reach the same level of confidence: in particular, the lower the within-herd prevalence the greater the difference.

For illustration purpose, it is possible to calculate the number of infected animals that the two monitoring system need before showing up a positive test. As said, in an optimistic scenario, the two monitoring systems would perform at the same 95% confidence level once the WHP is equal to 0.017. Note that the within herd prevalence is assumed to be the same in all herds. If we consider an average herd size across Europe (112), the actual MI would show a positive test only when the number of infected animals will be between 6 ($0.017 * 112 * 3$) and 19 ($0.017 * 112 * 10$), with herds affected between 3 and 10. On the contrary, under the VO option, the number of animals that the system would require need to be between 17 ($0.017 * 112 * 9$) and 44 ($0.017 * 112 * 23$), with affected herds being between 9 and 23.

Table 9: Number of positive herds (and related Confidence Interval) needed by the system to detect at least one of them when the prevalence is above the specified threshold (WHP)

WHP	N° pos Herds (CL=95%) Actual MI	N° pos Herds (CL=95%) VO option
0.001	12 (11-14)	35 (34-39)
0.002	12 (11-14)	35 (33-39)
0.003	11 (10-14)	34 (31-35)
0.004	11 (9-14)	31 (29-35)
0.005	11 (7-14)	31 (24-34)
0.006	11 (7-14)	30 (21-31)
0.007	9 (5-13)	27 (21-31)
0.008	9 (5-13)	23 (19-30)
0.009	8 (5-13)	23 (18-29)
0.01	8 (5-13)	23 (18-24)
0.011	8 (5-13)	23 (17-24)
0.012	7 (4-11)	19 (16-24)
0.013	7 (4-11)	19 (13-23)
0.014	7 (4-11)	18 (13-23)
0.015	7 (4-10)	18 (11-23)
0.016	6 (4-10)	17 (11-23)
0.017	6 (3-10)	17 (9-23)
0.018	6 (3-9)	16 (7-23)
0.019	6 (3-9)	16 (7-23)
0.02	6 (3-9)	16 (7-23)
0.021	6 (3-9)	16 (7-21)
0.022	6 (1-9)	15 (5-18)
0.023	6 (1-9)	14 (5-18)
0.024	6 (1-9)	13 (5-18)
0.025	6 (1-9)	13 (5-18)
0.026	6 (1-9)	13 (5-18)
0.027	6 (1-9)	11 (5-18)
0.028	5 (1-9)	9 (5-18)
0.029	4 (1-9)	9 (5-18)
0.03	4 (1-7)	9 (5-18)
0.031	4 (1-6)	9 (5-15)
0.032	4 (1-6)	9 (5-15)
0.033	4 (1-6)	9 (5-15)
0.034	4 (1-6)	9 (5-15)
0.035	4 (1-6)	8 (5-15)
0.036	4 (1-6)	8 (5-15)
0.037	4 (1-6)	8 (4-15)
0.038	4 (1-6)	8 (4-15)
0.039	4 (1-6)	8 (4-15)
0.04	4 (1-6)	8 (4-15)
0.041	3 (1-6)	8 (4-15)
0.042	3 (1-6)	8 (4-15)
0.043	3 (1-6)	8 (4-13)
0.044	3 (1-6)	8 (4-13)
0.045	3 (1-6)	8 (4-13)
0.046	3 (1-6)	7 (4-13)
0.047	3 (1-6)	7 (4-13)
0.048	3 (1-6)	7 (4-13)
0.049	3 (1-6)	7 (4-11)
0.05	3 (1-6)	7 (3-11)
0.051	3 (1-6)	7 (3-11)
0.052	3 (1-6)	7 (3-11)
0.053	3 (1-6)	6 (3-10)
0.054	3 (1-6)	6 (3-10)
0.055	3 (1-6)	6 (3-10)

