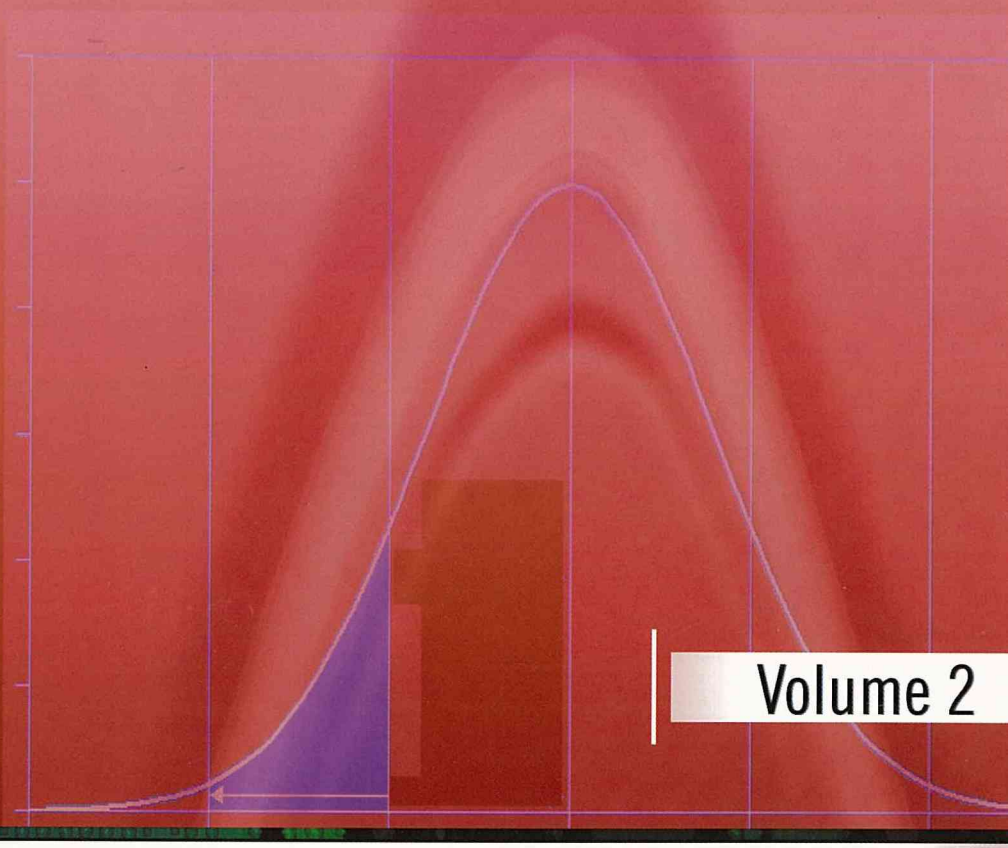
Organisation Mondiale de la Santé Animale
World Organisation for Animal Health
Organización Mundial de Sanidad Animal

Handbook on Import Risk Analysis for Animals and Animal Products



Volume 2

Quantitative risk assessment

Handbook on Import Risk Analysis for Animals and Animal Products

Volume 2.

Quantitative risk assessment

12, rue de Prony, 75017 Paris, France
Tel.: 33-(0)1 44 15 18 88
Fax: 33-(0)1 42 67 09 87
http://www.oie.int

ISBN: 92-9044-629-3
Vol. 1: ISBN 92-9044-613-7
Vol. 2: ISBN 92-9044-626-9

Cover photograph: © OIE

Contents

Author

Noel Murray
Ministry of Agriculture and Forestry
New Zealand

Editorial assistance

Stuart C. MacDiarmid
Ministry of Agriculture and Forestry
New Zealand
Marion Wooldridge
Veterinary Laboratories Agency (Weybridge)
United Kingdom
Bruce Gummow
Faculty of Veterinary Science
University of Pretoria
South Africa
Randall S. Morley
Canadian Food Inspection Agency
Canada
Stephen E. Weber
Centers for Epidemiology and Animal Health
Fort Collins
United States of America
Armando Giovannini
Istituto Zooprofilattico Sperimentale dell'Abruzzo e del Molise,
Italy
David Wilson
Head
International Trade Department
OIE, France

Acknowledgments

This second volume of the OIE ***Handbook on Import Risk Analysis: Animals and Animal Products*** draws heavily on David Vose's ***Risk Analysis: A Quantitative Guide*** (John Wiley & Sons, Chichester. 2000). Vose's book is an indispensable reference text for both the student and practitioner of animal health risk analysis. The text also draws on the book ***Import Risk Analysis: Animals and Animal Products*** (2002) by Noel Murray, published by the Biosecurity Authority, Ministry of Agriculture and Forestry, New Zealand.

