EFFICACY AND EQUIVALENCE OF PHYTOSANITARY MEASURES

A discussion and reference paper prepared for the
IPPC Expert Working Group on the Efficacy of Phytosanitary Measures
Imperial College, UK 2-4 July 2002

Ricardo Sgrillo
Cocoa Research Center (CEPLAC/CEPEC) - University of Santa Cruz (UESC)

CONTENTS

1. Introduction

The World Trade Organization Agreement on the Application of Sanitary and Phytosanitary Measures (the SPS Agreement) emphasizes the importance of a sound technical basis for measures affecting trade. In spite of the recognition of this concept in the SPS Agreement, very little has been done regarding the development and application of analytical methods to establish quantitative parameters for phytosanitary measures based on fundamental mathematical concepts and statistical analyses.

The increasing development and application of international standards associated with the SPS Agreement and increasing pressure by trading partners for greater transparency and recognition of equivalence has led to a shift in the paradigm of past approaches. This has resulted in greater interest and need for appropriate quantitative methods to measure, assess, and justify risk management decisions.

The main quantitative application widely used by the phytosanitary community in the past is the probit 9 approach, as it was proposed more than 60 years ago as the basis for measuring the efficacy of treatments for fruit flies (–Baker, 1939). The general rationale behind the probit 9 approach is very simple and straightforward: *if you kill a very high percentage (99.9968 %) of the pest population then you will decrease the chance of pest establishment to zero (or very close to zero)*. This rationale assumes a worst case scenario and results in a high-kill treatment requirement that may not be technically justified based on pest prevalence.

The probit 9 approach has been criticized, in particular because: "*While probit 9 was clearly designed to reduce the prevalence of pests by a predictable amount, it does not account for other variables contributing to pest risk. Natural survival rates, the likelihood of infestation, and the colonization potential of the pest are a few of the more important risk based considerations that are ignored by a direct estimation of mortality such as probit 9. Process parameters such as pre-shipment cultural practice, packing and shipping procedures, and distribution times or areas, are not considered when mortality is the sole criterion for determining quarantine security* (Liquido et al. 1997)

In recent years, the evolution of phytosanitary approaches to measuring the efficacy of measures has led to the recognition of another rationale: *if you decrease the number of fertile pest individuals in a commodity to a pre-established number assessed through risk analysis, then you will decrease the risk of establishment of this pest to an acceptable level.* A considerable effort and significant progress has been made through the application of this rationale to understanding options for achieving a desired level of quarantine security. This has opened the door for expanding the understanding and use of many important concepts for risk management, including:

Pest Free Areas (places and sites of production) - a risk management option based upon a sound pest risk assessment coupled with satisfactory evidence of effective, on-going surveillance and exclusion measures to maintain such areas pest free. (Liquido et al. 1997)

Systems Approaches - consecutive phytosanitary measures are used to decrease the number of fertile pest individuals in a commodity.

Alternative Treatment Efficacy - where the percentage of the population of the pest that has to be killed is variable, depending mainly on the initial infestation level and on the characteristics of the pest and of the commodity.

Maximum Pest Limits (especially for fruit flies) has developed as "the maximum number of immature fruit flies that may be present in consignments imported during a specific time to a specific location" (see Bartlett et al. 1996 and Liquido et al. 1995 for a deeper discussion on these concepts).

These initiatives demonstrate a trend toward exploring and applying more systematic approaches to evaluating (and justifying) the efficacy of phytosanitary measures. Yet a common understanding of the range of possible quantitative methodologies and the development and routine use of appropriate methodologies to evaluate and compare the efficacy of phytosanitary measures is far from being well-developed in the phytosanitary community.

While considerable progress has been made and guidance developed to measure the risk associated with a specific pest through Pest Risk Analysis, very little has been done on evaluating the efficacy of phytosanitary measures. The scientific literature includes several papers on the evaluation of efficacy for pest-commodity specific measures, but there is very little available on general concepts associated with

this subject. This continues to be a fundamental problem for the phytosanitary community where there is an important and growing need to fully apply the principles of technical justification and equivalence under the SPS Agreement.

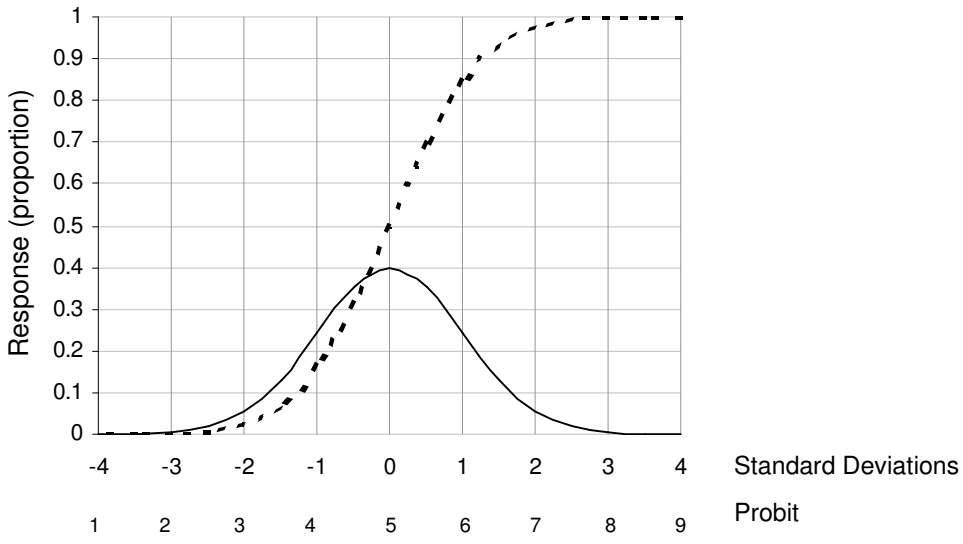
2. Probit - a brief review

The only quantitative evaluation of the efficacy of phytosanitary measures that is widely used today is the quantification of the efficacy of treatments (chemical, heat, etc), mostly for fruit flies.