Figure 14 shows the outcome of an additional graphical representation of the same modelling exercise. In this case, some parameter were fixed (i.e. TSeCL = 0.71; slaughtering rate = 0.4; TSeVO = TSeCL/3) and the total number of herds needed in the area of interest (dependent variable) is calculated as a function of the within-herd prevalence (assumed to be the same in all the positive herds). Different values of the prevalence of positive herds in the area are considered (from 0.001 to 0.1). As expected, the higher the prevalence of positive herds that needs to be monitored, the lower the number of herds that the system has to test in order to perform at a 95% confidence level.
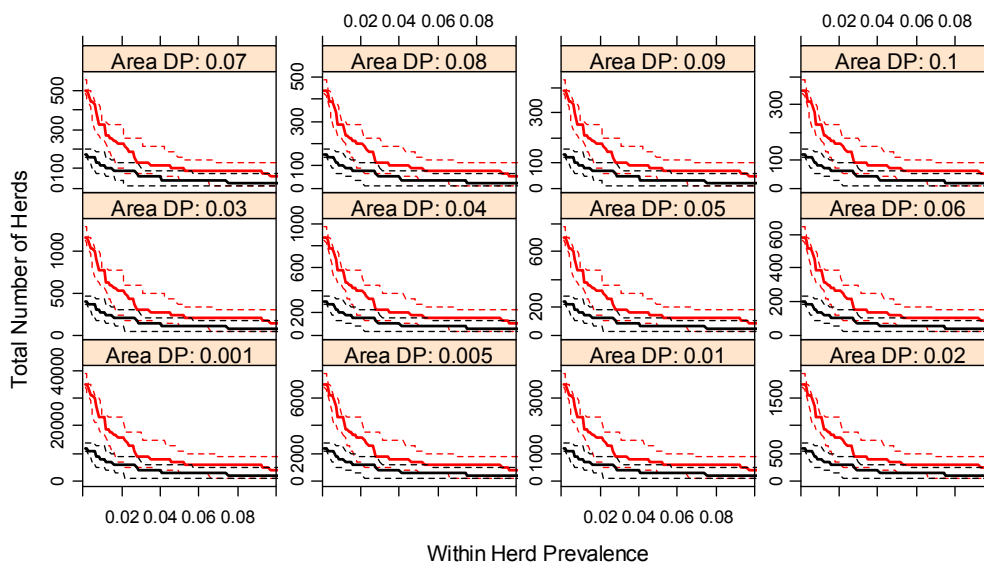
Legend: Black line=Total number of herds needed by the monitoring system under the current MI (dashed black line=confidence values); Red line= Total number of herds needed by the monitoring system under the VO option (dashed red line=confidence values)

Figure 14: Number of positive herds needed by the monitoring system as a function of the within herd prevalence. Actual meat inspection procedures (black) and Visual Only option (red) are shown. TSeCL = 0.71; slaughtering rate = 0.4; TSeVO = TSeCL/3

As predictable, the meat inspection cannot be used on its own as a tool able to detect as many cases as possible: in particular, the system will fail in case of small herds with a low prevalence of infected animals. However, in a surveillance system where more than one component (e.g. tuberculin test + meat inspection), the higher the Component Sensitivity, the higher the confidence achieved by the surveillance system. Therefore, also in non-OTF countries, the current meat inspection procedures appear as the most recommendable in view of eradication.

3.   Probability of Freedom (Pfree)

It must be noted that for this assessment the estimation of the probability of freedom (P_{free}) , based on the Bayesian principles, was not done for two reasons:

- The EU OTF-regulation only sets up the design prevalence at country (or part thereof) level and asks to substantiate that the prevalence is below that threshold **on a yearly basis**. On the contrary, the P_{free} is calculated across years, using a prior probability (System Sensitivity of the previous year) to calculate the present (i.e. following a new year of testing) probability of being free, conditional on the result form previous year(s). This parameter thus does not fit the purpose of substantiating OTF status each year;

- Usually, this parameter is useful when the aim is to achieve a "disease free" status. As an example, a surveillance system with 95% System Sensitivity (if the probability of introduction is not particularly high) will be able to reach a 99% probability of freedom in a few years. In this case, where the considered regions already achieved OTF status, the aim is to keep it.

CONCLUSIONS

The final results, presented in Figure 11 and in Figure 14, confirm what could be concluded intuitively: considering the meat inspection procedures as a test, the lower the sensitivity at individual level, the lower the confidence when stating that the prevalence is below the threshold (foreseen in the relevant regulation) given that all animals tested negative.

If the meat inspection test sensitivity is lowered, a greater number of OTF countries will not be able to achieve an acceptable level of confidence i.e. an acceptable probability of detecting at least one positive animal, should the prevalence be above the threshold set up by the law.

