

ACERA Project		
2006 Round 1, Project 08		
Title		
Strategies for the effective communication of probabilities		
Author(s) / Address (es)		
Carlo Caponecchia, School of Safety Science, The University of New South Wales		
Material Type and Status (Internal draft, Final Technical or Project report, Manuscript, Manual, Software)		
Project final report		
Summary		
<p>The aim of this project was to review literature on improving the communication of probability information, focusing primarily on medical decision aids because they were known to use graphical presentations and are presented regularly to lay people who have to make difficult decisions. Recommendations for presenting probabilities in non-graphical ways were also examined from diverse domains.</p> <p>Examples from the New Zealand Apples Import Risk Analysis procedure were used to demonstrate how tools could be used in future IRAs. Recommendations included:</p> <ul style="list-style-type: none"> • Using natural frequency information; • Using natural frequencies together with basic graphical features; • Providing a “key” and/or visual representation of what a probability word is intended to mean at several points in the document; • Developing probability interpretation tools which could be used at stakeholder workshops, or as a general resource independent of IRAs; • Identifying “key” probability values that could be shown graphically. <p>The report provides examples and makes suggestions regarding how these communication tools might be tested effectively.</p>		
ACERA Use only	Received By:	Date:
ACERA Use only	ACERA / AMSI SAC Approval:	Date:
ACERA Use only	ACERA / AMSI SAC Approval:	Date:

FINAL REPORT

Strategies for the effective communication of probabilities

Dr. Carlo Caponecchia

School of Safety Science, The University of New South Wales

Prepared for ACERA and Biosecurity Australia

February 2007



UNSW
THE UNIVERSITY OF NEW SOUTH WALES

Acknowledgements

Many thanks to Prof Mark Burgman from the School of Botany, University of Melbourne and ACERA, for comments on drafts of this report.

Thanks also to Professor Jean Cross, School of Safety Science, UNSW; and Dr. Will Rifkin, from the Science Communication Unit, UNSW, for discussions on this report that assisted in its development.

Finally, many thanks to our research assistants, Ursula Sansom-Daly and Amy Datyner, who contributed so much to gathering and classifying literature throughout the project.

Table of Contents

ACKNOWLEDGEMENTS	3
TABLE OF CONTENTS	4
LIST OF TABLES	5
LIST OF FIGURES	5
1. EXECUTIVE SUMMARY	6
2. INTRODUCTION AND SCOPE	8
3. METHODOLOGY	10
4. ISSUES AFFECTING HOW PEOPLE INTERPRET PROBABILITIES	12
4.1 LOW PROBABILITIES AND RELATIVE RISKS	12
4.2 REFERENCE CLASSES	13
4.3 NUMERICAL VERSUS VERBAL PRESENTATION	14
4.4 APRIORI LIKELIHOOD JUDGEMENTS AND SELF-SERVING BIAS	16
4.5 PRESENTATION OF PROBABILITIES IN EXPERT GROUPS: NUMBERS AND NATURAL FREQUENCIES	17
5. METHODS FOR IMPROVING THE COMMUNICATION OF PROBABILITIES	18
5.1 DECISION AIDS IN MEDICINE AND HEALTH PSYCHOLOGY	18
5.2 EXAMPLES OF PROBABILITY REPRESENTATION IN DECISION AIDS AND THEIR EFFICACY	19
5.3 STANDARDS AND RECOMMENDATIONS FOR PROBABILITY COMMUNICATION	30
5.4 OTHER STRATEGIES TO FACILITATE THE COMMUNICATION OF PROBABILITIES.....	33
5.4.1. <i>Translating verbal probability terms into the same numbers across individuals</i>	33
5.4.2 <i>Geographical models of risk magnitude</i>	35
6. APPLYING ASPECTS OF THE LITERATURE TO BIOSECURITY	37
6.1 STAKEHOLDER COMMENTS ON THE DRAFT IRA	37
6.2 OTHER PARTS OF THE IRA DOCUMENTATION	40
6.3 GENERAL FINDINGS AND RECOMMENDATIONS.....	41
7. RESEARCH RECOMMENDATIONS	43
7.1 MAJOR RESEARCH CONSIDERATIONS	43
7.3 SUGGESTED METHODOLOGIES FOR ANSWERING THESE QUESTIONS	44
8. REFERENCES	46
9. APPENDICES	50
APPENDIX A	50
APPENDIX B	51

List of Tables

Table 2.1 <i>Subjective concepts of probability</i>	8
Table 3.1 <i>Search terms and databases used in this project</i>	12
Table 4.1 <i>Means and standard deviation of numerical probability ratings associated with particular terms from one condition of Sutherland et als., (1991) study (n=25)</i>	20

List of Figures

<i>Figure 4.1. Incidence and relative risk formats for low probabilities as used by Stone, Yates and Parker (1994).</i>	16
<i>Figure 5.1 Stick figures for relative risk comparisons as used by Stone, Yates and Parker (1997). Risk aversion increases in relative conditions with the addition of stick figures (Adapted from Stone, Yates and Parker, 1997).</i>	30
<i>Figure 5.2. Example of a systematic oval style presentation for the effects of warfarin and aspirin (from Man-Son-Hing et al., 1999).</i>	31
<i>Figure 5.3. Systematic (left) and randomized (right) arrangements for the communication of probabilities (from Schapira et al., 2001).</i>	34
<i>Figure 5.4. Example of one version of the Paling Perspective scale (from Paling, 2003)</i>	35
<i>Figure 5.5. Example of the Paling perspective scale showing relative risks (from http://www.riskcomm.com/scales.htm)</i>	36
<i>Figure 5.6. Example of the Paling perspective scale for smallpox showing consequence information (from http://www.riskcomm.com/scales.htm).</i>	37
<i>Figure 5.7. Example of a risk ladder with action standard and comparative risks, as used by Sandman and Weinstein (from Lipkus & Hollands, 1999)</i>	39
<i>Figure 5.7. Example of a risk ladder with action standard and comparative risks, as used by Sandman and Weinstein (from Lipkus & Hollands, 1999)</i>	41
<i>Figure 5.9. Example of Hoffman et als., roulette wheel arrangement for the presentation of relative risk in prostate cancer screening (from Hoffman et al., 2006).</i>	42
<i>Figure 5.10. Comparison of probability information presented as conditional probabilities and natural frequencies (adapted from Girgerenzer & Edwards, 2003).</i>	44
<i>Figure 5.11. Demonstration of numerical probability membership functions of different probability phrases (from Karelitz & Budescu, 2004).</i>	50
<i>Figure 5.12. Photograph of the 3-dimensional Lego model used to communicate nuclear waste volume and location (from Drew et al., 2003).</i>	53
<i>Figure 6.1 Representation of data from the AAAG's submission on the draft IRA New Zealand Apples in a frequency format.</i>	60

1. Executive summary

The aim of this project is to review literature on improving the communication of probability information. Research on medical decision aids were used as a primary source, because they were known to include graphical presentations of complex information, and are regularly presented to lay people who have to make difficult decisions. Recommendations for presenting probabilities in non-graphical ways were also examined from diverse domains.

Several issues that influence how people interpret numerical probabilities were identified in the literature, including:

- low probabilities tend to be treated as if they are equal to zero;
- low probabilities in relative risk formats tend to lead to risk aversion compared to numerical formats;
- specifying the reference class on the probability in question is extremely important to facilitate understanding;
- there is a large variation in how interpret probability words (i.e., the same word can be taken to mean a wide range of numerical values);
- self-serving bias can affect the way people interpret probability words; and
- natural frequency presentation (e.g., 8 out of 100 apples) greatly enhances the understanding of probabilities by both field experts and lay people.

Despite an abundance of research on the efficacy of medical decision aids in general, little research has concentrated on their graphical elements. Other more basic research on graphical presentation of probabilities suggests that the systematic oval technique is particularly useful (see Figures 5.2 and 5.3). Other techniques such as the “Paling” Perspective Scale and risk ladders with action standards have received little research attention, but may be useful in situations where low probabilities need to be presented in the context of the likelihood of other events. Mapping probability representations onto dartboards or roulette wheels has been suggested as a way to present the relative risks of two alternatives, but no empirical analyses are available. Frequency-style presentation is widely viewed as the most appropriate format for complex numerical probabilities, and can be used effectively alongside other graphical methods.

Instances where there was potential for misunderstanding probability information in the New Zealand Apples Import Risk Analysis procedure are used to demonstrate how tools from the literature could be used in future IRAs or other related documentation.

Recommendations included:

- Using natural frequency information where possible, especially in the introductory parts of the IRA (or other document), where probability values and calculations are highlighted;
- Using natural frequencies together with basic graphical features (such as that exemplified in communicating Bayesian analyses, see Appendix B);
- Providing a “key” and/or visual representation of what a probability word is intended to mean at several points in the document;

- Developing probability interpretation tools (including graphical representations and natural frequencies of relevant risks etc), which could be used at stakeholder workshops, or as a general resource independent of IRAs;
- Identifying “key” probability values that could be shown graphically (because it is inefficient to represent them all graphically).

Several recommendations are made regarding future research on the basis of the literature reviewed and the problems identified in the New Zealand Apple case study. These include the need to determine exactly what problems stakeholders have with biosecurity probability information, and how they may be rectified; basic research on how previous research on low probabilities might apply in the biosecurity context; the development of methods to assess probability understanding (as distinct from decision making); and the efficacy of the different graphical representations for biosecurity risk communication.

2. Introduction and scope

Several sources indicate how people have difficulty interpreting the meaning of probabilistic information (see Girgerenzer & Edwards, 2003; Osman & Shanks, 2005; Lipkus & Hollands, 1999). The term “probability” is itself interpreted in various ways, as outlined by Teigen (1994; see also Anderson, 1998; Peterman, 2004), and summarised in Table 2.1.

Probability information from Biosecurity Australia’s operations needs to be communicated to several different stakeholder groups. In addition to people’s general difficulty with processing probabilities, representatives of these groups may have little scientific training or experience with probabilistic information, so misunderstanding of this information may be a significant source of strain on time and resources, and retard the progress of risk analyses. While difficulties with probabilities are sometimes attributed to lack of experience or education, others argue that the problem lies in how information is presented, rather than in some internal cognitive problem (Girgerenzer & Edwards, 2003). This project seeks to understand some of the ways in which probabilistic information can be communicated in a way that maximizes its comprehension by a wide range of individuals. A commonly held view is that transforming probabilistic information into graphical representation can facilitate its effective communication. The starting point on graphical probability representations in this project was medical decision aids, because these were known to represent complex information in graphical ways, and it was known that considerable research on this topic would be available.

Table 2.1

Subjective concepts of probability

Probability concept	Definition/example
Chance probability	Independent events, such as coin tosses
Tendency probability	Dispositions or propensities of the phenomena, external to the viewer. E.g., describing the likelihood of finding a particular species in a location as “improbable”, because that species is disposed to living elsewhere; estimating probability in terms of how “easily” an outcome could occur
Knowledge probability	Applies to the range of hypotheses under consideration. Hypotheses that have not been thought of are assigned as knowledge probability of zero
Confidence probability	Degree of belief in a hypothesis
Control probability	Refer to the degree to which the probability of an event depends not only on chance, but on what someone does to influence the event. E.g., feeling safer as a driver than as a passenger in a car

Plausibility probability Refers to the believability of an argument/story/event. E.g., a scientific presentation may be assigned a high plausibility probability because it was believable, with an effective mix of data, anecdote etc.

Note. Adapted from Anderson (1998)

As outlined in project proposals, this project aims to:

- Identify some of the issues that affect the interpretation of probabilities;
- review existing and emerging approaches in the use of visual tools for decision-making in medicine, public health and psychology;
- assess whether and how visual decision tools as used in health and medical research can be used to assist with the communication of biosecurity information; and
- identify a range of circumstances in which communication of probabilities is an important element of DAFF/BA Department of Agriculture, Forestry and Fisheries/Biosecurity Australia) operations, and where such visual tools could therefore be employed.

Additionally, recognising that it was unlikely that the literature would have already investigated visual tools in specific biosecurity applications, the current project aimed to make recommendations for future research on adopting and evaluating visual decision tools for biosecurity risks.

The initial focus of this review was on how visual tools in medicine could inform the presentation of probabilities in biosecurity concerns. However, 2 factors discovered in the process of the review meant that the scope of the review needed to be broadened:

- (a) there is little specific research on how visual tools¹ for the communication of probabilities actually affect the understanding of probabilities (as opposed to affecting decisions, or levels of satisfaction with decisions); and
- (b) the existence of several recommendations for different ways of presenting probabilities, with a combination of numbers, words, and graphics.