There are many methods to describe the response of a population to increasing doses of a treatment, including the Log-linear, Logit, and probit models. When the criterion for the population response is mortality, the Lethal Dose (LD), Lethal Concentration (LC), general mortality factor (Q), among others, have also been proposed. In plant quarantine, the expression most used to describe the efficacy of treatments is the probit.

The response of a population to a dose reflects the genetic variability of the natural populations. Usually few individuals will die with lower doses; most of the individuals will die with average doses; and few individuals will only die with higher doses. It is accepted (although not always true) that the statistical distribution of the individuals responding to increasing doses fits a normal or Gaussian distribution (bell shaped curve). The variability of the population (spread of the data points) is measured by its standard deviation(s). The graph in Figure 1 represents the cumulative and distributed response of a population exposed to increasing doses of a chemical product. The data were normalized, that is the zero value in the X axis represents the mean dose and the *s* is 1. The X scale is presented in multiples of the *s* (or normal equivalent deviates). The Y axis presents the proportion of individual that dies in each dose. The graph presents also a probit scale; that is, the same *s* scale, where 5 was added to avoid zero and negative values.

Figure 1 Cumulative (dashed) and distributed (continuous) dose-response curves



Probit only represents a proportion of a population, normally distributed, that responds to a dose. To have information on the number of individuals affected by the dose, the proportion has to be multiplied by the total number of individual in the population.

2.1 Probit 9

Probit 9 value is usually required for quarantine treatments for fruit flies and corresponds to a mortality of 0.999968 of the population. That is: for populations with less then 31,250 (1/{1-0.999968}) individuals there would be less than one survivor. The number of survivals is proportional to the population size. 10 survivors would be expected for a population of 312,500 individuals. The reasons for selecting these

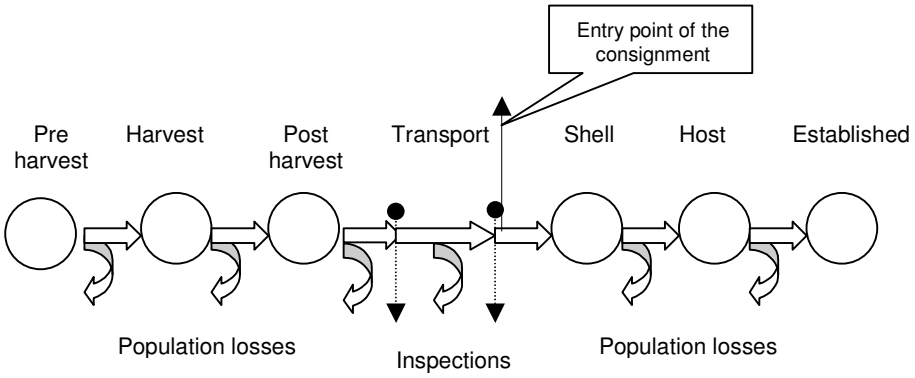
criteria for quarantine treatment were never given (Landolt et al. 1984) and may be relatively arbitrary to the extent that the objective was mainly to achieve an extremely strict requirement.

However, for products where the pest prevalence is lower, the requirement of a probit 9 treatment may not be technically justified. For instance, a consignment of 50,000 units with a prevalence of 1% would require a probit 8 (0.998 mortality) for less then one survivor. If the pest presents a probability of 0.4 to survive transport conditions, of 0.2 to be transferred to a suitable host, of 0.5 to find favorable environmental conditions, and of 0.3 to reproduce in the host, then a treatment with a probit 6 (0.834 mortality) would be enough to avoid establishment. The requirement of unnecessary higher probit treatments could involve higher costs, higher residues, lower product shell-live, and a unjustified barrier to trade (see Follett and McQuates 2001 for an interesting discussion on these points).

3. A holistic approach

From the field in one country to establishment in another country there is a chain of events that reduces the pest prevalence in one consignment of a plant or plant product as illustrated in simplified form below.

Figure 2 Population losses in the production/commercialization chain



For phytosanitary purposes it is convenient to have a holistic approach to this system. For example, a post-harvest treatment with a known mortality rate has no meaning if we don't know the initial prevalence[a] in the treated consignment, because we don't know how many individuals will survive. Also, the design of an inspection system could be optimized only if we know the probabilities associated with pest entry and establishment. It is not logical to design a very sensitive, expensive inspection system to detect a very low prevalence if the pest has a poor chance to become established.

The establishment of a pest has essentially two components:

- the prevalence of a pest in the consignment that enters the country, and
- the probabilities of establishment of the pest.

The entry of a pest does not mean introduction (establishment). Many factors, as described in ISPM No. 11 (Pest Risk Analysis for quarantine pests) could interact to allow, or not, the establishment, as in the following examples:

- probability of pest surviving existing pest management procedures;
- probability of transfer to a suitable host;
- probability of the environment be suitable in the PRA area, etc.

It is also interesting to note the observations of Liebhold et al. (1995): "*Founder populations are typically small and consequently are at great risk of extinction. Generally, the smaller the founder population, the less likely is establishment. Though many scientists have referred to a "minimum viable population," there is rarely a distinct threshold. Instead it is more realistic to consider the probability of establishment as being a continuous function of the initial population size. This function reflects many characteristics of the species, such as its intrinsic rate of reproduction, mate location abilities, and genetic diversity*". This is in agreement with the statement of Dr. Alexei Sharov (Virginia Polytechnic Institute and State University-

[a] *prevalence* in this paper means number of individuals (or infected units) capable of reproduction and spread, and not rate of infection.

personal communication): *In many cases, the dispersal pattern and the probability of finding a mate are critical for pest establishment. Insects that mate before dispersal have a higher probability of establishment. In the destination area, the initial population numbers are extremely low. So if insects disperse first, then they will probably never find a mate. This is true even for insects with very sensitive pheromone communication mechanisms (e.g. gypsy moth).*

Equation 1 estimates the probable number of establishments:

EQUATION 1. *Establishments = Prevalence * Probability*

Where *prevalence* is the number of individuals (or infected units, etc) in the consignment that enters the country and *probability* is the total probability of establishments (multiplication of each specific probability or probability distribution).