An additional consideration can be made looking at the modelling exercise on non-OTF countries. Also in this case, the only way to keep the same level of confidence when implementing a monitoring procedure (to demonstrate that the prevalence of positive units is below a given threshold) using a test with a lower sensitivity is to increase the sample size. As said, it is of course not possible to increase neither the slaughtering rate nor the number of herds. This means that a MI based on a lower test sensitivity will perform at the same level of confidence only when the WHP will be higher if compared to the classical MI or, which is the same, it will result in at least one positive test when the prevalence is above the same threshold, but with a lower confidence.

RECOMMENDATIONS

This modelling exercise showed what was expected. The estimation of the sensitivity of a monitoring system is based on the following parameters:

1. Test Sensitivity (Post Mortem Inspection in this framework)

2. Sample size (i.e. slaughtering rate)

3. Population size (i.e. number of herds per area and number of animals per herd)

As parameters 2 and 3 cannot be modified, the only parameter which is possible to control is the Test Sensitivity. A lower test sensitivity will lead to a lower area sensitivity, i.e. a lower ability of detection or, from another perspective, a lower confidence on the real prevalence in the population when all animals tested negative. Practically speaking, this means to potentially allow a greater number of infected animals before an action is undertaken, as a system based on the VO option would require more infected animals before detecting a positive animal.

REFERENCES

AHVLA (Animal Health and Veterinary Laboratories Agency) 2011. Project SB4500 Bovine tuberculosis: Infection status in cattle. Extract of the executive summary and figures from the annual surveillance report for the period January to December 2010. Available from: http://www.defra.gov.uk/animal-diseases/files/Surveillance-annual-report-2010.pdf.

Akaike H, 1974. A new look at the statistical model identification. IEEE Transactions on Automatic Control 19, 716–723.

Burnham, Kenneth P. and David R. Anderson. 1998. Model Selection and Inference: A Practical Information-Theoretical Approach. New York: Springer-Verlag.

Cameron AR, 2012. The consequences of risk-based surveillance: Developing output-based standards for surveillance to demonstrate freedom from disease. Preventive Veterinary Medicine, 105, 280-286.

Cameron AR and Baldock C, 1998. A new probability formula for surveys to substantiate freedom from disease. Preventive Veterinary Medicine, 34, 1-17.

Cannon RM, 2002. Demonstrating disease freedom – combining confidence levels. Preventive Veterinary Medicine, 52, 227-249.

Cannon RM, 2001. Sense and sensitivity – designing surveys based on an imperfect test. Preventive Veterinary Medicine, 49, 141-163.

COMISURV Report - Dupuy C, Hendrikx P, Hardstaff J and Lindberg A, 2012. Contribution of meat inspection to animal health surveillance in bovine animals. Available from: http://www.efsa.europa.eu/en/efsajournal/doc/322e.pdf.

EFSA BIOHAZ Panel (EFSA Panel on Biological Hazards), 2013. Scientific Opinion on the public health hazards to be covered by inspection of meat (bovine animals). EFSA Journal 2013;11(6):3266, 261 pp. doi:10.2903/j.efsa.2013.3266.

EUROSTAT (European Statistics), online. Available from http://epp.eurostat.ec.europa.eu/portal/page/portal/eurostat/home/

OIE Animal Production Food Safety Working Group. (2006). Guide to good farming practices for animal production food safety. *Rev. Sci. Tech*, *25*, 823-836.

APPENDIX

A. COMPLETE SET OF RESULTS FROM THE MODELLING EXERCISE FOR OTF AREAS

SI Rate	Test CL	nHerd+_CL	CL_CI_low	CL_CI_up	Test VO	nHerd+_VO	VO_CI_low	VO_CI_up
20%	0.286	63	58	69	3 fold lower	188	179	198
20%	0.38	47	43	52	3 fold lower	141	133	150
20%	0.71	25	22	30	3 fold lower	76	70	83
20%	0.92	19	17	23	3 fold lower	58	54	64
20%	0.286	63	58	69	5 folds lower	312	300	325
20%	0.38	47	43	52	5 folds lower	235	225	246
20%	0.71	25	22	30	5 folds lower	126	118	135
20%	0.92	19	17	23	5 folds lower	97	91	105
35%	0.286	34	32	38	3 fold lower	102	97	108
35%	0.38	26	24	29	3 fold lower	77	73	83
35%	**0.71**	**14**	**13**	**17**	**3 fold lower**	**41**	**38**	**45**
35%	0.92	11	10	13	3 fold lower	32	29	35
35%	0.286	34	32	38	5 folds lower	170	163	178
35%	0.38	26	24	29	5 folds lower	127	122	135
35%	**0.71**	**14**	**13**	**17**	**5 folds lower**	**69**	**64**	**74**
35%	0.92	11	10	13	5 folds lower	53	50	58
40%	0.286	29	27	33	3 fold lower	88	84	94
40%	0.38	22	20	26	3 fold lower	66	63	71
40%	0.71	12	11	14	3 fold lower	36	33	40
40%	0.92	9	9	12	3 fold lower	28	25	31
40%	0.286	29	27	33	5 folds lower	147	141	154
40%	0.38	22	20	26	5 folds lower	111	105	117
40%	0.71	12	11	14	5 folds lower	59	56	64
40%	0.92	9	9	12	5 folds lower	46	43	50