The changes in scope of the project are a natural development of the exploration of the literature. The purpose the literature review is to find sources of information that are most likely to assist in managing the problem in question. Discussion is restricted to import

¹ This does not refer to graphs, which were considered outside the scope of this review.

risk analyses (IRAs), though is easily transferable to any situation in which similar types of probabilities are conveyed.

3. Methodology

To find literature relevant to this project, the psycINFO, PubMed, ScienceDirect, and Web of Science databases were searched. The search terms that were used are detailed in the Table 3.1 below. Additional materials were found by searching the webpages of Biosecurity Australia (<http://www.affa.gov.au/biosecurityaustralia>), the International Patient Decision Aids Standards Collaboration (<http://ipdas.ohri.ca/>) and the Ottawa Health Research Institute (OHRI) decision aid website (<http://www.ohri.ca/DecisionAid/>).

Once the database searches of psycINFO, PubMed, ScienceDirect, and Web of Science were completed, a two-step process was used to select relevant journal articles. Abstracts from the database search results were read, and based on what was relevant to the aims of the project, full articles/chapters were sourced and read. Criteria for relevance were decided collaboratively between the project team, and based on the themes outlined above that were the focus of the review. Articles were subsequently categorised into topic areas that could facilitate both a comprehensive response to the project aims, and an efficient writing process. In addition to searches performed in databases, articles that were cited in other articles were also obtained and used as appropriate.

Medicine was chosen as a domain from which to draw material because it is well known that people have to make complex medical decisions after consulting probabilistic information. It was known that strategies exist within medicine to represent probabilities in graphical ways, and so it was appropriate to focus mainly on medical literature. However, materials on the use of graphical probability displays in other domains were consulted as necessary, and significantly added to the project.

Table 3.1
Search terms and databases used in this project

Database	Date of Search	Search Terms
psycINFO	7/06	Biosecurity Biosecurity AND decisions Biosecurity AND risk Decision aids Decision tools Risk communication AND probability Risk communication AND decision aids Risk communication AND decision tools Visual decision tools

Database	Date of Search	Search Terms
		Visual decision aids
PubMed	7/06	Biosecurity Biosecurity AND decisions Biosecurity AND decision-making Biosecurity AND Risk Decision aids Decision tools Risk communication AND Probability Risk communication AND decision aids Risk communication AND decision tools Visual decision tools Visual decision aids
Science Direct	7/06	Biosecurity Biosecurity AND risk Biosecurity AND decisions Biosecurity AND decision-making Decision aids Decision tools Risk communication AND probability Risk communication AND decision aids Risk communication AND decision tools Visual decision tools Visual decision aids
Web of Science	7/06	Biosecurity Biosecurity AND risk Biosecurity AND decisions Decision aids Decision tools Risk communication AND probability Risk communication AND decision aids Risk communication AND decision tools Visual decision tools Visual decision aids

4. Issues affecting how people interpret probabilities

The following section reviews aspects of probabilities that can confuse people, and strategies to remedy this. It also reviews with how people deal with low numbers, and their resulting behavioural intentions. The motivational biases that may affect how probabilities are interpreted are reviewed, as well as a consideration of how experts are not immune to the problems of dealing with probabilistic information.

4.1 *Low probabilities and relative risks*

While interpreting probabilities in general can be difficult, dealing with low probabilities represents a particular problem. This is perhaps because people have little experience with low probability events, because by their nature, they occur infrequently (Stone, Yates & Parker, 1994). Kahneman and Tversky's (1979) prospect theory essentially proposes that there is a function by which probabilities are associated with decisional weights. At low probabilities, this function is discontinuous, suggesting that any probability below a particular low value is given no decisional weight (i.e., they are treated as essentially zero).

Stone, Yates and Parker (1994) examined how the framing of low probabilities affects their interpretation, and accordingly, people's behaviour on the basis of that information. Participants were presented with incidence rates and relative rates for the safety of particular products. Their willingness to spend more money on the safer product was used as an indication of how they interpreted the probability information. Figure 4.1 shows the information presented to participants. Overall, participants indicated that they would pay significantly more in the relative rate condition than in the incidence rate condition. This was interpreted by Stone, Yates and Parker as evidence for the "editing hypothesis" whereby people edit small probabilities to essentially nil (similar to Kahneman and Tversky's (1979) prospect theory). When the reduction is from .006 to .003 (a difference of .003) it is interpreted as low but significant, but when it is from .000006 to .000003 (a difference of .000003) it is interpreted as zero. Stone, Yates, and Parker's further experiments, where the potential for base-rate overestimation was accounted for, were consistent with the editing hypothesis. While they note the problems of only dealing with intended rather than actual behaviour, and the need to consider the ethical issues when using different frames to influence people's behaviour, the major implication of Stone, Yates and Parker's study is that conveying low probabilities in a relative risk format leads to risk avoidance compared to conveying the same information as traditional probability estimates.

Incidence rate format

<p>STANDARD TYRES</p> <p>Cost \$225 for 4</p> <p>Annual blowout injury risk (per Michigan driver)</p> <p>.000060 probability of a serious injury</p>	<p>IMPROVED TYRES</p> <p>Cost \$? for 4</p> <p>Annual blowout injury risk (per Michigan driver)</p> <p>.000030 probability of a serious injury</p>
--	--

How much would you be prepared to pay for IMPROVED tyres? \$.....for 4 tyres

Relative risk format

<p>STANDARD TYRES</p> <p>Cost \$225 for 4</p> <p>Annual blowout injury risk (per Michigan driver) about average</p>	<p>IMPROVED TYRES</p> <p>Cost \$? for 4</p> <p>Annual blowout injury risk (per Michigan driver) half that for standard tyres</p>
---	--

How much would you be prepared to pay for IMPROVED tyres? \$.....for 4 tyres

Figure 4.1. Incidence and relative risk formats for low probabilities as used by Stone, Yates and Parker (1994).

4.2 Reference classes

Gigerenzer, Hertwig, van den Broek, Fasolo and Katsikopoulos (2005) examined 750 people's understanding of probabilistic weather forecasts across 5 different international cities. A "30% chance of rain tomorrow" was interpreted incorrectly by many respondents as indicating that it would rain in 30% of the region, or that it would rain for 30% of the time, as opposed to the correct meaning, that it would rain on 30% of the days that are similar to tomorrow. Differences in experience with probabilistic weather forecasts (because probabilistic weather information has been used for different lengths of time in weather reports across the world) was shown to influence understanding of the information to only a small degree ($r=0.2$). Gigerenzer et al., indicate that people do seem to have a basic understanding of what a probabilities mean, in terms of 30% relating to 30 out of 100, but they fail to understand the reference class when single-event probabilities are presented. Defining the reference class is thus suggested by Gigerenzer et al., to facilitate understanding the information. An example would be to say, along with the prediction of a 30% chance of rain tomorrow, that this percentage does not refer to the amount of rain (the "amount" reference class), or the area in which rain will fall (the "area" reference class), but that on days like tomorrow, it will rain 3 out of 10 times.

The importance of reference class information is not limited to weather information. Gigerenzer et al., (2005; see also Gigerenzer & Edwards, 2003) give the example of a psychiatrist who prescribed Prozac to patients, and informed them that there was a 30-50% chance of sexual problems, such as impotence, associated with the drug. Consequently, many patients became anxious and did not want to take the drug. When the psychiatrist started explaining that 3 to 5 out of every 10 *people* to whom he prescribed the drug may experience sexual problems, patients' attitudes changed. The psychiatrist had specified the "people" reference class, whereas the patients had previously used a time reference class (thinking that something may go wrong on 30-50% of their sexual encounters), which is quite a different outcome.

Gigerenzer et al., (2005) work with weather forecasts also highlighted an important consideration regarding people's meta-cognition on probabilities. When given the opportunity to respond freely regarding what the forecast meant, many just restated the probability in a different way, and indicated that despite having a feeling of knowing what the probability meant, they were unable to explain it in words. The idea that people have strong intuitions about probability information, and that these can be difficult to challenge, is also supported by others writing about teaching statistics and probability in university courses (see Konold, 1995). This indicates a potentially important issue in communicating probabilities effectively in several domains, including biosecurity. If people think they know what probabilities mean, attempts to change their interpretation of probabilities may be met with opposition or lack of appreciation of the need to challenge their interpretation. The necessity of any project attempting to do so may be questioned, because people may be unaware of their limitations in processing probabilistic information adequately. This represents a key challenge for future research that may follow from this review.

4.3 Numerical versus verbal presentation

Because uncertainty is sometimes presented in a verbal way, several researchers have examined the way in which verbal probability expressions map onto numerical expressions of a particular probability. People tend to prefer to express uncertainty verbally (e.g., using the term "likely" rather than saying "75% chance"), possibly because it is sufficiently vague not to result in being questioned about it (Budescu & Wallsten, 1985, 1990 in Smits & Hoorens). At the same time, people prefer to receive information numerically, because it appears more precise (Smits & Hoorens, 2005). Accordingly, knowing whether describing something as "highly likely", or as having a "low probability" will result in consistent numerical interpretations across individuals is an important consideration when communicating such information. Most of the research shows that the same term is represented by a wide range of numerical values across individuals within particular domains or professions (Karelitz & Budescu, 2004; see also Budescu & Wallsten, 1990; Biehl & Halpern-Felsher, 2001; Gurmankin, Baron & Armstrong, 2004). Budescu, Weinberg & Wallsten (1988) found that 11 probabilities

were described by 111 distinct phrases. Other examples of this phenomenon include how people tend to misinterpret verbal terms describing the probabilities of side effects of medication on pharmaceutical labels and leaflets (Berry, Raynor, Knapp and Bersellini, 2004). Table 4.1 shows the large standard deviations in numerical estimates of probability expressions given by a sample of 100 cancer outpatients (Sutherland, Lockwood, Tritchler, Sem, Brooks & Till, 1991).

Table 4.1

Means and standard deviation of numerical probability ratings associated with particular terms from one condition of Sutherland et als., (1991) study (n=25)^{2,3}

Term	Mean numerical rating (SD)
Always	97.3 (3.9)
Usually	83.7 (10.9)
Common	76.4 (16.1)
Frequently	67.8 (20.5)
Likely	70.1 (25.0)
Possible	40.6 (32.0)
May	37.7 (23.9)
Less common	21.2 (16.9)
Occasionally	26.4 (20.9)
Small chance	13.6 (18.9)
Unusual	26.4 (20.9)
Unlikely	29.4 (33.0)
Very rare	18.1 (26.0)
Rare	24.6 (30.5)
Never	18.8 (36.2)

Note. Adapted from Sutherland et al., 1991

These findings have important implications given that probabilities are often accompanied by textual/verbal expressions which may mean different things to different people (even if guides are presented to reduce differences in interpretation, see also Section 5.4.1).

² Results were presented by Sutherland et al in this format for one condition only, which represents one of the 16 possible combinations of 2 interviewers, 2 recording methods (magnitude estimation vs linear analogue scaling), 2 formats (paper and pencil vs computerised), and 2 trials.

³ Some of the terms above could be argued to be frequency terms rather than probability terms per se, but nonetheless illustrate the point that there is large variation in proposed numerical equivalents of such terms.

4.4 Apriori likelihood judgements and self-serving bias

Given that people appear to have various numerical interpretations of verbally presented uncertainties, other research shows that this can be modulated by personal motivations and apriori notions about the likelihood of events. For example, based on the belief that in winter, common colds are likely occurrences, in the expressions

“This winter you will probably develop a common cold”

and

“This winter you will probably develop a lung infection”,

the terms “probably” tend to be assigned different numerical expressions (higher in the context of the common cold than lung infections; Smits & Hoorens, 2005). Thus, the apriori conceptions of the likelihood of events can influence how a verbal uncertainty term is evaluated. This has implications when using verbal terms in contexts where people are quite experienced and would be likely to have preconceived notions of likelihood: even if verbal terms are used consistently, they still may be interpreted in different ways.

Unrealistic optimism is the tendency for people to think that negative events are less likely to happen to them than to the average person of the same age and gender (see Weinstein, 1980; 1982; 1983; Weinstein & Klien, 1996; Job, Hamer and Walker, 1995), and has been found in relation to several health conditions (such as coronary heart disease, HIV/AIDS and cancer), as well as in relation to road driving (Job, Hamer and Walker, 1995), being injured following a natural disaster (Spittal, McClure, Siegert & Walker, 2005), and experiencing environmental hazards (Hatfield & Job, 2001). Building on the tradition of unrealistic optimism research, Smits and Hoorens (2005) investigated how people assign numbers to verbally presented uncertainties pertaining to themselves as opposed to others.