Various people have offered critical comment on all or part of the modified text. In particular the Author and Editors wish to acknowledge:

David Vose
David Vose Consultancy
www.risk-modelling.com
Michael Roberts
Neil Cox
AgResearch Ltd
New Zealand
Lisa Gallagher
Tracey England
Louise Kelly
Rowena Jones
Veterinary Laboratories Agency (Weybridge)
United Kingdom
Sanping Chen
Carleton Quantitative Research
Ottawa
Canada

Anna Maria Conte
Istituto Zooprofilattico Sperimentale
dell'Abruzzo e del Molise,
Italy
Ziad A. Malaeb
Centers for Epidemiology and Animal Health
Fort Collins
United States of America

The Author and Editors thank Professor Vincenzo Caporale, Director of the **OIE Collaborating Centre for Epidemiology and Organization of Veterinary Services in Developing Countries**, Istituto Zooprofilattico Sperimentale dell'Abruzzo e del Molise, Teramo, Italy, for hosting their meetings and providing secretarial services during the drafting of the text.

Chapter 1
An introduction to quantitative risk analysis[1]

Introduction

In Volume 1 of this Handbook we stated that no single method of import risk assessment has proven applicable in all situations, and different methods may be appropriate in different circumstances[2]. In qualitative assessments, the likelihood the release and subsequent exposure to a hazard and the magnitude of the resulting consequences are expressed using non-numerical terms such as high, medium, low or negligible, and the qualitative approach has so far proved suitable for the majority of import risk assessments. However, in some circumstances it may be desirable to undertake a quantitative analysis, for example, to gain further insights into a particular problem, to identify critical steps or to compare sanitary measures.

The terms 'parameter', 'variable', 'input' and are often used interchangeably in quantitative risk assessments. In this Handbook, these terms are used as follows:

- **Parameter**

  In experimental statistics the term parameter represents a numerical descriptive measure that characterises a population, for example the population mean (μ), the population standard deviation (σ) and the binomial proportion (p). In spreadsheet computer software, it is often used to represent the arguments of mathematical, statistical or probability distribution functions such as the values required to define the shape of a Beta distribution or the mean and standard deviation of a normal distribution.

- **Variable**

  A variable is any characteristic that has a different value for different subjects or objects. If it can take on a different value as a result of a random process it is called a random variable. It can either be discrete, where it can only take on a limited number of values, or continuous, where it can take on any value within a given range. Examples of discrete variables include the number of infected animals, the number of test positive animals or the number of piglets in a litter, while examples of continuous variables include bodyweight or blood copper levels.

- **Inputs**

  An input is any information that is fed into a model. As a result parameters and variables, together with data and distributions, can be considered as inputs as they provide information that is used in a quantitative risk assessment model.

- **Model**

  A model is a simplified representation of the real world. Most models are symbolic because symbols represent properties of the system. In this handbook, a 'model' is a representation of an importation scenario in graphical or mathematical form where

---

1 The general reference for this chapter is Vose D. Risk Analysis, A Quantitative Guide. John Wiley & Sons Chichester, 2000

2 *Terrestrial Animal Health Code*, Article 1.3.1.1

equations are used to simulate the biological processes under study and the impact of risk management options.

– **Quantitative risk assessment**

A quantitative risk assessment is a mathematical model where the inputs and outputs are expressed numerically. In its simplest form, commonly referred to as a deterministic or point estimate analysis, both the inputs and outputs are expressed as single numbers or point values. These may represent a 'best guess', the 'average' or 'expected case' or perhaps the 'worst case'. When one wants to determine the impact of one or more of the input values on the output, one simply substitutes a new value into the model. This is effectively a 'what if', or scenario, analysis. For simple models with few inputs, this type of analysis can be easily undertaken using a calculator.