B. COMPLETE SET OF SCENARIOS

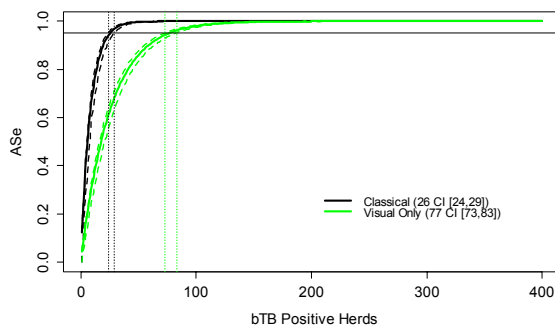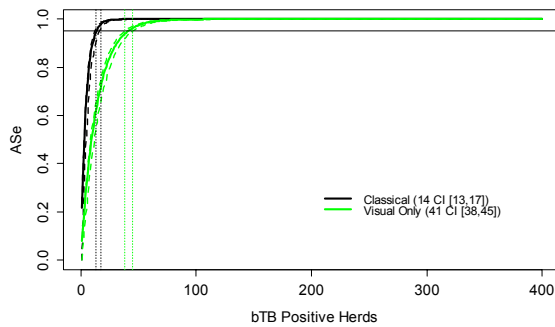- Slaughtering rate: 0.2 - Test Sensitivity (VO): 3 fold less
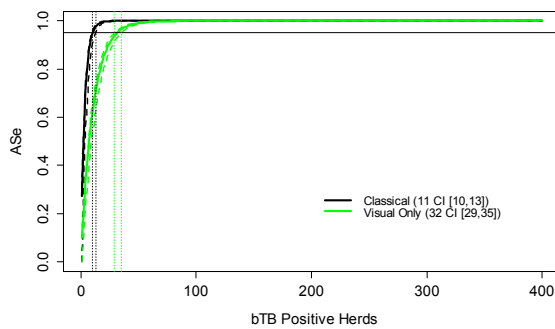
  o Test Sensitivity (CL): 0.286