Accordingly, Smits and Hoorens examined whether particular verbal uncertainties for positive and negative events would be given different numerical probabilities. In 2 studies using health, study and general life events, as well as more severe health events, likely (“likely” and “probably”) positive events occurring to oneself were given higher numerical probability ratings than the same likely positive events occurring to others. Unlikely (“doubtful” and “small chance”) negative events occurring to oneself resulted in lower numerical ratings than unlikely negative events occurring to others. Smits and Hoorens refer to this as “motivated reasoning”, which may come from the threat implied to oneself that people wish to avoid (though there are several accounts of why people engage in this kind of reasoning in the optimism bias literature, see Lee & Job, 1995). These data suggest that people’s motivations can influence how they interpret verbal uncertainty terms, which again is an important consideration in communicating probabilities to groups or individuals with clearly vested interests, as would be the case with much of the probability information generated by Biosecurity Australia in import risk analyses.

4.5 Presentation of probabilities in expert groups: numbers and natural frequencies

While it is commonly held that lay people do not interpret probabilities correctly, there is evidence to suggest that experts in their field have similar problems which can be fixed by a change in presentation. Ghosh and Ghosh (2005) review several studies that indicate how medical students and practitioners have difficulty explaining, interpreting or calculating probability statistics specific to medicine (e.g., the positive predictive value of a diagnostic test, which indicates the proportion of people with a disease among those who get a positive test result). For example, 95% of medical officers surveyed gave incorrect estimates of the chances of a woman having breast cancer, on the basis of a positive test (because tests are not 100% accurate) when the information was presented in a probability format (see Eddy, 1982).

Changing the format of the probabilistic information has a significant impact on how it is used. Hoffrage, Lindsey, Hertwig and Gigerenzer (2000) found that both HIV/AIDS counselors, and law professionals and law students have difficulty when using probability information encountered regularly in their jobs/study. A confederate was sent to 20 public health centre HIV/AIDS counselors in Germany. 15 out of 20 counselors told the low-risk “client” that it is 99.9 to 100% certain that he had HIV if he tested positive. In reality, the probability of having the infection given a positive test is 50%. This observation shows a lack of understanding of the positive predictive value of the test. Hoffrage et al., showed that in advanced medical students, representing information in natural frequency format (i.e., X patients out of 100), resulted in vastly better estimates of the likelihood of a disease given a positive test compared to the information presented as numerical probabilities. Similarly, in the legal profession, Hoffrage et al., asked 27 legal professionals and 127 students to evaluate DNA evidence. When data on the certainty of the DNA match was presented in numerical probabilities, 13% of professionals and 1% of the students abstracted the correct probability of a match. When the data was presented as natural frequencies, however, the correct probabilities were inferred by 68 and 44% of the samples respectively. Further examples of natural frequency formats, and comparisons to traditional formats are shown in Section 5.3 and Figure 5.10.

From this evidence it appears that just because someone is experienced does not mean that they know how to deal with probabilities relevant to their profession in a consistent and accurate manner (see also Steckelberg, Balgenorth, Berger & Muhlhauser, 2004). The implication of this for biosecurity applications is that not only do we need to consider the communication of probabilities to lay stakeholders (such as industry groups, farmers, importers etc), but also to internal people routinely involved with all aspects of risk evaluation and communication. The use of natural frequencies in making probabilities easier to interpret also needs to be considered (see also Section 5.3).

5. Methods for improving the communication of probabilities

The literature discussed above has identified some of the issues that affect how probabilities are interpreted, and suggested some strategies for improved communication. The following section reviews decision aids and visual tools as used in medicine which have attempted to solve some of the problems in communicating probabilities, as well as other recommendations for improved understanding. Medicine was chosen as a model from which to draw conclusions because it is an area in which people have to make complex, high consequence decisions after consulting probabilistic information, and it was likely that research on ways to assist in this process was available.

5.1 Decision aids in medicine and health psychology

The move towards actively involving patients in choices about their medical care has necessitated the development of materials to assist them with such decisions (see Kaplan & Frosch 2005). The term “decision aid” is given to many different types of materials, but they are defined as “interventions designed to help people make specific and deliberative choices among options by providing information about the options and outcomes that is relevant to a person’s health status” (O’Connor et al, 2003, pp. 4). The communication of probabilities for each option a patient has (e.g., possible side effects, survival rates), is central to the main aims of most medical decision aids. Decision aids typically also include general information about the treatment or screening test for which they are developed; information and/or exercises to help the decision maker clarify their values (e.g., what is more important, X% likelihood of hair loss and Y% survival with treatment A, vs Z% likelihood of impotence and W% likelihood of survival with treatment B); and may include examples of what other people have decided based on their values and options. Medical decision aids vary in terms of media (e.g., information boards, pamphlets, web-based information, group presentations, interviews); content (types of information, risks treatment options explored, values-based exercises); and implementation (e.g., following or preceding consultation with a doctor) (see O’Connor et al, 1999; Feldman-Stewart, Brundage & Van Manen, 2001). Examples of complete medical decision aids can be seen in Appendix A.

The diversity of types of decision aids and the diseases and tests for which they are employed can make it difficult to evaluate their use. Deciding on criteria by which to evaluate them is also complex. In many cases there is no “right” decision by which to judge the performance of the aid, because the decision is based on personal values. The efficacy of these aids has been tested according to patients’ satisfaction with their decision; the degree to which they experience decisional conflict; their knowledge and understanding of their disease/condition; the impact of decision aids on health status and quality of life; and even the financial cost of administering the aids (see Charles, Gafni, Whelan & O’Brien, 2005).

While intuitively the concept behind medical decision aids appears to be beneficial, they are quite controversial. There is potential for them to be not comprehensive, just by not representing *every* possible side effect of a treatment, and conversely, they may be too complex if they represent combinations and permutations of possibilities (see Cantillon, 2004). Others have concerns about the criteria used to evaluate decision aids; the value judgements implicit in the information/content inclusions and exclusions (when they are supposed to help clarify a patient's values); and the questionable use of aggregate level probabilities applied to individuals (because the probability of a particular outcome for the individual in question cannot be known, see Charles et al., 2005).

Despite these concerns, the Cochrane review of randomised control trials on the efficacy of decision aids indicates that they are beneficial in terms of increasing patient's knowledge of the relevant issues, in reducing their decisional conflict, and in being satisfied with their decision (see O'Connor et al, 2003). An extensive review of the efficacy of decision aids is beyond the scope of this project, and is available in several other sources (e.g., the Cochrane review by O'Connor et al, 2003). What is relevant to this project, however, is specific evidence regarding the efficacy of the probability information provided. While it is concluded that the inclusion of "descriptions of outcomes and probabilities" were likely to result in more realistic expectations regarding treatment/test outcomes (O'Connor, 2003, pp. 15), an examination of the studies used to support this idea reveals that this effect cannot be specifically linked to the use of graphical probability displays per se (see Schapira & VanRuiswyk, 2000; Man-Son-Hing et al., 1999; McBride et al., 2002). Thus, one of the difficulties in assessing decision aids is establishing which part of them may have generated a benefit (if any), because they typically involve so many aspects across different media. There is limited evidence that isolates the effectiveness of graphical probability displays within medical decision aids. However, it may be possible to understand how graphical features facilitate the communication of probabilities from more general work on how graphical presentations assist understanding in other areas (see Section 5.2 below).

5.2 Examples of probability representation in decision aids and their efficacy

Several different types of visual presentations of probabilities exist, which will be briefly outlined here. Not all of the visual/graphical representations of probabilities have been in the medical sector. Just as there is a lack of specific research on the efficacy of visual depictions in medicine (see also Barratt et al., 2005), little research has been done on visual probability depictions in other domains⁴. What little research does exist is diverse in terms of hazards and methods, and has been performed somewhat atheoretically

⁴ There is substantial information on the best way to present information in graphs, though this does not always relate to probabilities. Accordingly, guidelines for preparing graphs, such as supplying tick marks, action point references, and maximizing the data-ink ratio, were not considered central to the aims of this project. More information can be found in Lipkus & Hollands (1999).

(Lipkus & Hollands, 1999), with researchers/practitioners choosing what types of graphics to use in a seemingly arbitrary manner.

One application of stick figures to communicate relative risks found that they were particularly effective with low probabilities. Stone, Yates and Parker (1997) presented the annual risk of injury from tyre blow outs with standard and improved tyres (see Figure 5.1, similar to their presentation as detailed in Figure 4.1). The presentation was relative in that both standard and improved tyres were shown with the number of injuries out of 5 million drivers. One condition showed risks in words and numbers, while the other showed the number of injuries (the numerator) represented with stick figures (the 5-million drivers frame, the denominator, was presented in words). When stick figures were added, people were willing to pay significantly more for a safer product, that is, they became more risk averse. Further studies showed that this effect did not occur for higher probability events (see Stone and Rush, 1997). Stone, Yates and Parker offer potential explanations for the graphical presentation of relative risk being associated with risk aversion, and were able to discount the idea that the effect could have been due to the humanising nature of stick figures (because the same effect was observed with asterisks). It could be that the graphics are just more salient than mere numbers, and thus more deeply encoded. Alternatively, the graphics may call attention to the fact that there is a meaningful difference between the 2 conditions (as opposed to just emphasising the size of the relative risk per se; Stone, Yates & Parker). However, why the effect occurs for low probabilities and not high probabilities is still unclear.



Figure 5.1 Stick figures for relative risk comparisons as used by Stone, Yates and Parker (1997). Risk aversion increases in relative conditions with the addition of stick figures (Adapted from Stone, Yates and Parker, 1997).

Feldman-Stewart and Brundage (2004) report that the styles of presentation of graphical information affect ease of understanding depending on the level of information intended for abstraction. In their studies with students and cancer patients, vertical bars resulted in the most accurate and fastest gross level judgements, while numerical presentation resulted in the most accurate comparisons at a detailed level. The arrangement of “systematic ovals”, where a number of ovals (similar to the dots or stick figures used in other work) corresponding to the denominator are presented, and the numerator is coloured-in, starting at the bottom left corner, were found to be the second most easily encoded and accurate presentation style for both gross and detailed level abstraction (see Feldman-Stewart, Kocovski, McConnell, Brundage & MacKillip, 2000; Feldman-Stewart and Brundage, 2004). An example of a systematic oval arrangement can be seen in Figures 5.2 and 5.3.

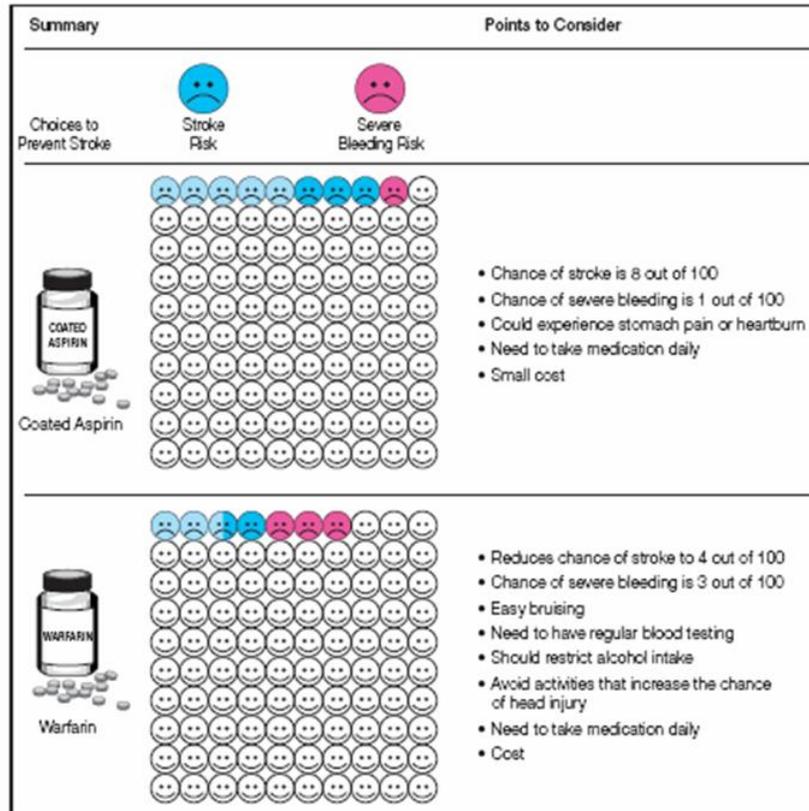


Figure 5.2. Example of a systematic oval style presentation for the effects of warfarin and aspirin (from Man-Son-Hing et al., 1999).

The other advantage of systematic ovals and other similar techniques is that they avoid framing effects. Framing refers to how numerical probabilities can be presented in a positive or negative context. For example, presenting the statement “there is a 94% chance of no side effects from treatment”, as opposed to presenting “there is a 6% chance of impotence associated with treatment” can have major effects on how people interpret the risk and behave accordingly. As Barratt et al., (2005) note, the systematic ovals technique avoids the potential for this because it simultaneously presents both probabilities (e.g., one being coloured-in, and one not being coloured-in). Systematic ovals remove the need for the reader/decision maker to do mental arithmetic to calculate the percentage of the alternative not presented (e.g., 100-94 in the example above), which may become taxing if several probabilities are presented. Similarly, the use of natural frequencies for both the positive and negative frames could lead to information overload (Barratt et al.). Systematic ovals and other similar techniques thus appear to be both an effective and economic mode of conveying probability information.