From Equation 1 we can estimate the necessary prevalence to allow for one (1) establishment, in function of the probabilities. If the values of the three probabilities listed above are 0.5, 0.2 and 0.7, then it is necessary for 14 individuals to enter (1/(0.5*0.2*0.7)) in order to have one establishment. This is because 14*0.5*0.2*0.7 = 1. However, if the importation occurs, for example, in the winter, then the third probability could be 0.1 and then it would be necessary for 100 individuals to enter for one establishment (100*0.5*0.2*0.1 = 1).

We conclude that, to reduce the number of probable establishments we can either decrease the prevalence or decrease the probabilities of establishments. The phytosanitary measures can be applied at any point in the chain shown in Figure 2, usually before the entry of the consignment. Other measures could also by applied to detect infected consignments.

4. Phytosanitary Measures (PMs)

Now we can group the PMs, according to their strategy:

Group 1 Reduction of the population of the pest in the consignment (prevalence)

The strategy of this group is to reduce the population of the pest in the consignment and consequently, to reduce the possibility of establishment. This can be achieved by treatments or other procedures.

- Pre Harvest (treatment in the field, pest free area, place or site of production, testing, etc)
- Harvest (removal of infested products, inspection for selection, stage of ripeness/maturity, etc)
- Post Harvest (handling, chemical/physical treatment, etc)
- Shipping and distribution (volume[a], transport environment, in transit or on arrival treatment, limits on distribution, etc.)
- Pest Free Areas (sites and places of production)

Group 2 Reduction of the probabilities of establishment

The strategy of the second group of PMs is to reduce the probabilities of establishment and so reduce the possibility of establishment. This group includes measures using import management:

- Frequency of importation
- Season timing
- Port of entry
- Restriction on the end use

Group 3 Detection of infested consignments

The strategy of the measures of the third group is to detect the consignments that present infestation (prevalence) above an established threshold. After the detection, the destruction, re-export or another PM of the other groups (treatment, restriction of the end use) can be applied. The following measures fall into this group:

- Inspection
- Testing
- Post-entry quarantine.

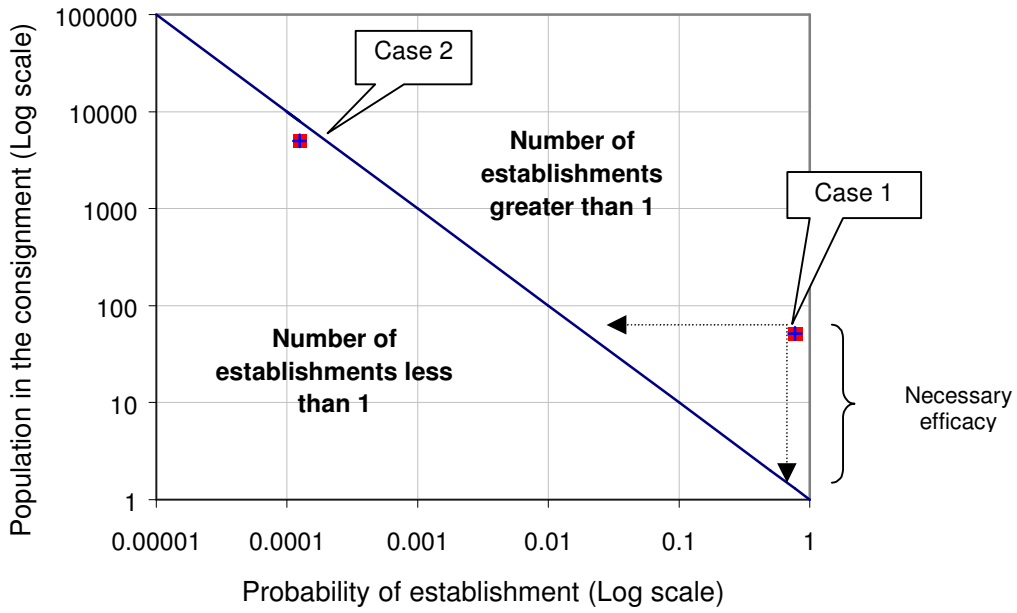
[a] Since the prevalence will be reduced proportionally with the volume of the consignment, a requirement for smaller consignments could be a valid and effective phytosanitary measure.

5. The meaning of efficacy

It is reasonable to consider that the PM should be effective to assure that the number of establishments of a pest derived from importation will be less than 1 in a certain period of time.

With equation 1 we can graph the "iso-establishments" line, where in the X-axis we have the probabilities of establishments and in the Y-axis we have the prevalence in the arrived consignment. The line represents the combination of prevalence-probability where the establishment is equal to one.

Figure 3 Iso-line for establishments
(NB: The line represents cases where the *prevalence* x *probability of establishments* = 1)



The graph in Figure 3 presents two different cases (dark squares). Case 1 represents a pest with low prevalence in the consignments (50 individual) and with a high probability of establishment (0.77). The efficacy of the PM has to be enough to move the point below the iso-line, either reducing the prevalence or the probability of establishment (or both). From equation 1 the prevalence has to be 1.3 (1/0.77) individuals or the probability of establishment equal to 0.02 (1/50). The efficacy required in the first case is 0.974 (1-1.3/50) and in the second is also 0.974 (1-0.02/0.77). So, for the PMs of the Groups 1 and 2 the efficacy is the same. For the PMs of Group 1 the efficacy represents a proportion of reduction in the prevalence (before the application of the PM). This is equivalent to the proportion of mortality (when the initial prevalence is known). For Group 2 the efficacy represents a proportion of reduction in the probability of establishment.

Case 2 represents a pest below the iso-line. Even considering that this pest has a much higher prevalence (5000 individual) it would not require any PM because the probability of establishment is very low (0.000125).

Equation 2 estimates of the necessary efficacy for Groups 1 and 2 PMs (proportion of reduction in the prevalence or in the probability of establishment) is:

EQUATION 2. $$NEfficacy = 1 - \frac{1}{Prevalence \times Probability}$$

Where *NEfficacy* is the necessary efficacy of the PM (proportion of reduction of the prevalence before the PM) and *prevalence* is the prevalence (number of individuals) before the application of the PM and *probability* is the probability of establishment.