  o Test Sensitivity (CL): 0.38
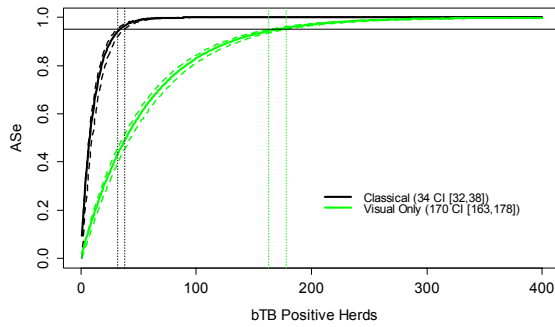

  o Test Sensitivity (CL): 0.71
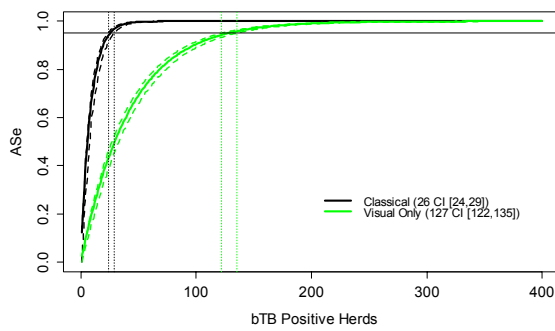

  o Test Sensitivity (CL): 0.92

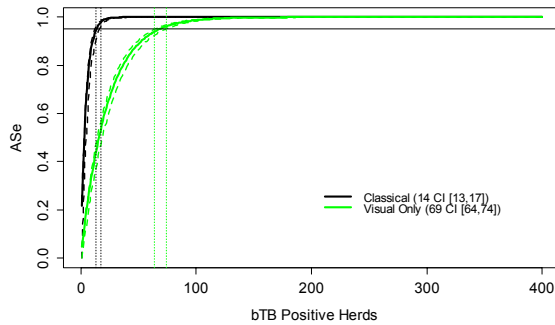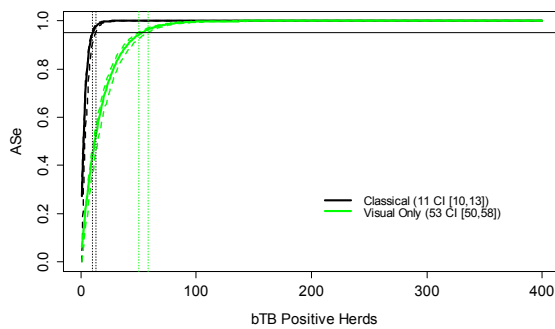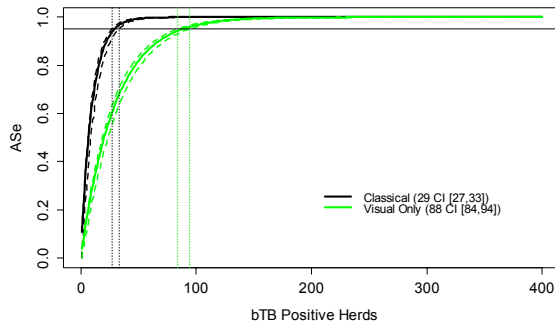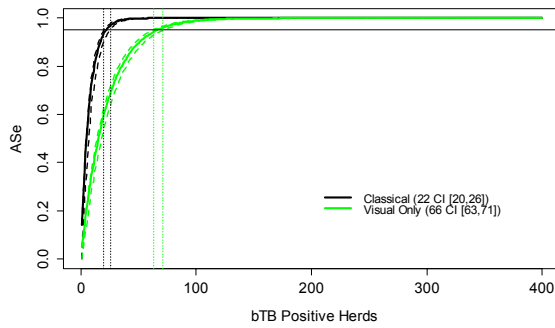- Slaughtering rate: 0.2 - Test Sensitivity (VO): 5 fold less

  o Test Sensitivity (CL): 0.286

  o Test Sensitivity (CL): 0.38

  o Test Sensitivity (CL): 0.71

  o Test Sensitivity (CL): 0.92

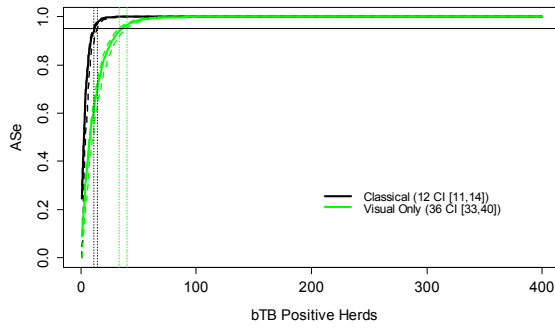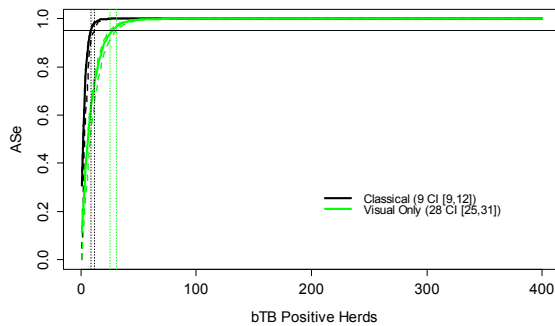- Slaughtering rate: 0.35 - Test Sensitivity (VO): 3 fold less