For more complex models, or in situations where one has more data to work with, probabilistic risk assessments are preferable. In these, inputs are described as probability distributions and a computer is essential for constructing the risk assessment model.

Deterministic (point estimate) risk assessment

Quantification of risk begins with considering an experiment, or trial with only two possible outcomes: success or failure. The trial may be repeated a number of times. For example, a trial may be a single embryo transfer from an infected animal to a susceptible recipient. A 'success' in this case would be where the infection is transmitted while a 'failure' would be a transfer where infection is not transmitted. If we observe no successes after ten transfers (trials) we may begin to suspect that the probability of transmitting infection by embryo transfer is low. As more transfers are undertaken without transmitting infection, the more confident we become that transmission is unlikely. This is shown in Table I, where confidence intervals[3] have been determined by consulting the statistical tables presented in Appendix 1.

Table I
Probability of transmitting infection following embryo transfer from a viraemic donor

Number of transfers (*n*)	Number of infected recipients I	Probability of transmitting infection $p_t = \left(\frac{r}{N} \times 100\right)$	Lower 95% confidence limit	Upper 95% confidence limit
10	0	0.00	0.00	30.85
20	0	0.00	0.00	16.84
30	0	0.00	0.00	11.57
40	0	0.00	0.00	8.81
100	0	0.00	0.00	3.62
1,000	0	0.00	0.00	0.37

If 100 experimental transfers were undertaken without transmitting infection, we could reasonably conclude, using the upper 95th percent confidence interval, that the probability

[3] A confidence interval is a range of numbers believed to include an unknown quantity with a specified level of confidence. For example, if we weighed 10 sheep we could calculate their average weight and the associated confidence intervals. If the average weight is 50 kg and the 95% confidence interval is ± 2.5 kg, this indicates that we could be 95% confident that the true average weight of all sheep in the flock lies somewhere within the interval bounded by 47.5 kg and 52.5 kg

of transmitting infection for each embryo transferred from an infected donor is 'at worst' 3.62%.

If we plan on undertaking an embryo transfer program we might like to estimate the probability that at least one recipient becomes infected or, alternatively, the average number of infected recipients we could expect.

To calculate the probability that at least one recipient becomes infected we proceed as follows:

- the probability of transmitting infection (a success) is p_t the probability of not transmitting infection (a failure) is $1-p_t$
- the probability that none of the recipients become infected is $(1-p_t)^e$, where e refers to the number of recipients (trials)
- so, the probability that at least one recipient becomes infected is $1-(1-p_t)^e$
- the probability is expressed in mathematical notation as $P(x \geq 1)$, where P refers to probability and x refers to the outcome, that is, an infected recipient
- and the final equation is then written as:

$$P(x \geq 1) = 1-(1-p_t)^e \qquad \text{Equation 1}$$

To calculate the expected number of infected recipients we multiply the probability of transmitting infection p_t by the number of recipients e:

$$\text{expected number of infected recipients} = p_t \times e \qquad \text{Equation 2}$$

If we assume a situation where the probability of transmission equals 3.62% (n=100) and the number of embryos transferred equals 30, we could determine the probability that at least one recipient becomes infected (Table II). For simplicity, we will assume that each recipient is implanted with only one embryo and that each donor produces a single transferable embryo. As a result the number of recipients equals 30.

$$P(x \geq 1) = 1-(1-0.0362)^{30} = 0.6692 = 66.92\%$$
$$\text{expected number of infected recipients} = 0.0362 \times 30 = 1.086$$

This scenario is essentially a 'worst case' as we have assumed that all the donors are infected. If we had some information on the prevalence of disease among the donors we could incorporate this into the model. Suppose a survey had been recently undertaken in a donor flock of sheep and 5 I animals out of 100 (n) tested were found to be infected. By consulting the statistical tables in Appendix 4 we could estimate that the true disease prevalence, with a 95% level of confidence, is likely to be between 1.64% (lower 95% confidence limit) and 11.28% (upper 95% upper confidence limit) with an expected value of 5%. We could include these estimates of disease prevalence in the model to determine three possible outcomes (Table II) using the following formulae:

$$P(x \geq 1) = 1-\left(1-p \times p_t\right)^e \qquad \text{Equation 3}$$

$$\textit{expected number of infected recipients} = p \times p_t \times e \qquad \text{Equation 4}$$

where: p = prevalence,
p_t = probability of transmitting infection and
e = number of recipients.