For example, if we have a consignment of 50.000 units, with an infestation rate of 5%, and the pest has .02 probability of establishment, then the efficacy of a PM has to be greater than:

$$NEfficacy = 1 - \frac{1}{50.000 \times 0.05 \times 0.02} = 0.98$$

For the Group 3 PMs, it is necessary to consider the efficacy of the sampling and inspection system. The sampling model for inspection should be based in the hypergeometric distribution (see draft ISPM on Inspection Methodology). Transforming the Schilling (1968)[a] approximation to estimate the number of samples required to detect a known infection rate, we can get the proportion of a consignment that has to be sampled to detect a specific prevalence:

EQUATION 3. $$Np = Size \times (1 - (1 - CL)^{(1/prevalence)})$$

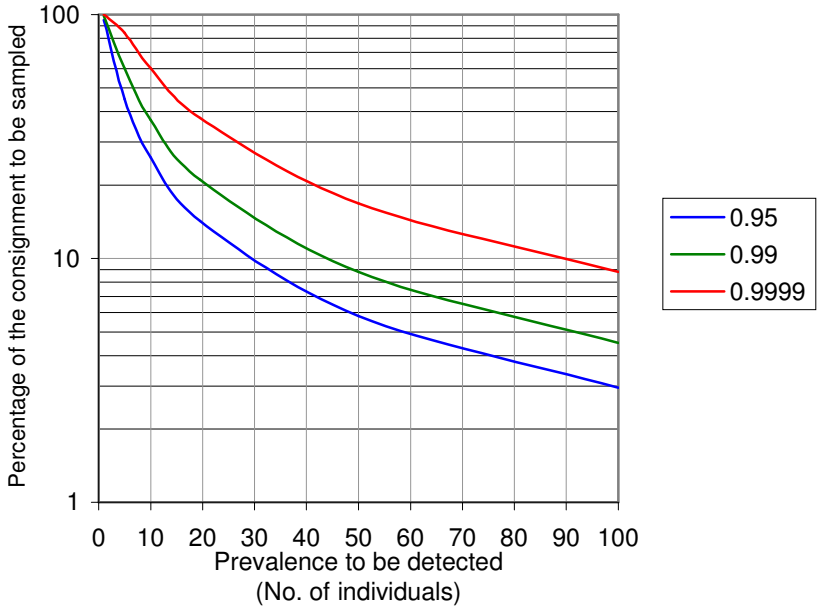
Where *Np* is the number of unities to be sampled, size is the size of the consignment, *CL* is the confidence level (0.95, 0.99, etc) and *prevalence* is the minimum number of individual to be detected.

Using the same numbers of the example above the minimum number of individuals to be detected is 50 (50.000 x 0.05 x 0.02). If we use a 0.95 confidence level then the number of units to be sampled is:

$$Np = 50000 \times (1 - (1 - 0.95)^{1/50}) = 2908$$

The graph in Figure 4 presents the relation between the prevalence to be detected and the necessary percentage of samples.

Figure 4 Percentage of sampling as function of the prevalence to be detected



For Group 3 PMs, the efficacy may be thought of as the proportion of consignments with prevalence above the fixed threshold that shall be detected. This corresponds to the confidence level of the hypergeometric sampling method, as presented in equation 3.

[a] **http://www-ist.massey.ac.nz/61325/sg/chap4-5.htm**

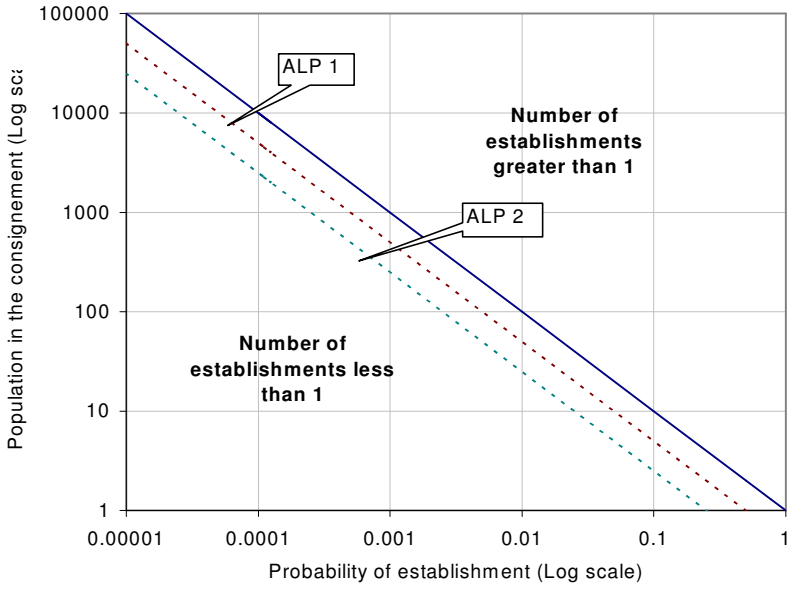
Summarizing we have:

- For PMs that reduces the population in the consignment (Group 1) the efficacy could be measured by proportion of reduction in the prevalence (or percent mortality, probit, etc)
- For PMs that reduce the probability of establishments (Group 2) the proportion of reduction could measure the efficacy.
- For PMs that detect infested consignments (Group 3) the efficacy could be the proportion of consignments that will be detected.

6. Appropriate Level of Protection (ALP)

The concept of ALP introduced in Figure 3 above can also be represented by the following graph:

Figure 5 Iso-lines for one establishment, considering the concept of Appropriate Level of Protection



According to the nature of the pest and/or the strategic importance of the crop, or other legitimate reason, a country may increase or decrease the level of protection desired. Many different parallel lines can be drawn as iso-lines to represent different levels of protection, as shown in figure 5 by the dashed lines. These lines actually represent different confidence levels and can be adjusted to meet the ALP requirements. The use of such presentations add transparency to the phytosanitary requirements and facilitates the choice of a PM based on efficacy.