  o Test Sensitivity (CL): 0.286

  o Test Sensitivity (CL): 0.38

  o Test Sensitivity (CL): 0.71

  o Test Sensitivity (CL): 0.92

- Slaughtering rate: 0.35 - Test Sensitivity (VO): 5 fold less

o Test Sensitivity (CL): 0.286

o Test Sensitivity (CL): 0.38

o Test Sensitivity (CL): 0.71

o Test Sensitivity (CL): 0.92

- Slaughtering rate: 0.40 - Test Sensitivity (VO): 3 fold less

  o Test Sensitivity (CL): 0.286

  o Test Sensitivity (CL): 0.38
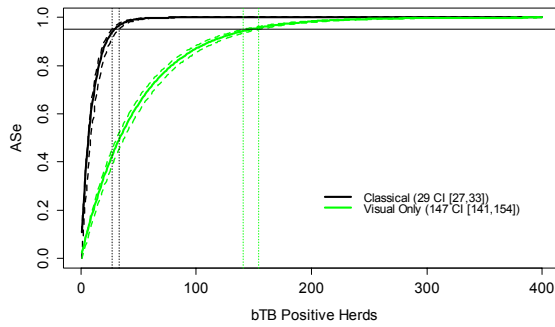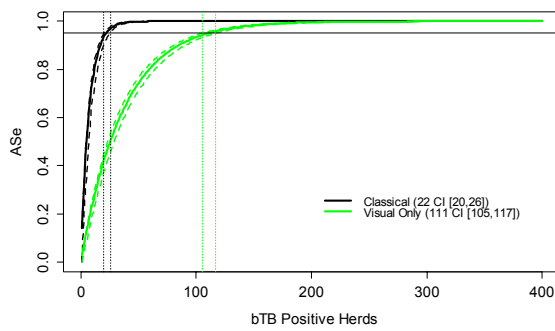

  o Test Sensitivity (CL): 0.71

  o Test Sensitivity (CL): 0.92

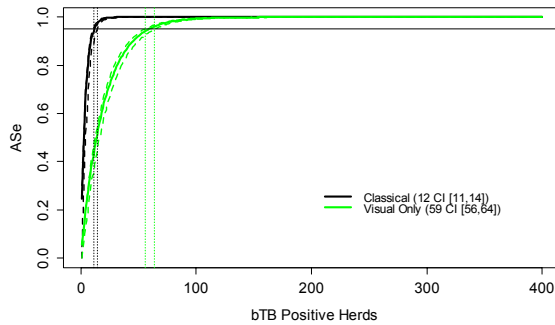- Slaughtering rate: 0.40 - Test Sensitivity (VO): 5 fold less
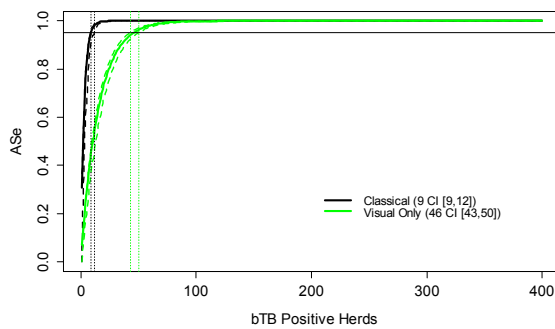
  o Test Sensitivity (CL): 0.286

  o Test Sensitivity (CL): 0.38

  o Test Sensitivity (CL): 0.71

  o Test Sensitivity (CL): 0.92

GLOSSARY AND/OR ABBREVIATIONS

AHAW	Animal Health and Welfare
AWO	Animal welfare officer
AMI	*Ante-mortem* inspection
BIOHAZ	Biological Hazards Panel
bTB	Bovine Tuberculosis
CI	Confidence interval
CL	Classical / Actual meat inspection
CONTAM	Contaminants in the Food Chain Panel
CVO	Chief Veterinary Officer
DFD	Dark firm dry meat
EFSA	European Food Safety Authority
(Ag-)ELISA	(Antigen-capture) Enzyme-linked immunosorbent assays
EU	European Union
FBO	Food business operator
FCI	Food Chain Information
FSA	Food Standards Agency
GI	Gastro intestinal Tract
HACCP	Hazard Analysis and Critical Control Point
I	Incision
Lnn	Lymph nodes
Non-OTF	Non officially tuberculosis free
MI	Meat inspection
ML	Most likely – which is equivalent to mode
MS	Member state
OIE	World organization for animal health
OTF	Officially tuberculosis free
OV	Official veterinarian
P	Palpation
PCR	Polymerase chain reaction
Pfree	Probability of Freedom
PHR	Public health risk
PMI	*Post-mortem* inspection
SAS	Scientific Assessment Support
TOR	terms of reference
V	Visual inspection
VO	Visual only meat inspection
WG	Working group