In a qualitative design, Schapira, Nattinger and McHorney (2001) examined participants’ perceptions of various graphical and numerical probability displays, including the systematic oval style. Graphical displays of stick figures coloured-in to represent a probability were rated as easy to identify with, and understandable, and the context they

provided made them seem meaningful, as opposed to bar graphs which were perceived as analytical, difficult to understand and as having less impact. However, bar graphs were preferred for comparing multiple risks in a simple manner. When displaying larger denominators (1000 compared to 10 in previous presentations), with coloured and non-coloured faces, participants found abstracting the risk magnitude to be more difficult in the random highlighting method, where figures are coloured-in at random throughout the display, as opposed to the systematic method, where highlighted figures are grouped together (see Figure 5.3).

How systematic ovals and similar arrangements work for low probabilities, however, is an issue that requires further consideration. Graphically representing the numerator for low probabilities, such as Stone, Yates and Parker (1997) did, removes the need for showing large numbers of symbols representing the denominator (which would be impractical to do), but may also bias results⁵. This is something that requires further research, possibly by piloting whether and how people understand low probabilities in numerical and graphical presentations, by getting them to rank the level of likelihood implied by each display. Such an arrangement would remove the aspect of behavioural intention, focusing rather on comprehension of the probability represented, and would also remove the demand characteristics inherent in Stone, Yates and Parker's (1999) design (i.e., the imperative to "buy" a safer product).

⁵ Some studies have presented tens of thousands of dots representing the denominator across several different pages, though having to present them on separate pages appears to weaken their effects (see Lipkus & Hollands, 1999).

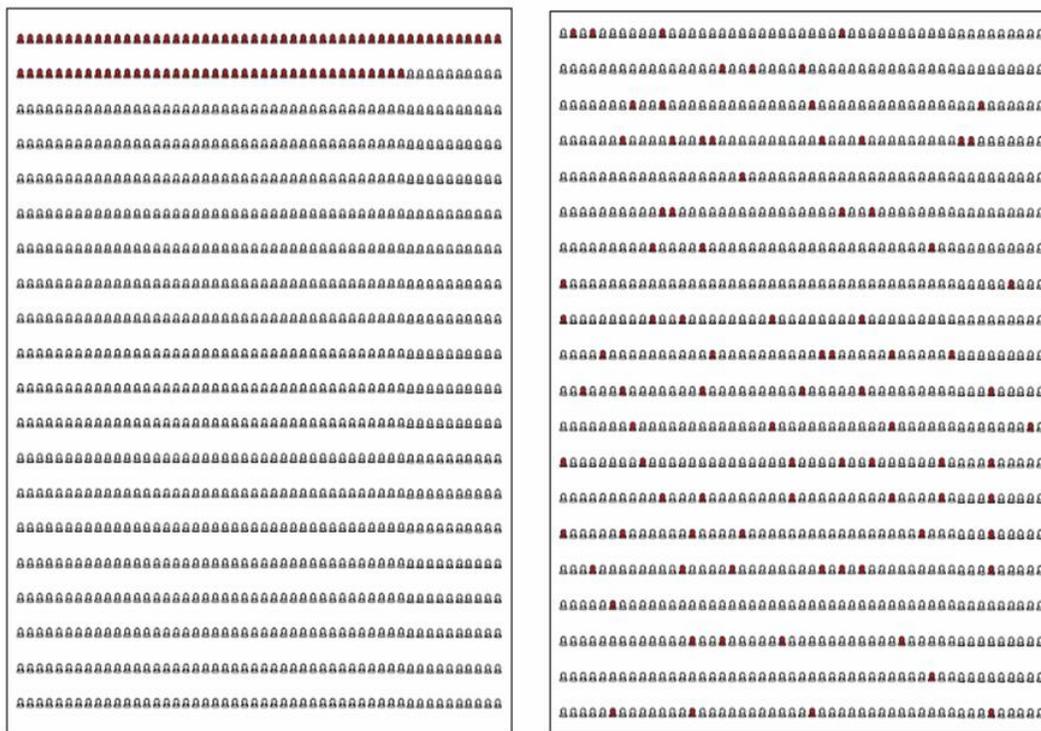


Figure 5.3. Systematic (left) and randomized (right) arrangements for the communication of probabilities (from Schapira et al., 2001).

Variations of the “Paling” Perspective Scale may be useful as an alternative to the systematic oval style presentation for presenting low probabilities. Paling’s scales (see examples in Figures 5.4 to 5.6) present a range of likelihoods, from certain to extremely low, with text and numerical descriptions of risk along the scale. Risks that a particular group may already be familiar with can be placed on the scale to provide comparison. In evaluating the efficacy of the “Paling” Perspective Scale for communicating blood transfusion risks, both written risk formats (in natural frequencies, i.e., 1 in 250,000) and the Paling Perspective Scale improved understanding, but no difference between the formats was observed (Lee & Mehta, 2003). However, it is not clear to what extent Lee and Mehta tested the understanding of probability information per se (e.g., by asking participants to put different probability presentations in rank order), because their focus was more on perceptions of risks in blood transfusions. Further, whether the effects of written versus graphical probability information are retained over time, and whether the effect of them combined is greater than either alone was not addressed, and both of these issues are important for future research.

Helping the Public Understand Risks From Water

Making Sense of the Regulators' Figures to Explain Water Safety

Lifetime risks of death at levels permitted by EPA

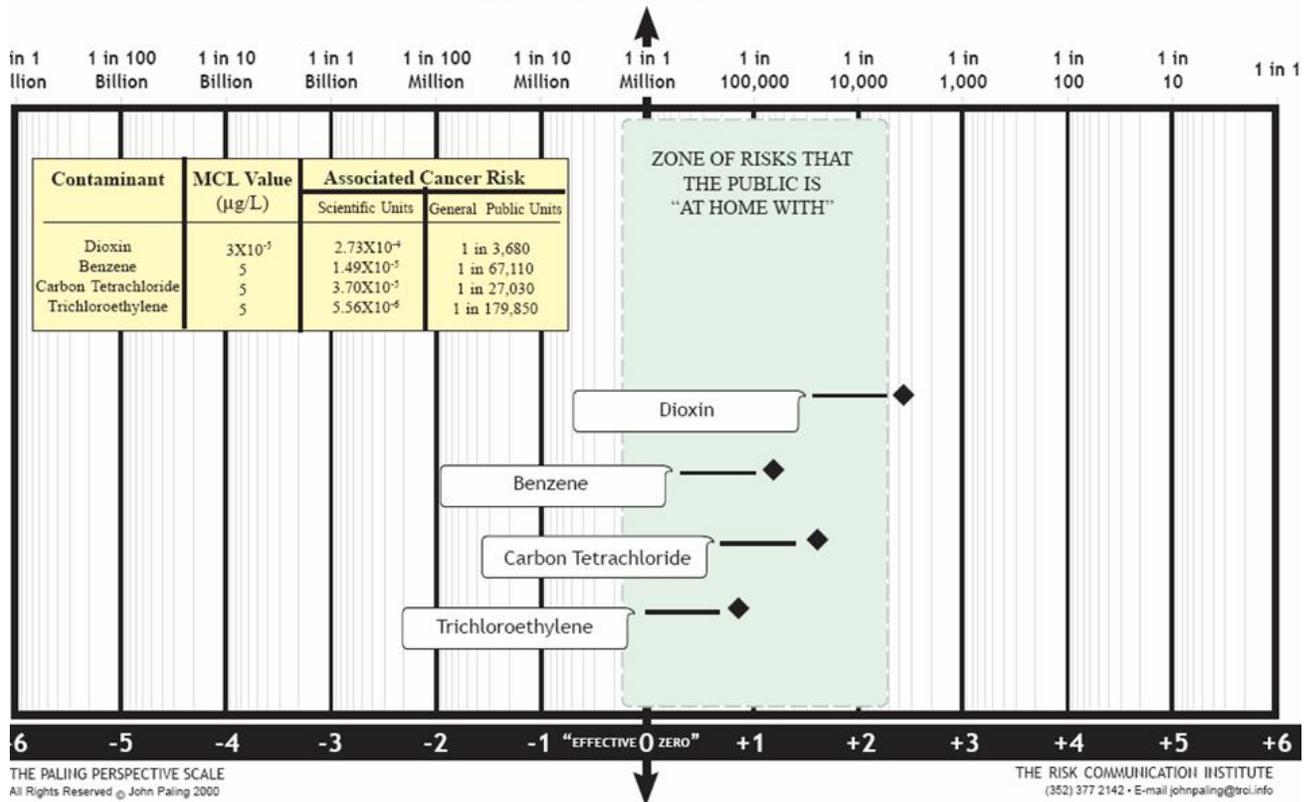


Figure 5.5. Example of the Paling perspective scale showing relative risks (from <http://www.riskcomm.com/scales.htm>)

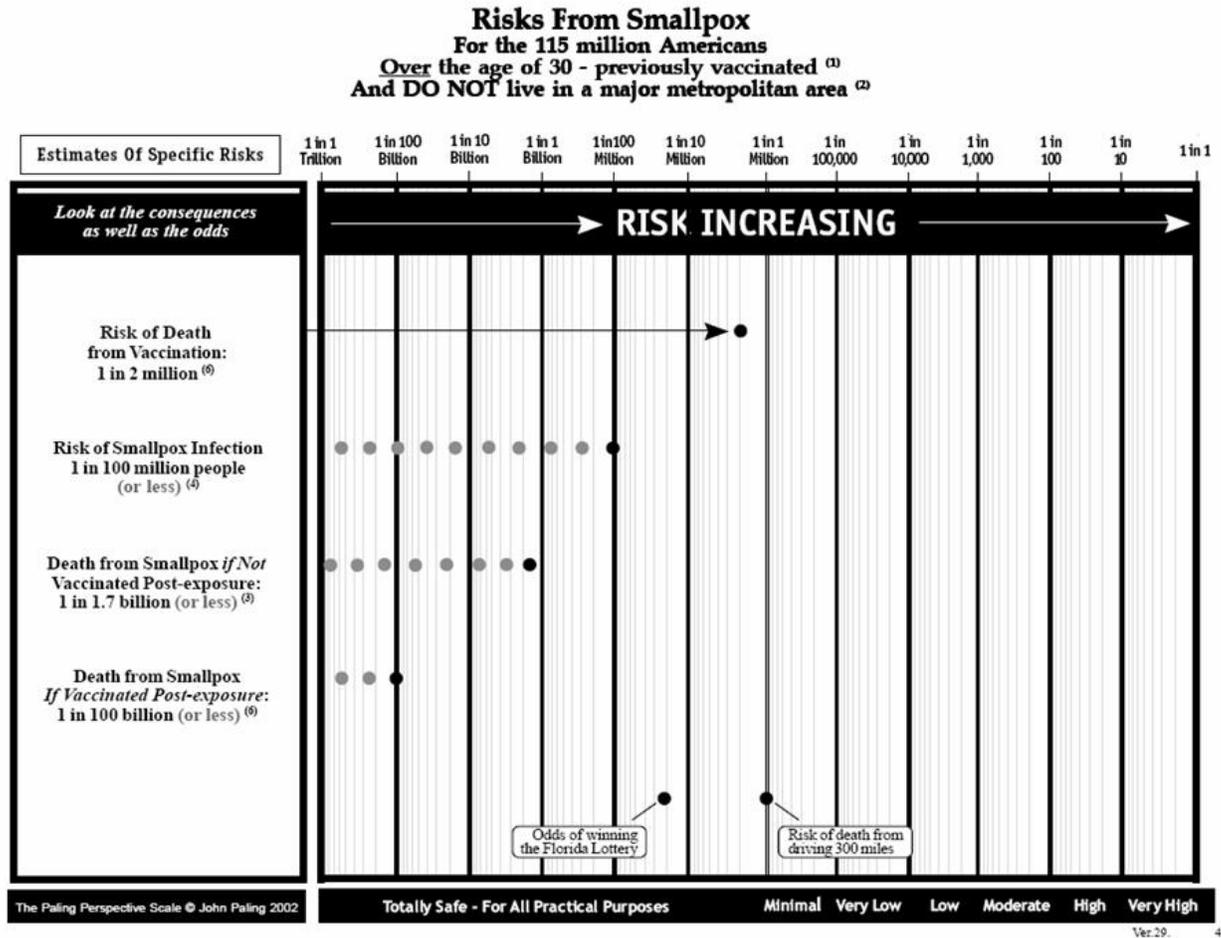


Figure 5.6. Example of the Paling perspective scale for smallpox showing consequence information (from <http://www.riskcomm.com/scales.htm>).