Table II
Probability of transmitting infection to at least one recipient and the expected number of infected recipients if thirty embryos are transferred

Scenario	p = prevalence in the flock of origin	p_t = probability of transmitting infection via embryo transfer	Probability ≥ 1 recipient infected (Equation 3)	Expected number of infected recipients (Equation 4)
Minimum	1.64% (lower 95% CL*)		1.77%	0.017 (17 out of every 1,000)
Most likely	5% (expected value)	3.62% (upper 95% CL)	5.28%	0.054 (54 out of every 1,000)
Worst case	11.28% (upper 95% CL)		11.55%	0.122 (122 out of every 1,000)

* CL = confidence limit

After considering the probabilities that one or more recipients would become infected, we might consider that the likelihood is too high and that some risk management measure is desirable. So, we might then decide to test the donors and discard any that are positive. If we test a potential donor, chosen at random, we could calculate the probability that it is infected D^+, given that it is test negative T^-. This is a conditional probability, which is expressed as $P(D^+|T^-)$. For a perfect test, this probability would be zero. However, since all tests are imperfect (with a sensitivity[4] of less than 1), we can expect that the test will fail to detect some infected animals. In addition, some uninfected animals will be incorrectly classified as positive, since the specificity[5] will also be less than 1. In these circumstances we calculate the $P(D^+|T^-)$, by firstly determining the predictive value of a negative test *NPV* as outlined in Chapter 4 and then calculate its complementary probability (1-*NPV*). This represents the prevalence of infection within the group of donor animals we accept. That is, the prevalence of infection amongst the test negative animals as a result of discarding test positive animals. From Equation 40 in Chapter 4 the *NPV* is calculated as:

$$NPV = P(D^-|T^-) = \frac{Sp(1-p)}{p(1-Se)+(1-p)Sp} \qquad \text{Equation 5}$$

where: p = the prevalence of infection in the flock of sheep
Se = test sensitivity
Sp = test specificity

So the prevalence of infection within the test negative group is calculated as:

$$P(D^+|T^-) = 1 - NPV \qquad \text{Equation 6}$$

If we use a test with a sensitivity of 90% and specificity of 99% and reject any positive animals, we could calculate the probability of infection for a test negative animal by substituting these values into Equation 6 (Table III):

[4] Sensitivity of a test is its ability to correctly classify an infected animal as test positive. It is calculated as the proportion of infected animals that yield a positive test result $P(T^+|D^+)$

[5] Specificity of a test is its ability to correctly classify an uninfected animal as test negative. It is calculated as the proportion of uninfected animals that yield a negative test result $P(T^-|D^-)$

Table III
Prevalence of infection among test negative donors

Scenario	p = prevalence in the flock of origin	Se = test sensitivity	Sp = test specificity	Prevalence among test negative donors (Equation 6)
Minimum	1.64% (lower 95% CL*)			0.17%
Most likely	5% (expected value)	90%	99%	0.53%
Worst case	11.28% (upper 95% CL)			1.27%

* CL = confidence limit

Since 1-*NPV* is the prevalence of infection within the test negative group, we can replace '*p*' in Equation 3 with '1-*NPV*' to determine the probability of transmitting infection to at least one recipient:

$$P(R^+ \geq 1) = 1 - (1 - (1 - NPV) \times p_t)^e \qquad \text{Equation 7}$$

where: R^+ = infected recipient
and the expected number of infected recipients:

$$(1 - NPV) \times p_t \times e \qquad \text{Equation 8}$$

The results of these calculations are shown in Table IV.