7. Measurement of efficacy

7.1 General considerations

Prevalence

It is usually difficult to estimate the population size in absolute terms. In some cases, it would be possible to estimate the infection rate in the field and to multiply it by the estimated survival rate. The estimation of the rates should have sound statistical bases to provide reliable numbers for the mean values and variability. The most known and used statistical methods for estimations of this kind are based on the assumption that the population is normally distributed. However this is often not the true. In general, pests have a clustered spatial distribution and, in these cases, the appropriate statistical approach should be identified and applied. This is not a difficult procedure in statistical sciences.

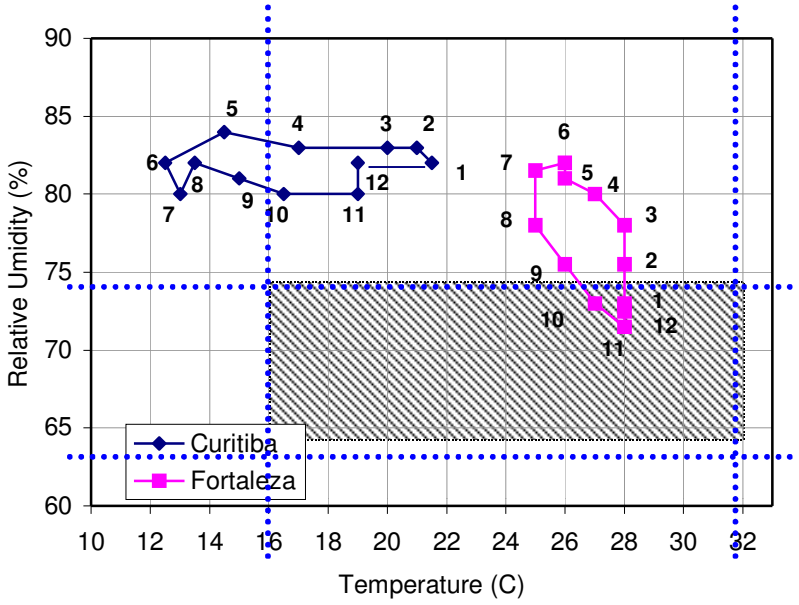
Chemical, heat, radiation etc, treatments; transport environment (cold, etc.)
The statistical methods for evaluation of the mortality rate of physical and chemical treatments are well known. For quarantine purposes, the probit analysis is the de facto standard. However the appropriateness of this technique for commodity treatment bioassay data has been strongly questioned. The number of pest individuals to be tested is proportional to the expected mortality rate. A probit 9 treatment will require a much larger population size (~100.000 individuals) than a probit 8. See Liquido et al. (1997) for a very interesting discussion on this matter, and Bartlett (1996) - pages 29-42, for the methodology and mathematical example.

Pest Free Areas (sites and place of production)
It is expected that the effectiveness of these methods will be close to 100%. However a well planned statistically-based sampling program for the commodity is useful to directly estimate the efficacy of the measure.

Probability of establishment
The measurement of the efficacy in reducing the probability of establishment, as discussed in point 4.2, should be examined with special care because this has not been well explored.

Some techniques used in the study of population ecology could be used to illustrate the possible effect of this kind of PM. As an example, Figure 6[a] below shows a monthly graph of the humidity versus temperatures in two locations in Brazil. The dashed square represents the zone of optimum development *for Ceratitis capitata*[b] (16-32°C and 65-75% RH) as found elsewhere in the literature.

Figure 6 Climogram (Temperature x Relative Humidity) for Fortaleza (red) and Curitiba (blue), Brazil and zone of optimum development for *Ceratitis capitata* (dashed square).



The graphic shows that in Curitiba the environmental conditions are sub-optimum all year around, while in Fortaleza there are favorable conditions from September to February.

The country could permit the importation entering in Fortaleza only from March to August, or only accept Curitiba as the entry port. These measures reduce the probability of establishment.

The graph above does not allow for any inference on the efficacy of the measure but the use of spatial models could be a tool to evaluate the reduction of the probabilities as function of environmental

[a] Araujo, K.R.P. 2000. Modelo matemático para simular a aplicação da técnica do inseto estéril e etapas de implementação de um programa de controle da mosca do mediterrâneo. PhD Dissertation, CENA/USP, 2001.
[b] *Ceratitis capitata* is not a quarantine pest for Brazil. It is used here only as an example.

conditions (see the electronic paper of Alexei Sharov[a]). Also, Dr. Sharov stated (personal communication): *Modeling the probability of population establishment is definitely a difficult task but to my mind not impossible. Estimating model parameters in other countries seems a reasonable strategy. Of course parameter values may be biased, but biased information is better than no information at all. There are various software that compare climate conditions in various parts of the world. Specialized natural enemies may be present in the country of origin but absent in the destination country.*

Another PM that could be used to reduce the probability of establishment is restriction on the end use. If we consider, for example, a virus in potatoes that will be planted, then the probability that the pest gets to suitable environmental conditions (soil) could be nearly 1 (100%). However if this commodity is used for consumption then the probability will certainly be much lower. The proportion of reduction in the probability could be used as a measure of efficacy, considering of course all the necessary statistics.

Detection

The efficacy of a detection system has two components: the efficacy of the sampling system (how well the samples represent the consignment) and the efficacy of the detection method (visual identification, testing, post-entry quarantine, etc.). As discussed before, the confidence level of the sampling method is important for measuring the efficacy of the sampling system.

The efficacy of the detection method could be estimated by the relation between the number of samples and the number of infested samples in one consignment. Efficacy is necessarily pest-specific because visual identification depends on the pest characteristics (size, symptoms, etc.). It is also important to consider that the efficacy will depend on the ability of the inspectors and therefore special care should be placed on the variability of the estimations. For laboratory testing, the efficacy of the diagnostic method, considering false-positive/negative response, etc., could be used.