In some ways, the Paling perspective scales are similar to the risk ladder technique, whereby a range of risks are presented in ascending order. The risk ladder has been used to communicate risk magnitude, for example, asbestos and radon exposure risk, and to indicate particular levels at which exposure is likely to become a problem (Lipkus & Hollands, 1999). From their series of studies performed for the American Environment Protection Authority, Sandman and Weinstein (1994) summarise their main conclusions when investigating various risk ladder presentation formats. They used hypothetical data on exposure to risks such as radon and asbestos and their impact on lung cancer. People felt more “at risk” when presented with only an action standard, telling them what to do with a particular level of risk, compared to when the action standard was presented at the mid-point of a risk ladder, in the context of other risks (see Figure 5.7). Sandman and Weinstein suggest that the risk ladder acts as an anchor, against which relative levels of risk are judged. People are thought to be reassured that things could be worse. Sandman and Weinstein recommend the use of the risk ladder in situations where the communicator wishes to avoid panic. Action statements alone, without a risk ladder

appear to generate risk aversion (which could be useful in some situations (e.g., if you want to encourage people to protect themselves from UV rays). In addition, the position on the ladder at which the risk is placed was found to have a significant effect on perceived risk, equivalent to a 10x difference in actual risk. Placing a risk high or low on the page can thus influence the degree to which people perceive something to be a risk. Providing comparative risks for something familiar, such as cigarette smoking made people think that the information was more helpful and that they understood it better, however, this had no effect on their ability to distinguish between high and low risks nor between probability estimates.

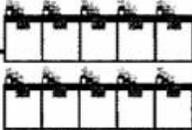
Cancer Deaths from Lifetime Radon Exposure			
Radon Level (pCi/L)	Extra Cancer Deaths (out of 1000 people)	Equivalent Smoking Risk	Advice
100	500 in 1000	 10 Packs per day	HIGH TO VERY HIGH RADON LEVELS Measurements near the upper end of this range are much higher than the EPA action guideline. Exposure to such levels is very dangerous. For residents living in homes at the higher end of this range, action should be taken within the next couple weeks to substantially reduce their exposure. If prompt action is not possible or is not effective, they should consider moving until the radon levels are reduced. Exposure to levels at the lower end of the range is also unsafe. Residents living in homes at these levels should act to reduce the readings within the next couple of months.
40	200 in 1000		
20	100 in 1000	 2 Packs per day	
10	50 in 1000		MODERATE TO HIGH RADON LEVELS Measurements in this range are above the EPA action guideline. Exposure to these levels is a significant risk if it extends over many years. Residents should carefully evaluate the causes of their elevated levels and make plans to reduce the levels permanently. To minimize the cumulative risk, this permanent action should be completed in the next year or two. In the meantime, residents may want to avoid prolonged exposure to areas of the home where the levels are highest.
4	20 in 1000	 8 cigarettes per day	
2	10 in 1000		LOW TO MODERATE RADON LEVELS Measurements in this range fall below the EPA action guideline. Radon levels at the lower end of this range present a low health risk. Radon levels at the higher end of this range, extended over a lifetime, present a moderate health risk. Any plan to lower the levels should be carefully evaluated to be sure that it is likely to be effective, since it is often difficult to reduce levels below this range. Many authorities do not recommend trying to reduce levels in this range, especially for homes near the lower end. Residents who decide to try to reduce their levels below this range can take several years to act without adding significantly to their risk.
1	5 in 1000	 2 cigarettes per day	
0.5	2.5 in 1000		VERY LOW TO LOW RADON LEVELS Measurements in this range are no higher than the outdoor "background" level in many areas. Exposure to these levels does not call for action. Even at these low levels, there is a small risk associated with lifetime exposure to radon. However, authorities agree efforts to reduce radon levels still further are likely to be expensive and ineffective.
0.1	0.5 in 1000		

Figure 5.7. Example of a risk ladder with action standard and comparative risks, as used by Sandman and Weinstein (from Lipkus & Hollands, 1999)

Hoffman, Wilkes, Day, Bell & Higa (2006) also suggest the use of dart board and roulette wheel style diagrams for the communication of probability information (see examples in Figures 5.8 and 5.8, or the online demonstration at http://edoctoring.ncl.ac.uk/System_Check/psa_detect_html). Hoffman et al., argue that relative risk information, such as the probabilities of different side effects of medical treatments are not optimally communicated with pie-charts because it is difficult for patients to associate themselves with one of two populations presented in the chart. Presenting the information on a dart board capitalises on the idea that a throw of a dart is like the patients decision. Coloured regions around the circle are in proportion to the likelihood of different outcomes, so the decision is akin to deciding on which board one would rather “play”. One must be careful about likening health decisions to a game of chance, and the biases that may be associated with such pursuits (e.g., gambler’s fallacy, superstitions, framing effects with gains or losses, etc), but they may provide a useful tool for communicating probabilities and assisting people to clarify their values regarding medical treatment. Currently, no systematic evaluation of these tools compared to others appears to be available.

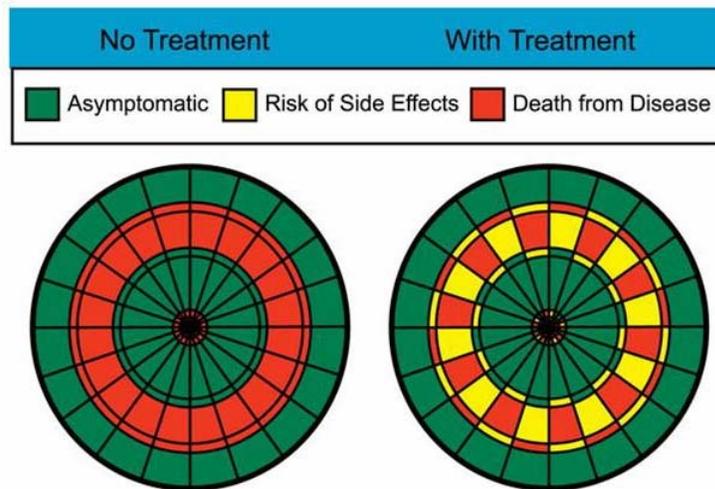


Figure 5.8. Example of Hoffman et als., dart board arrangement for the presentation of relative risks (from Hoffman et al., 2006).

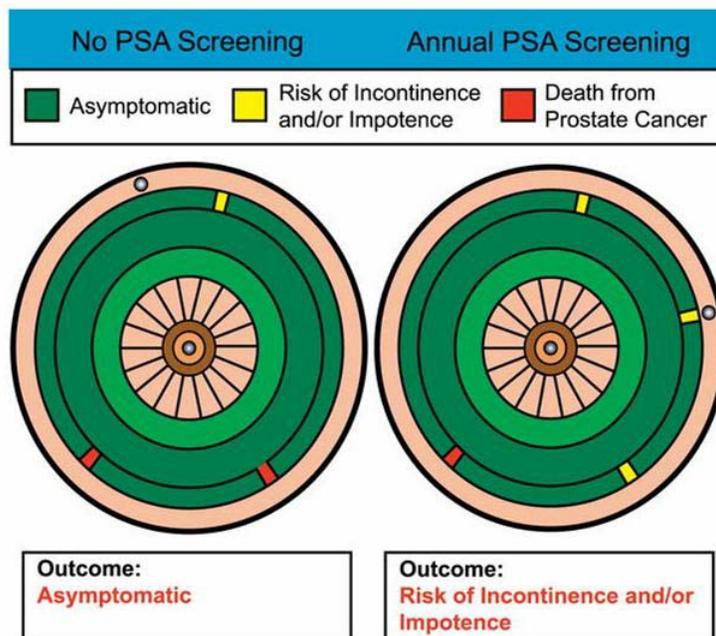


Figure 5.9. Example of Hoffman et al., roulette wheel arrangement for the presentation of relative risk in prostate cancer screening (from Hoffman et al., 2006).

5.3 Standards and recommendations for probability communication

While visual aids are thought to be useful in communicating probabilities despite limited specific evidence for their efficacy (Lipkus & Hollands, 1999), several other non-graphical strategies for communicating probabilities are supported in medicine. The following section briefly outlines recommendations for non-graphical presentations.

- (a) *Natural frequencies.* The use of natural frequencies or event rates has been previously covered in this review. Briefly, evidence suggests that natural frequencies are more easily understood than simply presenting probabilities as percentages, because they give clearly defined reference classes (Barratt et al., 2005), and prompt people to think about sets of cases that can be visualised and counted (Peterman, 2004). Mathematically, it is thought that frequency information maps onto much more basic skill sets, possessed by even young children, whereas single event probabilities require more sophisticated skills that are taught later in mathematical education (Anderson, 1998). Anderson also cites evidence that skills for frequency information have a longer evolutionary history than do skills for single-event probabilities. Girgerenzer and Edwards (2003) recommend the use of natural frequencies for both single event probabilities, and conditional probabilities (e.g., the probability of a positive test if the patient actually has the disease; the probability of a negative test if the patient does not

have the disease; the probability of the disease if the patient has a positive test). Examples of natural frequency formats, and how they differ to numerical percentage formats are shown in Figure 5.10. However, whether natural frequencies possess the same benefits for the communication of low probabilities does not seem to have been extensively considered in the literature. Given the idea that natural frequencies work because they allow people to visualise a reference class, whether the difficulty of imagining a high denominator (e.g., 100,000) makes this approach less useful remains an empirical question. Barratt et al., suggest that analogies such as “one person in a football stadium crowd” may be useful for presenting low probabilities, but this compromises accuracy and is likely to be interpreted differently by different people (depending on their experience of football stadiums, and stadium attendances!). How frequency presentations of small risks affect people’s understanding of them is thus an issue for future research.

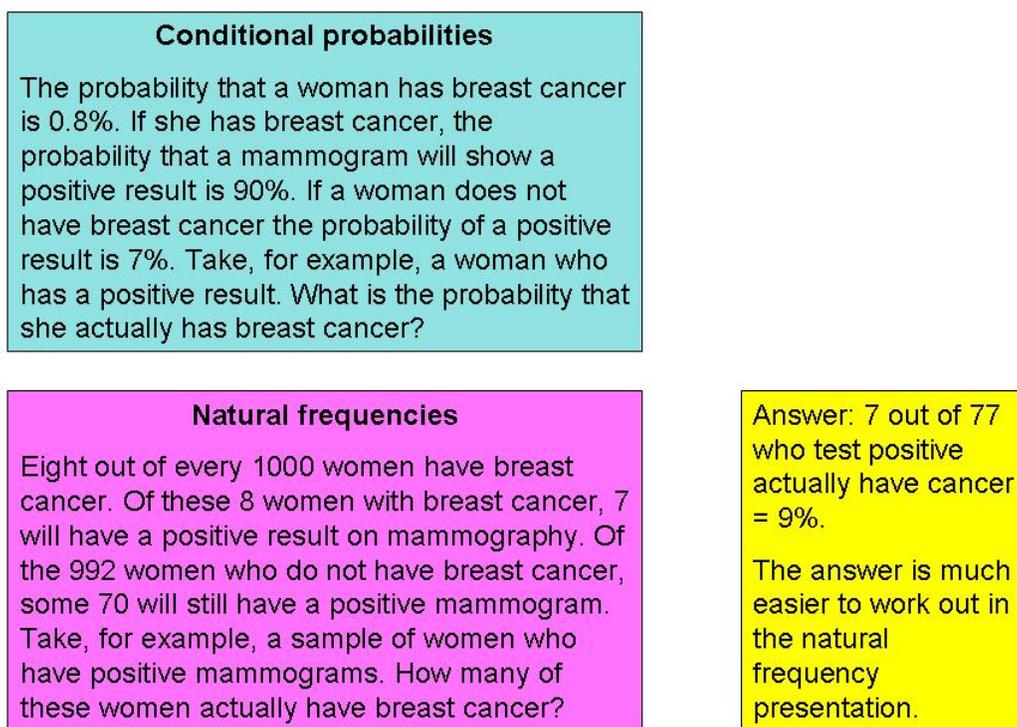


Figure 5.10. Comparison of probability information presented as conditional probabilities and natural frequencies (adapted from Girgerenzer & Edwards, 2003).