Table IV
Probability of transmitting infection to at least one recipient and the expected number of infected recipients if thirty embryos are transferred

Scenario	(1-NPV) = prevalence in the group of test negative donors (from Table III)	P_t = probability of transmitting infection via ET	Probability ≥ 1 recipient infected (Equation 7)	Expected number of infected recipients (Equation 8)
Minimum	0.17%		0.18%	0.002 (2 out of every 1,000)
Most likely	0.53%	3.62% (upper 95% CL*)	0.57%	0.006 (6 out of every 1,000)
Worst case	1.27%		1.37%	0.014 (14 out of every 1,000)

* CL = confidence limit

So, by making use of a statistical table and a calculator, we have been able to undertake a simple deterministic or point estimate analysis that has given us a very good idea of the risks we face. We could go on adding to this model, for example by including an estimate of the probability that a randomly chosen flock is actually infected and the effect of quarantining and testing recipients to screen out positive animals.

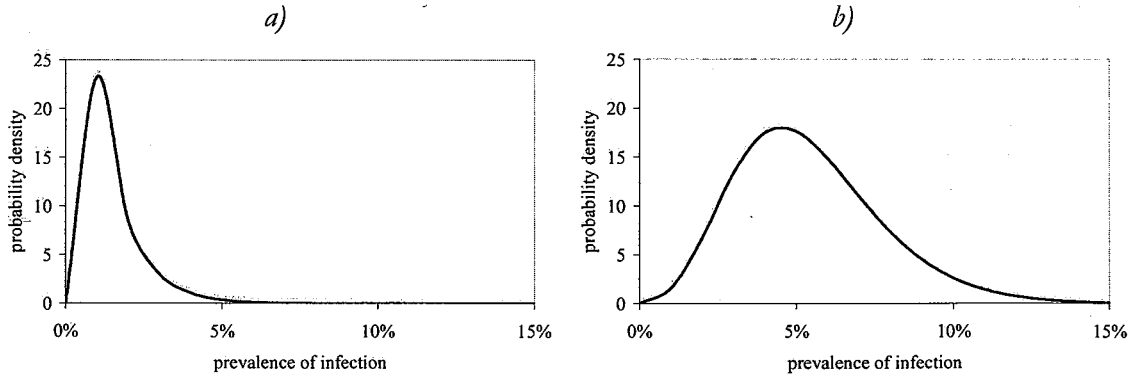
Probabilistic risk assessment (Monte Carlo simulation)

The embryo transfer model under discussion could be refined further. Just as we have estimated the probability of transmitting infection by embryo transfer, and the prevalence of infection within the flock of origin, we could include confidence intervals of the estimates of sensitivity, specificity and the probability that the flock of origin is infected.

However, as the number of such variables[6] increases there will be a rapid escalation in the number of potential combinations or 'what if' scenarios. For example, if we had four variables, each with a mean and upper and lower 95th percent confidence limits, we would have 3^4, or 81 possible scenarios. Such an approach has significant drawbacks. It can rapidly become impractical to analyse the results. In addition there is no weighting for each of the values chosen. For example, our 'best guess' might be far more likely to happen than the 'worst case'.

If we had information about the range of values and the likelihood of each value, we could assign a probability distribution to each variable, which we can now describe as random variables as they can take on a different value as a result of a random process. In our embryo transfer example we could use the Beta distribution (Chapter 4) to define a probability distribution for each input variable (Fig. 1). Such a model is called a stochastic model and we can calculate the combined impact of the variation in each of the model's input distributions to determine a probability distribution of the possible model outcomes. The simplest way to do this is to perform a simulation. This involves randomly sampling values from each distribution and combining the values generated, according to the mathematical logic of the model, to produce a result for that particular scenario. This process is repeated many times and the results from each scenario, which are also known as iterations, trials or realisations, are combined to produce a probability distribution of possible model outcomes.

Throughout this text, probability distributions will be described in terms of functions used in the risk assessment computer software @RISK[7] and the spreadsheet software Microsoft Excel[8]. For example, the notation Binomial( ) is an @RISK function while BINOMDIST( ) is a Microsoft Excel function and is distinguished by capital letters.



a) a Beta distribution of the probability of transmitting infection by ET if 100 transfers from infected donors to susceptible recipients were undertaken without transmitting infection: Beta (0+1,100-0+1)

b) a Beta distribution of the prevalence of infection if 5 infected animals were detected in a sample of 100: Beta(5+1,100-5+1)

Figure 1
An example of two probability distributions that could be assigned to the input variables in the embryo transfer quantitative risk assessment example

An ascending cumulative frequency plot (Fig. 2a) is often used to display the results of a simulation. It shows the probability of being equal to or less than a certain value. For

6 A variable is any characteristic that has a different value for different subjects or objects

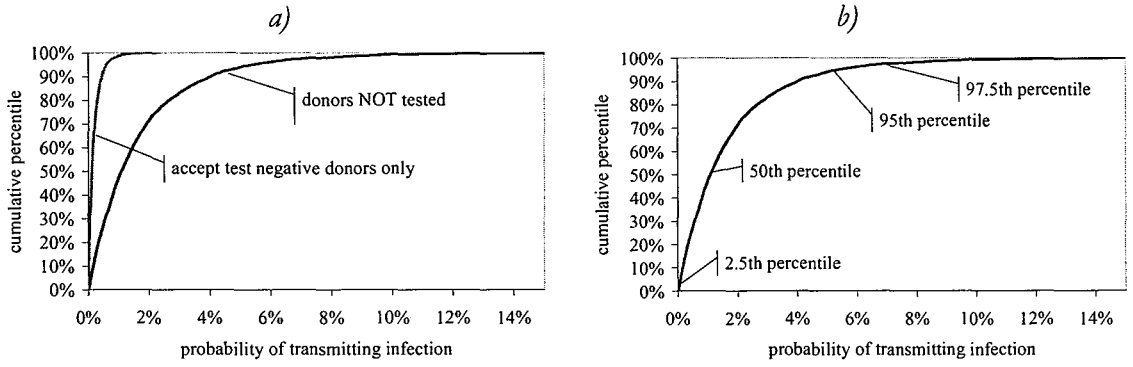
7 Palisade Corporation, Newfield, New York

8 Microsoft Inc., Redmond, Washington

example, we could report the results as follows, by reading from the 95th cumulative percentile:

In 95% of iterations, the probability of transmitting infection to at least one recipient is equal to or less than 5.4% if test positive donors are not rejected and less than 0.61% if test positive donors are rejected.

Alternatively, we might choose to report the median result (50th percentile) and the associated 95% confidence intervals. In the case of testing and rejecting positive donors the median is 0.12% with lower and upper 95% confidence limits of 0.004% and 0.8% respectively. It is important to note that the 95th percentile does not represent the upper 95% confidence limit. The upper and lower 95% confidence limits about the 50th percentile are represented by the 97.5th and 2.5th percentiles respectively (Fig. 2b). The area under the curve embraced by these percentiles is equal to 95% of the total area, which is the relevant area for the 95% confidence interval.



a) with and without testing donors

b) percentiles for the probability without testing

Figure 2
Ascending cumulative frequency plots of the probability of transmitting infection to at least one recipient if thirty embryos are transferred

Sampling values from a probability distribution

Sampling values from probability distributions is most commonly undertaken by either Monte Carlo or Latin hypercube sampling. The Monte Carlo method is based on simple random sampling from the entire distribution, which represents the sampling frame for each iteration. It is sampling with replacement, as it is possible for the same values to be selected more than once. Latin hypercube sampling, on the other hand, involves stratified sampling without replacement. The range of the distribution is divided up into a number of intervals, equal to the number of iterations to be performed and a simple random sample is then chosen from within each interval. Each interval is only selected once during a simulation. As a result, Latin hypercube sampling ensures that values from the entire range of the distribution will be sampled proportional to the probability density of the distribution. Fewer samples are usually required to reproduce the probability distribution so it is more efficient than Monte Carlo sampling for the same number of iterations. It is generally the preferred method of numerical simulation since fewer iterations are required for a particular level of accuracy.