7.2 Comparing the efficacy of different phytosanitary measures

a. Compare two different PM with the same endpoint

To compare different PMs with the same endpoint should not be complicated. To compare, for example, the efficacy of a fumigation treatment with the efficacy of a heat treatment it should be considered that, in both cases, the response (mortality) is a result of a dose. In fumigation, the dose is characterized by the concentration of the chemical and the period of exposure. As the pest would change its metabolic rate with the environmental temperature, the response to the same dose will change with the environmental temperature, and different tests should be made in different temperatures to consider this additional cause of variation. In heat treatment, the dose is characterized by the temperature and also by the period of exposure. The additional cause of variation, in this case, would be the size of the fruit, as larger fruits will take more time to transfer the treatment temperature to its interior and therefore different tests should be made with fruits of different sizes. As the concepts are the same, the efficacies may be directly compared by any suitable statistical technique for the comparison of means of different treatments. Note that most statistical tests will require that the data be drawn from the same population.

b. Compare one PM with two or more PMs (systems approach) with the same endpoint

This case is similar to the previous one. If the efficacy of the systems approach is measured after the application of the last PM of the system, then the data will reflect also the efficacy and variability of each of the PMs applied previously and the statistical comparison between the mean efficacies is equal to the previous case. However if the available data corresponds to the efficacy of each of the PMs that compose the systems approach, then the statistical requirements are slightly more complicated and have to take into account the propagation of errors, as discussed in section 8.2 below.

c. Compare two different PMs with different endpoints

It is useful to consider that there are only three PM endpoints: Group 1 - to reduce the prevalence in the commodity (see footnote on page 5); Group 2 - to reduce the probability of establishment; and Group 3 - to increase the detection capability, as discussed in points 4 and 5 above. For example, to compare the efficacy of a fumigation treatment that kills 97% of the population with a treatment that uses irradiation to sterilize 98% of the population, it is necessary to consider that, conceptually and for quarantine purposes, a sterilized pest individual is equivalent to a dead one and so the efficacies could be directly compared.

An additional consideration that could be raised at this point is that the right question seems to be where or when the efficacy should be measured. If it is decided to measure the efficacy of different PMs of group

[a] Modeling Forest Insects Dynamics. http://www.ento.vt.edu/~sharov/popechome/model/model.html

1 (see points 4 and 5) just before the commodity is delivered to the importing country (see Figure 2), with all the required statistical care, then it would be possible to have a reliable direct comparison between these efficacies.

Considering PMs that have very different endpoints, as the PMs of Groups 1, 2 and 3, it is not very likely that any of the measures of Groups 2 or 3 could substitute any of the PMs of Group 1. However, possibilities may exist which are worth exploring. Further discussion on this topic needs to be considered.

8. Error Analysis

(see the papers of Trinh,1996 and Muscat, 2001 for a deeper discussion on this subject.)

8.1 Standard Error and Confidence Level

When we need a quantitative estimation of, for example, pest prevalence in a consignment, we would sample the specific consignment with an appropriate sampling design. However, the result (the *mean* prevalence) is a description of the sample and does not necessarily represent the true population (consignment) mean. The *variance* is a measure of the dispersion of the data, or how the data spreads around the *mean*. The *mean* and the *variance* give us the "best" values and describe how the data is dispersed around the best value, respectively. To answer the question of how confident we can be that the sample mean is truly the best value, we need another statistic - the *confidence level*, as expressed in Equation 4 below. This is a common statistical application used to measure the confidence in the sample mean.

EQUATION 4. $$x \pm t * \sqrt{\frac{s^2}{n}}$$

Where x is the sample *mean*, s^2 is the sample variance, n is the number of units sampled and t is the student's *"t"* statistic. The values of x, s and n are estimated from the sample numbers and the value of t is taken from tables as function of $n-1$ and of the confidence limit (*cl*) (95 %, 99%, etc.). The square root of the *variance* divided by n is called *standard error*. A confidence interval is a range defined between a lower and an upper bound, which quantifies the percentage of time that this range contains the parameter of interest.

For example, if the *mean* prevalence (estimated from sampling 42 units) is 86.9 infested units, the *variance* is 95.06, the value of t (cl = 95%, n-1=41) is 2.02 then the *confident level* is 3.0 infected units. This is en expression of the *confidence interval* in absolute terms. This number means that we can be 95% sure that the true population *mean* is between 83.9 (86.9-3.0) and 89.9 (86.9+3.0). These intervals can be also expressed in *relative* terms (percent of the mean). In these case it would be expressed as 86.9 ± 3.5%

If we need greater confidence then we could choose the t value for 99% *cl*. In this case we are 99% confident that the true population mean would be between 82.8 and 91.0 (86.9 ±. 4.6%). In other words, if we sample the consignment 100 times we would expect that in 99 cases the sample mean would be between the values 82.8 and 91.0. However it should be noted that Equation 4 is only valid for normally distributed data. There are other methodologies to calculate confidence levels for other distributions.

Other methods to estimate the confidence levels of the efficacy of phytosanitary treatments is described by Baker et al. (1990), Couey and Chew (1986), Vail et al. (1993), and others.

8.2 Error propagation

The discussion on this point is based on the papers of Trinh (1996) and Muscat (2001).

If we need to estimate the confidence level for a variable calculated from the mathematical operation of other variables estimated from sampling, then we need to know how the error associated with each sampled variable propagate to the calculated variable.

The simplest approach is to calculate the square root of the sum of the squared individual errors. If the operation is an addition or subtraction then the *absolute* errors are used. In multiplication and division the *relative* errors are used.