- (a) The use of frequency information has also been suggested as a way of improving how people interpret Bayesian analyses. People in several fields are not experienced with Bayesian statistics, in many cases relying solely on classical statistics. Anderson (1998) reviews the case for incorporating Bayesian analysis into ecology, but notes that ecologists may not be familiar with Bayesian

- formulae. Following the recommendations of Girgerenzer and Hoffrage (1995), Anderson compares Bayes' formula for a particular ecological problem with a graphical representation of the same information in a natural frequency format (see example of both formats in Appendix B). The graphical natural frequency format is much more accessible for anyone not experienced with Bayesian statistics, and may solve some of the problems of graphically representing large denominators mentioned above.
- (b) *Use of constant denominators.* Evidence suggests that constant denominators in natural frequency presentations are more readily understood than constant numerators. For example, saying "1 in 100 compared to 5 in 100" is more easily understood than saying "1 in 100 compared to 1 in 20" (Barratt et al., 2005). The finding that people tend to neglect base rates is important when presenting the probabilities of two events together in natural frequency format. Yamagishi (1997) found that people thought that a mortality rate of 12.86% was more dangerous than a rate of 24.14% when the rates were presented with different denominators. For example, participants rated cancer as more dangerous when it was expressed as killing 1,286 out of 10,000 people (which equates to 12.86%), compared to when it was presented as killing 24.14 out of 100 people (24.14%). This highlights the importance of using the same denominators if probabilities are to be compared.
- (c) *Timeframes.* Risks are often presented across the span of a lifetime. The estimation of lifetime is variable, and may be subject to bias. For example, to frame a risk in a positive way, presenting it as a figure across a lifetime may make it appear smaller. 10 year time frames are recommended (Woloshin, Schwartz & Welch, 2002). This may be important in biosecurity because presenting low probabilities may be facilitated by representing them across a 10 year risk, rather than just a low probability at the time of import/export (or over a single year).
- (d) *Communicating uncertainty.* In the medical field, communicating that there is a degree of uncertainty in probability estimates is seen as important. Barratt et al., suggest using phrases such as "our best guess is..." or using 95% confidence intervals. Even with issues where a lot of research has been done, and the likely outcomes are quite well known, the fact that it cannot be known which group a patient falls into (e.g., the 95% who survive, or the 5% who die) is seen as an important piece of information to convey.
- (e) *Communicating false positives and false negatives.* If a screening process can result in false positive or false negative results, the likelihood of this should be communicated (Barratt et al., 2005). This is particularly important in medicine, where screening for several types of cancers is not 100% accurate, and not providing this information can lead to false hope or false grief/worry. This information should also be presented in frequency format, and within the context

of subsequent events and their likelihood (e.g., presenting event rates for subsequent screening and treatment). These considerations may be relevant to biosecurity applications if diagnostic or statistical procedures where an error rate is known. For example, biosecurity procedures may fail to detect the presence of diseases or pathogens. It may be important to communicate the failure rate, and to use an appropriate format when doing so (a graphical format or natural frequency).

- (f) *Presenting evidence for probability values.* Being transparent about where probability estimates came from, the source of the data from which they were derived, and how they were calculated is recommended when communicating probabilities (Barratt et al., 2005). This removes the skepticism that some may have regarding the data, and may be particularly important when using event rates. Some research has suggested that when risk are presented in event rates (such as 1 in 10) people may question the validity of the data, just because it sounds so simple that it could be contrived in some way (Schapira, Nattinger, & McHorney, 2001). This may have to do with the rounded nature of multiples of 10. Rounding should be acknowledged (e.g., by saying “*about* 1 in 10”), and the source of the data provided, if for nothing else than to assure those reading the data of its veracity and due process.

5.4 Other strategies to facilitate the communication of probabilities

The following section reviews suggestions from diverse fields regarding other ways to improve the ways in which probabilities are communicated.

5.4.1. Translating verbal probability terms into the same numbers across individuals

As previously reviewed, people vary widely in the numbers they assign to verbal probability terms. This is a particular problem given that verbal terms are often used to exchange likelihood information. It is not just an issue when communicating to lay people, but experts within the same fields also show large variability with how verbal terms are used (see Karelitz and Budescu, 2004). Solving this problem by eradicating the use of verbal probability terms, and just using numbers does not account for the difficulties people have with numbers, nor does it cater for people’s preference for verbal terms to express their opinions when there is some degree of uncertainty (see Smit and Hoorens, 2005). While a predefined “verbal vocabulary” (like an operational dictionary linking verbal terms to particular numbers as they are intended to be used in that context) seems the next most appropriate option, this does not account for how people find it

difficult to suppress the meanings they usually associate with particular terms (Karelitz & Budescu). Studies of weather forecasters provide a good example of this idea. Weather forecasters are taught to use terms in particular standardised ways when generating forecasts, but do not interpret those same terms in the same way when they are part of a medical diagnosis, suggesting that other personal preferences or strategies over-ride the standard they have used in other contexts (Wallsten, Fillenbaum & Cox, 1986).

Accordingly, Karelitz and Budescu (2004) describe a method for “translating” different people’s verbal probability terms into the same numbers. To do this, they developed “membership functions” of a particular person’s phrases to show which numerical probabilities belong to the phrase and which do not (see Figure 5.11). From here, it is possible to map different probability lexicons onto the same scale. Participants selected and ranked several probability terms, and then rated the extent to which numerical probabilities applied to each them. Participants also rated the likelihood, both numerically and verbally, of a dart hitting the target of several dart-boards, showing different shaded proportions. From this information, Karelitz and Budescu were able to create membership functions, and tested 3 methods of translating the verbal into the numerical for pairs of participants (alternating in the roles of communicator and receiver). The details of the translation procedures are beyond the scope of this review, but all methods improved the communication of verbal probability terms.

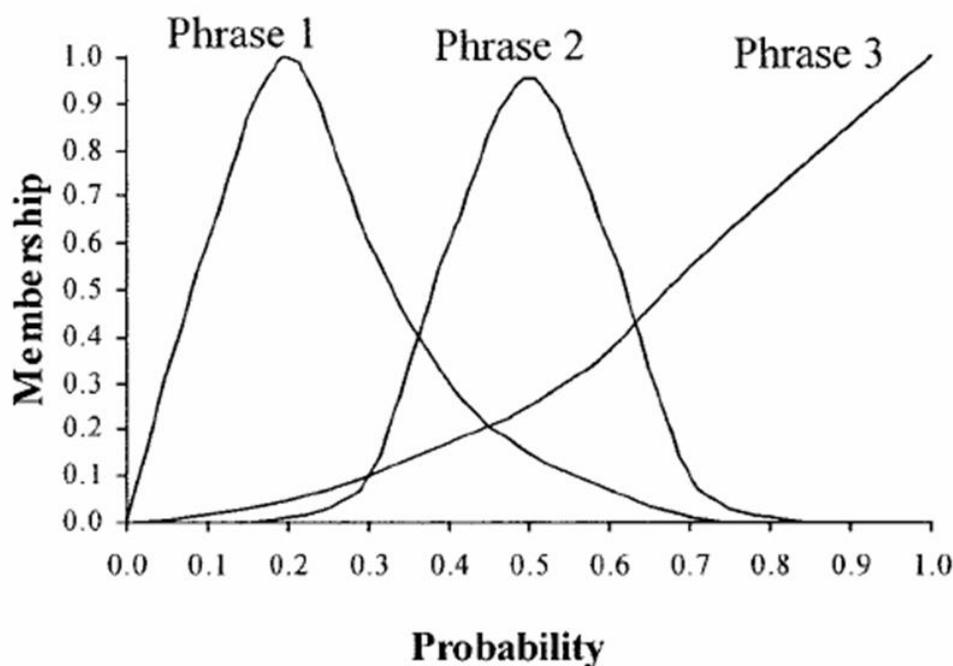


Figure 5.11. Demonstration of numerical probability membership functions of different probability phrases (from Karelitz & Budescu, 2004).

5.4.2 Geographical models of risk magnitude

Communicating about the risks of nuclear waste, including how to transport it and where to place it, is a significant challenge for governments worldwide. Community views on nuclear waste are not always commensurate with reality, and considerable community consultation is required on these issues. Drew et al., (2003) report on a project to change the method of community consultation where geographical models of the placement and transportation of nuclear waste were employed at stakeholder information meetings. Different coloured Lego blocks were used to represent different types and volumes of waste, and were placed on a detailed map of the USA, showing state borders and key road and rail transportation routes (see Figure 5.12). With the interactive 3-dimensional maps, stakeholders were able to simulate the movement of the waste around the country and discuss and model different storage and transportation plans. Overall, the Lego map was rated as the most helpful device used in the consultative process. It helped participants to conceptualise the types of waste and their locations, as well as the volume of waste in each location, and discuss the concerns of particular community groups around the storage and transportation routes (Drew et al.,). While this is not strictly a probability communication device, proportions were conveyed effectively with this strategy. In cases where geographic location is important, this technique could be useful. The probability of an event in a particular location, given the presence of particular stimuli at that location could be represented in much the same way as volume was represented in Drew et al., description (e.g., number of dots on the blocks). Similarly, colour or size of the blocks could represent relative areas, or severity of consequences. A simplistic example of how this technique could be adapted to biosecurity applications would be mapping the probability of infection X in crop Y at locations A, B and C in New South Wales, given the parameters of the particular location, and the amount of livestock carrying infection X being imported. This could be particularly useful in community and stakeholder consultation regarding an import decision.



Figure 5.12. Photograph of the 3-dimensional Lego model used to communicate nuclear waste volume and location (from Drew et al., 2003).

6. Applying aspects of the literature to Biosecurity

The following section uses IRA documentation to identify situations where the recommendations from the literature reviewed in this project may be useful in improving the communication of probabilities. Instances where there was evidence of misunderstanding of probabilities, or potential for future misunderstandings were identified for this purpose. This analysis is illustrative rather than comprehensive, due to the sheer number and volume of IRAs conducted by Biosecurity Australia.

The basic procedure employed was to become familiar with the IRA procedure, read through the appeals and responses to the draft documents, as well as the final IRAs and supporting materials. The importation of New Zealand Apples was chosen. All references to IRA documentation below refer to documents available under the New Zealand Apples IRA link on the Biosecurity Australia website, unless otherwise stated.

6.1 Stakeholder comments on the draft IRA

The final inventory of stakeholder comments (Final inventory of issues raised by stakeholders in response to the draft IRA on NZ apples, no date; hereafter referred to as FI) was used to identify situations where stakeholders appear to have difficulties with probability information. These comments are reviewed here with suggestions for how the issue may be resolved.

(a) Some stakeholders have concerns with the verbal probability terms used.

E.g., Description of likelihoods: “Where do the classifications and associated descriptions come from? Are they internationally recognized?” (FI pp.25). (see also other examples in the section on fire blight; probability of entry, pp.34-35; and Arthropods: Mealybugs pp.45).

The final IRA Part B (November, 2006) also recognises the confusion that qualitative terms (such as “high”, “low”, “medium”) caused for stakeholders in the 2004 draft (see pp. 42). It is recognised in the IRA that stakeholders will apply their own meanings to these terms, rather than the intended meanings set out in the guidelines on IRAs. These observations mirror the literature on the difficulties of translating verbal probability expressions between people (see Karelitz and Budescu, 2004; Wallsten, Fillenbaum & Cox, 1986; Sutherland et al., 1991), and how experience and motivations can influence the interpretation of such terms (Smits & Hoorens, 2005).

This does not mean that verbal terms should not be used – they are an important part of qualitative analyses, but strategies for enhanced understanding could be trialed. Providing a “key” to the meaning of terms in a separate location is not enough – people will still be

confused and are still likely to use their own interpretations of the terms (Wallsten, Fillenbaum & Cox, 1986). Providing the key in several locations where the terms are used may help. Alternatively, a simple graphical representation accompanying the key, which shows the relative difference between the terms may also be useful (much like a “progress bar” used in many computer software applications, or even a graphical representation of the numerical probability associated with each particular term). Table 12 in the Final IRA (November 2006, pp.43) includes numerical estimates of semi-qualitative likelihood terms, so this may be an appropriate location to place graphical representations. It remains likely that people will still use their own interpretation when they encounter such terms, because the subjective meaning of the term is engrained, and some people are not likely to look at reference information repeatedly. However, stakeholder workshops and the development of training resources for the interpretation of probabilities, as discussed below, may help to counter this effect (see section 6.3, General recommendations).

(b) Combining qualitative terms

Some of the stakeholder comments also express concern over combining likelihoods of distinct events.

E.g., In splitting the probability of entry into effectively 8 components, BA stated that the IPPC definition of entry is preserved. However, the model is not robust enough to overcome the mathematical outcome of this splitting. This is because the overall probability of an event is effectively determined by the lowest probability of any sub-component. (FI pp.26)

and

“The probability of importation and distribution should not be combined together...probability of fire blight infected fruit in fruit ready for shipment in NZ – high; probability of failure to detect fire blight at the border – high; therefore distribution of fire blight around Australia is high” (FI, pp.27).

These comments also relate to the table of “rules” for combining descriptive likelihoods in the final IRA (pp.43). This table is quite difficult to understand. It is not immediately clear, for example, that when a high likelihood and a low likelihood event are combined, that the resulting likelihood should be low.