Differentiating variability and uncertainty

The way in which variability and uncertainty have been described by risk analysts has led to a degree of confusion. To understand what is meant by these terms, it is important to appreciate that risk assessment is essentially a tool aimed at predicting the probability of an

outcome of a particular action or actions. For example, we might want to predict the likely height of a person chosen at random. We know from our own observations that there is a great deal of natural variation among individuals in the population. While we might have a good 'feel' for its range and what the average might be, it is only by measuring several people that we can begin to make some accurate predictions about the heights of people in the general population. As more measurements are collected, more knowledge is acquired. We can begin to describe the variation in people's heights with increasing certainty, enabling us to be more and more confident in our predictions. If we measured everybody in the population, we would have a perfect understanding and we would be able to state exactly what the population parameters, such as the average height and standard deviation (a measure of the amount of variation that exits), were. Obviously, this is impractical and we need to achieve a balance between acquiring perfect knowledge and obtaining reasonable estimates upon which we can base our predictions with a reasonable level of confidence.

Table V
A hypothetical example of the height of ten adults chosen at random and the associated statistics

Height in centimetres (x_i)									
152.3	118.4	158.5	168.8	163.4	162.9	180.7	99.5	188.9	198.5

Sample average = $(\bar{x}) = \frac{\sum_{i=1}^{n} x_i}{n} = 159.2$

Sample standard deviation = $s = \sqrt{\frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n}} = 30.3$

Standard error of the mean = $s_{\bar{x}} = \frac{s}{\sqrt{n}} = 9.6$

t value with $(n-1)$ degrees of freedom = 2.262 (from the student's t distribution)

Confidence interval = $\pm t \times s_{\bar{x}} = \pm 2.262 \times 9.6 = \pm 21.7$

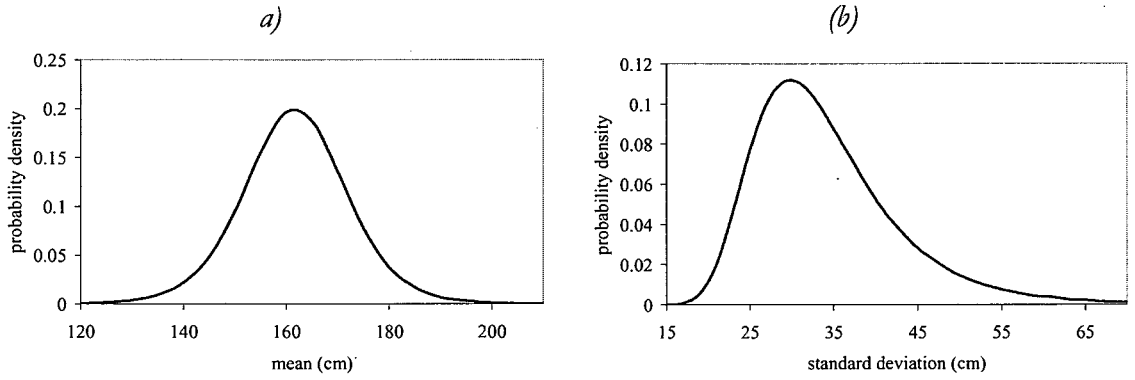
Upper 95% confidence limit = $\bar{x} + t \times s_{\bar{x}} = 159.2 + 2.262 \times 9.6 = 180.9$

Lower 95% confidence limit = $\bar{x} - t \times s_{\bar{x}} = 159.2 - 2.262 \times 9.6 = 137.5$

Note: sample statistics are represented by $\bar{x}$ (average) and s (standard deviation) while the corresponding population parameters are represented by μ and σ

If we choose ten adults at random and measure them, we can calculate their average height and standard deviation. These are actually *sample statistics*, rather than *population* parameters because we have collected data from a subset of the population only (Table V). If we deduce, from previous observations, that height is a normally distributed variable, we could use these sample statistics in a normal distribution function (Chapter 3) to enable us to describe the distribution of height in the general population and make some predictions. However, because of the small sample size we might be concerned that these sample statistics do not adequately reflect the population parameters. That is, the population parameters are uncertain. As shown in Figure 3 we could develop a sampling distribution for both the mean and standard deviation (see Chapter 6 for details). A sampling distribution enables us to capture the uncertainty associated with the estimate of a population parameter based on the data we have collected. For example, we can calculate

confidence intervals, which allow us to determine how confident we can be that the true population parameter lies within so many units either side of the corresponding sample statistic. Confidence intervals are determined from the area under the curve surrounding the average value of the distribution. The 95% confidence interval, for example, corresponds to ± 47.5% of the area under the curve either side of the average value. In our case the 95% confidence interval is ± 21.7 cm about the sample average of 159.2 cm (Table V). This indicates that we could be 95% confident that the true population average lies somewhere within the interval bounded by 137.5 cm to 180.9 cm.



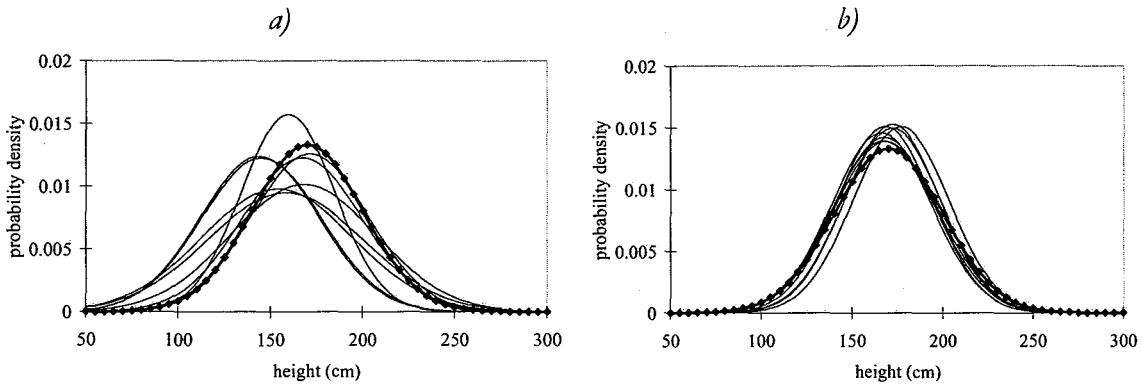
a) hypothetical sampling distributions of the mean

b) standard deviation

Figure 3
Hypothetical sampling distributions of the mean and standard deviation based on the data in Table V

If we randomly select a value from each sample distribution of the mean and standard deviation in Figure 3 and insert them into a normal distribution function, plot its graph and repeat this exercise a number of times, we could build up a picture of possible distributions of height (Fig. 4a). Each of these distributions separately represents a *first order distribution*, while together they form a *second order distribution.* These distributions, which enable variability and uncertainty to be modelled separately, are explored in more detail in Chapter 7. The thick black line in Figure 4a represents the hypothetical situation where we have perfect knowledge. It can be seen that there is a certain degree of uncertainty associated with the small sample size, because there are a number of different possible distributions.

What happens if we increase the sample size to 100 adults? By repeating the exercise just outlined, we can see from Figure 4b, that by collecting some additional information we have reduced the uncertainty considerably as the range of possible distributions is very close to the distribution representing perfect knowledge. We appear to have achieved a good balance between acquiring perfect knowledge and obtaining reasonable estimates.



a) sample size = 10

b) sample size = 100

Figure 4
A hypothetical normal distribution of the height of adults in Great Britain.
The thick line represents perfect knowledge where the average height of all adults is 170 cm with a standard deviation of 30 cm. Each thin line represents one possible distribution of height

Uncertainty, then, may be thought of as a measure of the incompleteness of one's knowledge or information about an unknown quantity. It is important to remember that even with perfect knowledge variability still exists.

As was observed in Volume 1, even though quantitative risk assessments involves numbers, they are not necessarily more objective, nor are the results necessarily more 'precise' than with qualitative assessments. Choosing an appropriate model structure, which pathways to include or exclude, the level of aggregation or disaggregation, the actual values used for each input variable and the type of distribution applied to them, all involve a degree of subjectivity. Further, because data are often lacking, models may need to incorporate expert opinion, which by its very nature is subjective.

The means by which this inherent subjectivity is countered in a good risk assessment is by ensuring that it is *transparent.* All the information, data, assumptions, uncertainties, methods and results must be comprehensively documented and the discussion and conclusions supported by a reasoned and logical discussion. The assessment should be fully referenced and subjected to peer review.