Suppose that, in a systems approach, we need to estimate the probability (and the confidence level) for one infested unit in a consignment, after consideration of the mortalities due to a field treatment, due to post harvest handling and due to the transport conditions. Also suppose that we have the following information about the mean value of the proportion of the mortality (and/or reduction in pest prevalence) and the respective 95% confidence levels:

M1 (reduction due to field treatment)	= 0.9 ±. 0.05	(relative: ±. 5.56%)
M*2* (reduction due to post harvest handling)	= 0.7 ±. 0.07	(relative: ±. 10%)
M3 (reduction due to transport conditions)	= 0.5 ±. 0.20	(relative: ±. 40%)

The respective survivals S (S=1-M) are:
S1 = 0.1 ±. 0.05
S2 = 0.3 ±. 0.07
S3 = 0.5 ±. 0.20

The final survival (probability to have one infested unit) is:

$$S = 0.1 \times 0.3 \times 0.5 \pm \sqrt{5.56^2 + 10^2 + 40^2}$$

S = 0.015 ± 41.6% or
S = 0.015 ±. 0.00624

As discussed before, these number don' t mean much if we don' t know the prevalence in the consignment. So, after measuring the size of the consignment and the initial infestation rate, we could go further to add phytosanitary meaning to the mortality/survival rates:

C (size of consignment) = 5000 units ±. 50 units (±.1%)
I (infection rate (10%)) = 0.1 ±. 0.002 (±. 20%)
N (number of infected units) = C x I x S (note that N is the same prevalence in equation 1)

The number of units expected to be infected, after the effect of M1, M2 and M3 in C will be:

$$N = 5000 \times 0.1 \times 0.015 \pm \sqrt{1^2 \times 20^2 \times 41.6^2}$$

N = 7.5 ±. 44.8%
N = 7.5 ±.3.36 units

If the confidence levels were estimated within the 95% limit then we would be 95% sure the consignment will have between 4 and 11 infested units.

9. Equivalence of Phytosanitary Measures

The concepts presented below were completely borrowed from *Determination of Efficacy Product Evaluation* American Dental Association Council on Scientific Affairs 1999[a]. These could be useful to begin a discussion on the equivalence of PMs.

In comparing two different PMs, it could be demonstrated that one is *equivalent* or *at least as good as* the other.

9.1 Demonstration of *equivalency*

The equivalence of a given PM to another implies that the efficacy provided by those PMs are so close as to suggest that the two PMs could be used interchangeably without any meaningful effect on the outcome.

In demonstrating equivalence, a quantitative definition of what is meant by *closeness of efficacies* must be provided. This is done by specifying two percentages, L% and U%, where L% (for “lower”) is less than 100% and U% (for “upper”) is greater than 100% (together, these two percentages are sometimes referred to as defining the “range of equivalence”). The PM_2 is considered equivalent to the PM_1 if its observed mean outcome response lies within the range from L% to U% of the true mean outcome response for the PM_1.

[a] www.ada.org/prof/prac/stands/efficacyguidelines.pdf

For example, if L% is set equal to 90% and U% is set equal to 110%, then a PM_2 will be deemed equivalent to a PM_1 if the observed mean outcome response for the PM_2 is greater than 90% of the response for the PM_1 and less than 110% of the response for the PM_1.

The results of the study must support the inference that the *observed* mean score associated with the PM_2 lies between 90% and 110% of the *observed* mean score associated with the PM_1. Among the available statistic methods the Fieller Confidence Interval Approach and the "Two One-sided Tests" approach could be used.

9.2 Demonstration of *at least as good as*

To say that a given PM (PM_2) is *at least as good as* a another (PM_1), the benefit provided by PM_2 must be adequately large enough so that the switch from PM_1 to the PM_2 will not result in a meaningful loss of efficacy and, in fact, may enjoy a greater efficacy than provided by PM_1.

At least as good as might better be understood by considering its relationship to the property of *equivalence*. For equivalence of a PM_2 and PM_1, the mean outcome score for the PM_2 is not too much higher nor too much lower than that of the PM_1, that is, the observed mean responses are *close*. The property of *at least as good as* requires that the mean of the PM_2 not be too much higher than that of PM_1. It does not require that the mean of the PM_2 be not too much lower than the mean for the PM_1. Thus, this property requires that the observed mean score for the PM_2 does not vary too much from that for PM_1 *on the ineffective side*.

In demonstrating that a given PM2 is *at least as good as* a given PM_1, a quantitative definition of what is meant by *large enough benefits* must be provided. This is done by specifying a single percentage, which is designated as U% (note that U% is greater than 100%). The PM_2 is considered to be *at least as good as* the PM_1 if the PM_2 observed mean outcome response is no greater than U% of the observed mean outcome response for PM_1.

For example, if U% is set as 110%, then a PM_2 will be deemed as *at least as good as* a PM_1 if the observed mean outcome response for the PM_2 is no greater than 110% of the response for the PM_1.

The results of the study must support the inference that the observed mean score associated with the PM_2 lies at or below 110% of the observed mean score associated with the PM_1.

This criterion can be addressed in various ways. For example, the Fieller approach to equivalence described above can be invoked, with the criterion satisfied when the entire 90% Fieller Confidence Interval consists of values no greater than 110%. Alternatively, a single, one-sided test can be performed to determine if the observed mean score associated with the PM_2 is less than 110% of the observed mean score associated with PM_1.

10. Putting it all together

Conceptually the steps for deciding a phytosanitary requirement are very clear as shown in Figure 7, and presented below.

Considerations of the importing country:

1. Choice of an Appropriate Level of Protection
2. Determination of the final probability of establishment by risk assessment
3. Estimation of the maximum pest limit (prevalence) in the consignment (see figure 3 and 5). This estimation should consider the application of measures to reduce the probability of establishment, according to point 4.2.

The steps above correspond to processes that occur after a consignment arrives in the importing country (right side of Figure 2). Other steps are relative to the left side of figure 2, and are of concern for the exporting country:

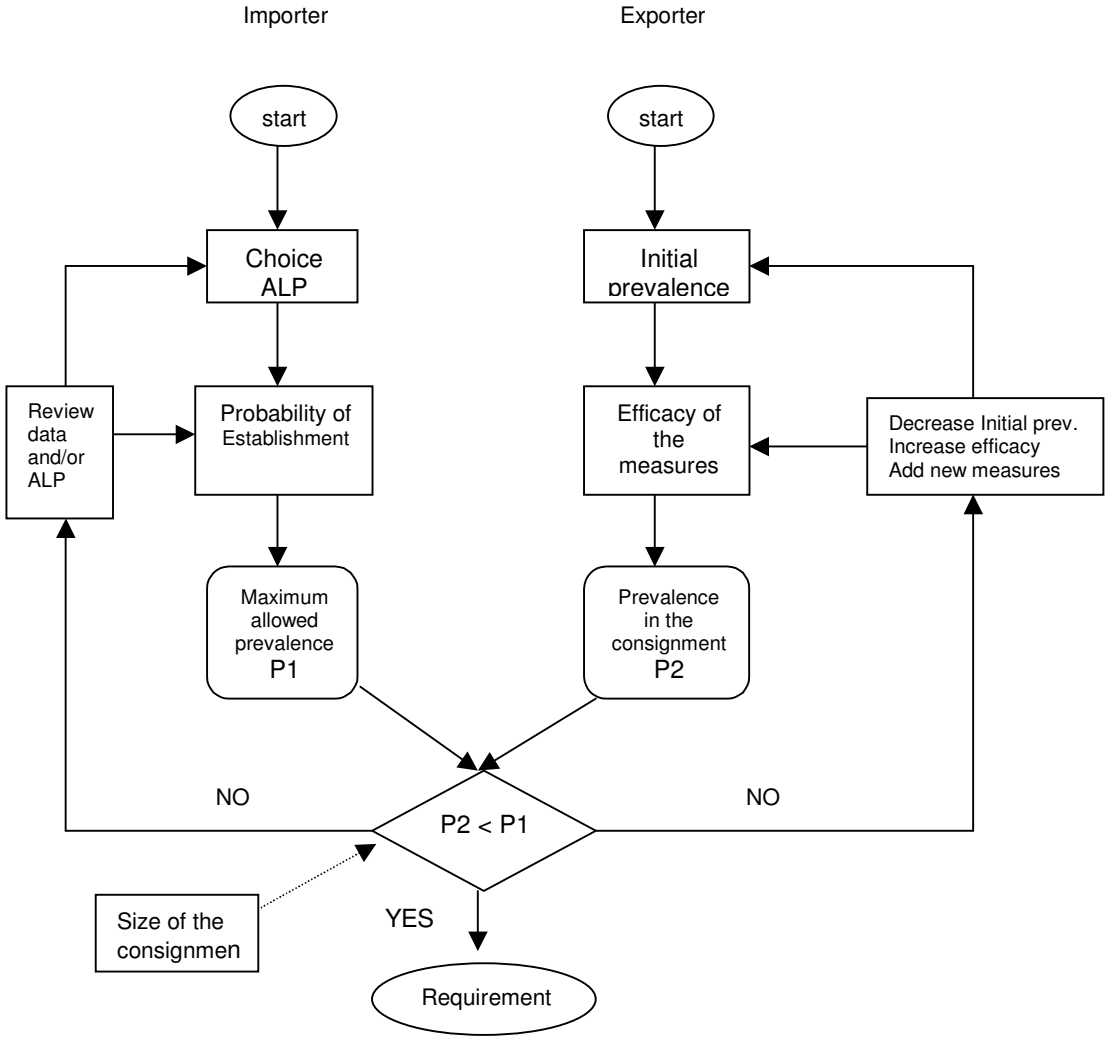
4. Estimation of the initial pest prevalence in the product

5. Estimation of the efficacy of the locally available phytosanitary measures
6. Estimation the pest prevalence in the delivered consignment

The final step is to compare the prevalence estimated in the steps 3 and 6, considering the volume to be traded.

If the prevalence estimated in point 6 is greater then the one estimated in point 3 then the importing country could review its data or even its ALP level. On the other side, the exporting country could increase the efficacy of any of the available PM already in the system to add a new PM or to try to decrease the initial prevalence.

Figure 7 Conceptual flowchart of the establishment of a phytosanitary requirement



11. References

Baker, R.T., D.S. Cowley, D.S. Harte, and E.R. Frampton. 1990. Development of a maximum pest limit for fruit flies (Diptera: Tephritidae) in produce imported into New Zealand. *J. Econ. Entomol.* 83:13-17.

Bartlett, P.W., G.R. Chaplin, and R.J. Van Velsend (eds.). 1996. Plant Quarantine Statistics. A Review. Proceedings of an International Workshop. Sydney, Australia - December 1995. HRDC, 93 pp.

Couey, H.M., and V. Chew. 1986. Confidence limits and sample size in quarantine research. *J. Econ. Entomol.* 79:887-890.

Follett, P. A.. and G. T. McQuates. 2001. Accelerated development of quarantine treatments for insects on poor hosts. *J. Econ. Entomol.* 94(5):1006-1011.

Landolt, P.J., D. L. Chambers, and V. Chew. 1984. Alternative to the use of probit 9 mortality as a criterion for quarantine treatment of fruit fly (Dipetra:Tephritidae) infested fruit. *J. Econ. Entomol.* 77:285-287.

Liebhold, A.M., W.L. Macdonald, D.Bergdahl, and V.C. Mastro. 1995. Invasion by Exotic Forest Pests: A Threat to Forest Ecosystems. *Forest Science Monographs* 30. 49 p.

Liquido, N.J., R.L.Griffin, and K.W.Vick (eds.) 1997. Quarantine security for commodities: current approaches and potential strategies. Proceedings of Joint Workshops of the Agricultural Research Service and the Animal and Plant Health Inspection Service, June 5-9 and July 31-August 4, 1995. USDA, ARS, 1996-04, 56 pp.

Muscat, A.J. 2001. A practical guide to the propagation and analysis of experimental errors. <**www.che.arizona.edu/Directory/Faculty/Muscat/ChEE202/Programs/error.pdf**>.

Vail, P.V., J.S. Tebbet, B.E. Mackey, and C.E. Curtis. 1993. Quarantine treatment: a biological approach to decision-making for selected hosts of codling moth (Lepidoptera:Tortricidae). *J.Econ. Entomol.* 86:70-75.

Trinh, P. (ed.) 1996. Error Analysis.
<**http://www.seas.upenn.edu/courses/belab/ReferenceFiles/stat.pdf**>