Perhaps some of this confusion can be solved by further justification for this procedure. The final IRA does refer to the guidelines on IRAs for information about how these classifications were developed, however, communicating these more immediately may be more effective (see also general comments below regarding stakeholder workshops). Additionally, presenting the information in natural frequencies, or semi-graphical natural frequencies, such as suggested by Anderson (1998) regarding Bayesian analyses (see Appendix B) may assist in communicating the combination of a sequence of events.

(c) One of the stakeholder comments shows a general lack of understanding of probability concepts.

“No reference or reasons are given for a high or extreme probability being close to one or a negligible probability being very close to zero” (FI, pp.29)

This kind of comment reflects the need to be aware of the diverse backgrounds of stakeholders. Presenting material in several formats may be the best way to communicate with such a range of stakeholders, and this approach is supported in the education literature regarding the different ways in which people prefer to receive information (Gardener, 2006). It may also be useful to develop background materials that can be used with all Biosecurity documents that explain probability concepts, including:

- numerical probabilities;
- verbal probability terms and their intended meanings; definitions of reference classes in natural frequency arrangements, and;
- the introduction of graphical formats with relevant examples.

(e) Other stakeholder comments

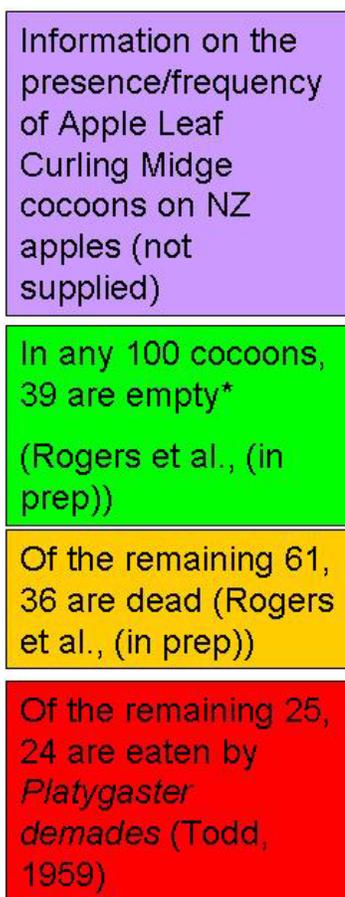
Parts of the response to the revised draft IRA by the Australian Access Action Group (AAAG; 29/3/2006) are also consistent with issues raised in the literature and/or could improve the communication of probabilities by changes in format. The discussion on page 6 regarding appropriate levels of protection (ALOP) demonstrates the tendency for people to deal with low probabilities as if they were zero, as predicted by Kahneman and Tversky's (1979) prospect theory. The probability range of 0.001 to 0.000001, is dismissed as essentially zero, and seen as a trade barrier. Several future research issues could be pertinent to this. It would be interesting to determine whether and how people's perception of low probabilities change if these probabilities are presented on the Paling perspective scales, or a risk ladder with information about other more familiar risks. This research, on the effect of the contextualisation of low probability could impact on arguments of the kind made by the AAAG (see also Section 7.1 B).

The following paragraph from the AAAG document on Apple Leaf Curling Midge seems akin to the information presented in the review on conditional probabilities, and could be conveyed in a more convincing and readily comprehensible manner:

Central to our grievance is the IRA's apparent assumption that any cocoon present on apple fruit presents is a viable risk. This is a spurious assumption in the face of an enormous amount of data disputing the fact.

*Recent research by New Zealand scientists working on this issue shows that 36-42% of the cocoons present on New Zealand apple fruit are empty (Rogers et al (in prep)). Additionally any cocoons with pupae present show a 59% mortality rate (Rogers et al (in prep) and importantly Todd (1959) showed that up to 95% of the remaining pupae were parasitized by *Platygaster demades* (AAAG, 2006, pp.15).*

This could have been expressed in natural frequencies, or in a diagram of natural frequencies similar to Anderson's (1998) example (see Figure 6.1).



*average of 36 and 42 was taken for simplicity

Figure 6.1 Representation of data from the AAAG's submission on the draft IRA New Zealand Apples in a frequency format.

These kinds of presentations should be kept in mind not just for IRAs but in all forms of communication about biosecurity risks.

6.2 Other parts of the IRA documentation

(a) Similar to the AAAG document referred to above, some parts of the pest risk assessments presented several numerical probabilities together in one paragraph (e.g., Fire blight pp 56). This has to be taken in context, because this was part of a literature review, and without the probability information, the conclusions could not be made. It is also impractical to present all of the numerous probability values in graphical form. Perhaps at least some of them could be presented in natural frequencies or with a clear

definition of reference classes, especially when they are intended to be compared between studies or locations.

(b) The use of scientific notation in the pest risk assessments also requires consideration. Scientific notation is a convenient way of representing small probabilities, but may not be optimal for conveying probabilities to all stakeholders, particularly in the case of relative probabilities. While the principle that the probability becomes lower as the negative superscript becomes higher is clear, this is not easy to visualise. The difference in the index may only be 1 (i.e., 10^{-4} compared to 10^{-3}), but this equates to a very large difference in probability. Again it may be impractical to represent these graphically every time, considering the frequency with which probabilities with scientific notation are used. However, it may be possible to represent some of the probabilities in natural frequencies (with the same denominator if they are to be compared), or to provide a tool to guide the interpretation of scientific notation (e.g., $1 \times 10^{-3} = 0.001 = 1$ in 1000; $10^{-4} = 0.0001 = 1$ in 10,000; $10^{-5} = 0.00001 = 1$ in 100,000...).

6.3 General findings and recommendations

Despite the abovementioned practical difficulties with using the graphical representations of probabilities in each case that a probability is conveyed, opportunities where graphical representations would be useful can still be identified. Similarly, particular stakeholders for whom different formats may be beneficial could be identified, along with appropriate communication methods.

For example

- Using a natural frequency or graphical format on only key probability findings may cue stakeholders to the reference class and aid understanding when other probabilities are presented. Examples of this might include the $P_{\text{importation (apple)}}$ (the probability that an individual apple will be infected or infested),
or the PEES (probability of entry, exposure and spread), and several other important probability estimates.
- The methodology for the Final IRA on NZ Apples (November, 2006) gives information on how probabilities to be reported in the document are calculated. While the information in the methodology section appears to be designed to assist with the interpretation of information later presented in the document, it does not take advantage of graphical presentation or the use of natural frequencies (see for example Tables 8, 9, 12, and 13). The recommendations above on the use of natural frequencies, relative risk ladders or perspective scales, and methods to graphically represent probability terms could be applied to this information and included at least at the beginning of an IRA document, perhaps giving graphical examples of how the probability estimate in question should be interpreted (e.g., in terms of reference class etc).

It may be appropriate to use graphical displays at stakeholder meetings or workshops, rather than in the IRA documentation. Providing handouts that stakeholders can take away from such meetings, similar to the format of a medical decision aid, where the risks (presented graphically or otherwise) are contextualised within the benefits and consequences, may enhance people's understanding of the information, and make it more accessible. In addition, the stakeholder workshop approach is likely to increase perceptions of transparency, consultation, and adherence to scientific protocol, all of which were issues that received negative comments from stakeholders. It may even be possible to hold workshops not specific to any one IRA, but to educate stakeholders on the general principles with which Biosecurity Australia makes qualitative and quantitative classifications. Handouts, examples from a range of IRAs, and graphical presentations of probabilities are likely to be of use in such situations, and foster better understanding of future IRA proceedings.

7. Research recommendations

Several issues covered in this review require further research. Some of these come from the state of research in the general domain on visual tools in probability communication, but the same questions can be appropriately answered in a biosecurity context. The advantage of this approach is that the wider literature on visual tools would be advanced, and the imperative to assess how it may serve an applied situation would also be satisfied. The following is an outline of suggested research objectives and accompanying preliminary methodological considerations for conducting this research.

7.1 Major research considerations

- A) Which visual displays result in optimal understanding of probability information (specific to understanding probabilities, rather than just measuring decision making)?
 - i) Research should focus on the Paling perspective scale or Hoffman et als., (2006) dart board/roulette arrangement, which have had little or no research attention to date.
 - ii) An important methodological consideration is an improved means of testing the understanding of probabilities per se, which does not focus on decisional outcomes, nor make participants feel like they are completing a math test⁶. This is important because to date, research appears to have used behavioural outcomes as an index of probability understanding. With a change in focus to what people understand, and thus what comments they are likely to make, a different approach is required.
- B) How can visual displays be best used for the optimal understanding of low probabilities?
 - i) What consequences might there be for representing only the numerator graphically?
 - ii) How effective are natural frequency presentations for low probabilities, compared to other representations of probabilities?
 - iii) What are the effects of contextualising risks (e.g., risk ladders; action standards and information) on the perception of low probabilities?
 - iv) How can the increase in risk aversion associated with the presentation of relative risks (and relative risk with graphics) be used effectively? Will this effect understanding of risks as well as decisional outcomes/behavioural intentions?
 - v) Particular attention needs to be given to different ways of expressing probabilities in scientific notation (e.g., assessing whether people have difficulties interpreting this information and if so how to rectify this).

⁶ The phenomenon of test-anxiety is a significant concern in this type of research, particularly if performed with stakeholders/participants who have little scientific or mathematical backgrounds, or differing education levels (see Spielberger 1980 for information on the test anxiety concept, scale and literature).

- C) What effects (if any) are there in making people aware of their motivational bias in converting verbal probability terms to numerical probabilities?
- D) Research needs to be conducted across a range of stakeholder groups, including people internal to Biosecurity Australia, including
 - i) Baseline testing of understanding of probability information
 - ii) Investigating if changes in presentation of probabilities result in changes in understanding among experts, building on research discussed in Section 4.5.
 - iii) Identifying stakeholders who would most benefit from changes in probability presentations, and communication modes.
- E) Assessment of specific difficulties encountered with biosecurity probabilities is fundamental to most of this research. Some evidence in the IRA stakeholder comments highlights misunderstandings and concerns, though it is important to evaluate what parts of the IRAs (or similar documents) they do not understand and explore their perceptions of ways to rectify this.
- F) How can membership functions be operationalised and used for the translation of verbal probability terms between people involved in an import risk analysis (for example) and the stakeholders?

7.3 Suggested methodologies for answering these questions

Some of the research questions raised above lend themselves to experimental research, while others would be appropriately answered by non-experimental research, such as focus groups and interviews. Findings from both types of inquiry are likely to influence one another. The following is just an outline of how the experimental and non-experimental approaches could be used. Detailed methodologies would need to be designed for each research question with careful consideration of control variables, reliability and validity of measures, and appropriate statistical analyses.

Experimental work would be useful for testing basic premises, such as what generally happens with low probabilities when they are contextualised (B(iii)); issues with scientific notation (B(v)); and comparing the effectiveness of different visual tools (A). A potential solution to the problem outlined in A (ii), the measurement of understanding of probabilities, would be ranking different probability expressions. Experimental studies could use materials from a biosecurity context to develop various visual aids (numerous examples are available from the various IRAs), and assess how understanding changes with different presentations. Such experiments could be conducted with a group of lay participants, or sample stakeholders. Any findings from the basic experimental work could then be applied (experimentally or otherwise) into “real” applications to assess how they transfer out of the laboratory.

Initial focus group or survey research is important in ascertaining the degree of problems that stakeholders have with probability information (E), and identifying groups with particular needs (D(iii)). Specific recommendations made in Section 6, regarding including graphics in tables; using natural frequencies more often; and developing and

evaluating stakeholder workshops and training resources could be piloted with a group of stakeholders in a quasi-experimental design.

Researching how the visual and other formats discussed in this project can enhance probability communication is an important pursuit, and one that can be readily achieved.

8. References

- AAAG. (2005). *Submission in response to "Revised draft IRA for apples from New Zealand"*. Released by Biosecurity Australia.
- Anderson, J. L. (1998). Embracing uncertainty: the interface of Bayesian statistics and cognitive psychology. *Conservation Ecology*, 2(1).
- Barratt, A., Edwards, A., Trevena, L., McCaffery, K., Woloshin, S., Bekker, H., Timmermans, D., Charvet, A. (2005). Presenting probabilities. In A. M. O'Connor, Llewellyn-Thomas, H., Stacey, D. (Ed.), *International Patient Decision Aids Standards (IPDAS) Collaboration Background Document*.
- Berry, D., Raynor, T., Knapp, P., Bersellini, E. (2004). Over the counter medicines and the need for immediate action: a further evaluation of European Commission recommended wordings for communicating risk. *Patient Education and Counseling*, 53, 129-134.
- Biehl, M., Halpern-Felsher, B. (2001). Adolescents' and adults' understanding of probability expressions. *Journal of Adolescent Health*, 28, 30-35.
- Biosecurity Australia (2006). *Final IRA for apples from New Zealand, Part B*. Canberra: Biosecurity Australia
- Biosecurity Australia (no dates). *Final inventory of issues raised by stakeholders in response to the draft IRA on New Zealand apples*. Canberra: Biosecurity Australia.
- Budescu, D. V., Wallsten, T.S. (1985). Consistency in interpretation of probabilistic statements. *Organizational Behavior and Human Decision Processes*, 36, 391-405.
- Budescu, D. V., Wallsten, T.S. (1990). Dyadic decisions with numerical and verbal probabilities. *Organizational Behavior and Human Decision Processes*, 46, 240-263.
- Budescu, D. V., Weinberg, S., & Wallsten, T.S. (1988). Decisions based on numerically and verbally expressed uncertainties. *Journal of Experimental Psychology: Human Perception and Performance*, 14, 281-294.
- Cantillon, P. (2004). Is evidence based patient choice feasible? *British Medical Journal*, 329, 39.
- Charles, C., Gafni, A., Whelan, T. & O'Brien M. (2005). Treatment decision aids: Conceptual issues and future directions. *Health Expectations*, 8, 114-125.
- Drew, C. H., Grace, D.A., Silbernagel, S.M., Hemmings, E.S., Smith, A., Griffith, W.C., Takaro, T.K., Faustman, E.M. (2003). Nuclear waste transportation: Case studies of identifying stakeholder risk information needs. *Environmental Health Perspectives* 111(3), 263-272.
- Eddy, D. M. (1982). Probabilistic reasoning in clinical medicine: Problems and opportunities. In D. Kahneman, P. Slovic & A. Tversky (Eds.), *Judgement under uncertainty: Heuristics and biases* (pp. 249-). Cambridge: Cambridge University press.
- Feldman-Stewart, D., Brundage, M.D. (2004). Challenges for designing and implementing decision aids. *Patient Education and Counseling*, 54, 265-273.

- Feldman-Stewart, D., Kocovski, N., McConnell, B.A., Brundage, M.D., Mackillop, W.J. (2000). Perception of quantitative information for treatment decisions. *Medical Decision Making*, 20, 228-238.
- Gardner, H. (2006). *Multiple Intelligences: New Horizons*: Basic Books.
- Ghosh, A. K., and Ghosh, K. (2005). Translating evidence-based information into effective risk communication: Current challenges and opportunities. *Journal of Laboratory and Clinical Medicine*, 145(4), 171-180.
- Gigerenzer, G. (2002). *Calculated risks: How to know when numbers deceive you*. New York: Simon and Schuster.
- Gigerenzer, G., Edwards, A. (2003). Simple tools for understanding risks: from innumeracy to insight [Electronic Version]. *British Medical Journal*, 327, 741-744. Retrieved 16 January 2007 from <http://www.bmj.com/cgi/content/full/327/7417/741>.
- Gigerenzer, G., Hertwig, R., van den Broek, E., Fasolo, B. & Katsikopoulos, K.V.. (2005). "A 30% chance of rain tomorrow": How does the public understand probabilistic weather forecasts? *Risk Analysis*, 25(3), 623-629.
- Gigerenzer, G., Hoffrage, U. (1995). How to improve Bayesian reasoning without instruction: frequency formats. *Psychological Review*, 102(4), 684-704.
- Gurmankin, A. D., Baron, J. & Armstrong, K. (2004). Intended message versus message received in hypothetical physician risk communications: Exploring the gap. *Risk Analysis*, 24(5), 1337-1347.
- Hatfield, J., and Job, R.F. S. (2001). Optimism Bias about Environmental Degradation: The Role of the Range of Impact of Precautions. *Journal of Experimental Psychology*, 21, 17-30.
- Hoffman, J. R., Wilkes, M.S., Day, F.C., Bell, D.S., Higa, J.K. (2006). The roulette wheel: An aid to informed decision making [Electronic Version]. *PLoS Medicine*, 3, 743-748.
- Hoffrage, U., Lindsey, S., Hertwig, R., Gigerenzer, G. (2000). Communicating Statistical Information [Electronic Version]. *Science*, 290, 2261-2262.
- Job, R. F. S., Hamer, V., & Walker, M. (1995). The effects of optimism bias and fear on protective behaviours. In D. T. Kenny, Job, R.F. S (Ed.), *Australia's Adolescents: a health psychology perspective* (pp. 151-156). Armidale, Australia: University of New England Press.
- Kahneman, D., Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2), 263-292.
- Kahneman, D., & Tversky, A. (1984). Choices, values and frames. *American Psychologist*, 39, 341-350.
- Kaplan, R. M., Frosch, D. L., (2005). Decision making in medicine and health care. *Annual Review of Clinical Psychology*(2005.1), 525-556.
- Karelitz, T. M., Budescu, D.V. (2004). You say "probable" and I say "likely": Improving interpersonal communication with verbal probability phrases. *Journal of Experimental Psychology: Applied*, 10(1), 25-41.
- Konold, C. (1995). Issues in assessing conceptual understanding in probability and statistics. *Journal of Statistics Education*, 3(1).
- Lee, D. H., Mehta, M.D. (2003). Evaluation of a visual risk communication tool: effects on knowledge and perception of blood transfusion risk. *Transfusion*, 43, 779-787.

- Lee, S. H. V., & Job, R. F. S. (1995). The effect of information on optimism bias. In D. T. Kenny & R. F. S. Job (Eds.), *Australia's Adolescents: A health psychology perspective* (pp. 157-163). Armidale University of New England Press.
- Lipkus, I. M., Hollands, J.G. (1999). The visual communication of risk. *Journal of the National Cancer Institute Monographs*, 25, 149-263.
- Man-Son-Hing, M., Laupacis, A., O'Connor, A., Biggs, J., Drake, E., Yetisir, Elizabeth., Hart, R. G. (1999). A patient decision aid regarding antithrombotic therapy for stroke prevention in atrial fibrillation: A randomised control trial. *The Journal of the American Medical Association*, 282(8), 737-743.
- McBride, C. M., Bastian, S., Halabi, S., Fish, L., Lipkus, I.M., Bosworth, H.B., Rimer, B.K., Siegler, I.C. (2002). A tailored intervention to aid decision making about hormone replacement therapy. *American Journal of Public Health*, 92(7), 1112-1114.
- O'Connor, A. M., Légaré, F., & Stacey, D. (2003). Risk communication in practice: the contribution of decision aids. *British Medical Journal* 327, 736-740.
- O'Connor, A. M., & Stacey, D., Entwistle, V., Llewellyn-Thomas, H., Rovner, D., Holmes-Rovner, M., Tait, V., Tetroe, J., Fiset, V., Barry, M., Jones, J. (2003). Decision aids for people facing health treatment or screening decisions. *The Cochrane Database of Systematic Reviews*, 1.
- Osman, M., Shanks, D.R. (2005). Individual differences in causal learning and decision making. *Acta Psychologica*, 120, 93-112.
- Paling, J. (2003). Strategies to help patients understand risks. *British Medical Journal*, 327, 745-748.
- Peterman, R. M. (2004). Possible solutions to some challenges facing fisheries scientists and managers. *ICES Journal of Marine Science*, 61, 1331-1343.
- Sandman, P. M., Weinstein, N.D. (1994). *Communicating effectively about risk magnitudes: Bottom line conclusions and recommendations for practitioners* Washington (DC): Environmental Protection Agency.
- Schapira, M. M., Nattinger, A.B., & McHorney, C.A. (2001). Frequency or probability? A qualitative study of risk communication formats used in health care. *Medical Decision Making*, 21, 459-467.
- Schapira, M. M., VanRuiswyk, J. (2000). The effect of an illustrated pamphlet decision-aid on the use of prostate cancer screening tests. *The Journal of Family Practice*, 49(5), 418-424.
- Smits, T., Hoorens, V. (2005). How probable is probably? It depends on whom you're talking about. *Journal of Behavioral Decision Making*, 18(2), 83-96.
- Spielberger, C. D. (1980). *Test Anxiety Inventory*. Redwood City: Mindgarden.
- Spittal, M. J., McClure, J., Siegert, R.J., Walkey, F.H. (2005). Optimistic bias in relation to preparedness for earthquakes. *The Australasian Journal of Disaster and Trauma Studies*, 1.
- Steckelberg, A., Balgenorth, A., Berger, J., Mühlhauser, I. (2004). Explaining computation of predictive values: 2 x 2 table versus frequency tree. A randomized controlled trial [Electronic Version]. *BMC Medical Education*, 4.
- Stone, E. R., Rush, C.J. (1997). *Risk communication: the effectiveness of graphical modes depends on the risk magnitude*. Paper presented at the annual meeting for the Society for Judgement and Decision-making.

- Stone, E. R., Yates, J.F., Parker, A.M. (1997). Effects of numerical and graphical displays on professed risk-taking behaviour. *Journal of Experimental Psychology: Applied*, 3(4), 243-256.
- Stone, E. R., Yates, J.F., Parker, A.M. . (1994). Risk communication: Absolute versus relative expressions of low-probability risks. *Organizational Behavior and Human Decision Processes* 60, 387-408.
- Sutherland, H. J., Lockwood, G.A., Tritchler, D.L., Sem, F., Brooks, L., Till, J.E. (1991). Communicating probabilistic information to cancer patients: Is there 'noise' on the line? *Social Science & Medicine*, 32(6), 725-731.
- Teigen, K. H. (1994). Variants of subjective probabilities. In G. Wright, Ayton, P. (Ed.), *Subjective probability* (pp. 211-238). Chichester; New York: Wiley.
- Wallsten, T. S., Fillenbaum, S., & Cox, J. A. (1986). Base rate effects on the interpretations of probability and frequency expressions. *Journal of Memory and Language*, 25, 571-587.
- Weinstein, N. D. (1980). Unrealistic optimism about future life events. . *Journal of personality and social psychology*, 39(5), 806-820.
- Weinstein, N. D. (1982). Unrealistic Optimism About Susceptibility to Health Problems. *Journal of Behavioural Medicine*, 5(4), 441-460.
- Weinstein, N. D. (1983). Reducing Unrealistic Optimism About Illness Susceptibility. *Health Psychology*, 2(1), 11-20.
- Weinstein, N. D., Klein, W M. (1996). Unrealistic Optimism: Present and Future. *Journal of Social and Clinical Psychology*, 15(1), 1-8.
- Woloshin, S., Schwartz, L. M., & Welch, H. G. (2002). Risk charts: Putting cancer in context. *Journal of the National Cancer Institute*, 94(11), 799-803.
- Yamagishi, K. (1997). When a 12.86% mortality is more dangerous than 24.14%: Implications for risk communication. *Applied Cognitive Psychology*, 11, 495-506.

9. Appendices

Appendix A

Examples of medical decision aids

Many examples of medical decision aids are available online. Because they are in pdf format it is difficult to include them here.

Suggested URLs

Download “Hormone replacement therapy – Making decisions. Should I use hormone replacement therapy?” from

<http://www.nhmrc.gov.au/publications/synopses/wh35syn.htm>

and look at

- Should you take Traditional non-steroidal anti-inflammatory drugs (NSAIDs) or Cox2 inhibitors for osteoarthritis?
- Making Choices: Life Changes to Lower Your Risk of Heart Disease and Stroke
- [Ankylosing spondylitis decision aid](#): Should I take biologic agents such as infliximab or etanercept?

at <http://decisionaid.ohri.ca/decaids.html>

Appendix B

Examples of Bayesian analysis converted into natural frequency format (adapted from Anderson, 1998).

An ecologist needs an indicator for the presence of breeding Marbled Murrelets in the forest. A good indicator should be a simple observation that indicates with high probability something about the ecosystem that is important but not easy to observe. The existence of suitable nest sites (big horizontal branches with lots of moss and lichen) is chosen as such an indicator. A survey is developed to detect possible sites, and it is tried out on 1000 stands.

The data is in the Table

	Murrelets are nesting in the stand	Murrelets are not nesting in the stand	Marginal total
Survey data: Potential nest sites not seen	4	808	812
Survey data: Potential nest sites were seen	46	142	188
Marginal total	50	950	1000

The ecologist wants to know the probability that murrelets are nesting in stand X, given the observation of suitable nest sites in the survey.

Bayes' theorem:

$$\begin{aligned}
 P(H_1|D) &= \frac{P(H_1) * P(D|H_1)}{P(H_1) * P(D|H_1) + P(H_2) * P(D|H_2)} \\
 &= \frac{0.05 * 0.92}{(0.05 * 0.92) + (0.95 * 0.15)}
 \end{aligned}$$

$$= 0.244$$

So if potential nest sites are observed, the probability that Murrelets are actually nesting there is 0.24.

The above information can be represented graphically as follows:

