CURRENT OPINION

Drug Safety 2003; 26 (1): 1-11
0114-5916/03/0001-0001/$30.00/0

Patients' Understanding of Risk Associated with Medication Use

Impact of European Commission Guidelines and Other Risk Scales

Dianne C. Berry,[1] *D.K. Raynor,*[2] *Peter Knapp*[2] and *Elisabetta Bersellini*[1]

1 University of Reading, Reading, United Kingdom
2 University of Leeds, Leeds, United Kingdom

Abstract

Patients want and need comprehensive and accurate information about their medicines so that they can participate in decisions about their healthcare. In particular, they require information about the likely risks and benefits that are associated with the different treatment options. However, to provide this information in a form that people can readily understand and use is a considerable challenge to healthcare professionals. One recent attempt to standardise the language of risk has been to produce sets of verbal descriptors that correspond to specific probability ranges, such as those outlined in the European Commission (EC) Pharmaceutical Committee guidelines in 1998 for describing the incidence of adverse effects.

This paper provides an overview of a number of studies involving members of the general public, patients, and hospital doctors, that evaluated the utility of the EC guideline descriptors (very common, common, uncommon, rare, very rare). In all studies it was found that people significantly over-estimated the likelihood of adverse effects occurring, given specific verbal descriptors. This in turn resulted in significantly higher ratings of their perceived risks to health and significantly lower ratings of their likelihood of taking the medicine. Such problems of interpretation are not restricted to the EC guideline descriptors. Similar levels of misinterpretation have also been demonstrated with two other recently advocated risk scales (Calman's verbal descriptor scale and Barclay, Costigan and Davies' lottery scale).

In conclusion, the challenge for risk communicators and for future research will be to produce a language of risk that is sufficiently flexible to take into account different perspectives, as well as changing circumstances and contexts of illness and its treatments. In the meantime, we urge the EC and other legislative bodies to stop recommending the use of specific verbal labels or phrases until there is a stronger evidence base to support their use.

Over the past decade there has been an increasing realisation that patients want and need information about their medicines that is authoritative, accurate and understandable so that they can fully participate in decisions about their healthcare. Enabling people to make informed judgements is a cornerstone of the philosophy of concordance in medicine taking[1] and is a key aspect of the UK National Plan for the National Health Service (NHS).[2] However, this move towards patient empowerment also presents a considerable challenge as many patients are cognitively and emotionally ill-equipped to understand, retain, and use the necessary information.[3] Effectively communicating even the simplest and most unthreatening of messages to a diverse audience is difficult enough. The problems of communicating complex medical information, involving risk and uncertainty, are considerable.

Consumers need to know many types of information about the medicines that they take. This includes the medicine's name, purpose and how it works. People also need to know about its contraindications and its possible adverse effects, including how likely they are to happen and what to do if they do occur. One approach to simplifying and standardising the presentation of probabilistic information (such as when informing patients about the benefits and risks associated with particular medicines) has been to produce sets of verbal descriptors that correspond to specific probability ranges.[4] Undoubtedly the most influential of these in relation to medicine taking was outlined in the European Commission (EC) Pharmaceutical Committee guidelines in 1998.[5] This followed an earlier published EC directive[6] which stated that every medicine supplied in the UK should be accompanied by a comprehensive patient information leaflet, which must include a list of all adverse effects referred to in the medicine's summary of product characteristics (previously known as the data sheet), in a form that the patient would understand. The 1998 guideline recommended that the frequency of adverse effects could be banded into five groupings based on five verbal descriptors, ranging from 'very rare' to 'very common', with each term being associated with a specified range of frequencies of occurrence (see table I). Unfortunately, despite advocating in the guideline that patient information leaflets should be subject to user testing,[7] the EC appears to have failed to use such an evidence-based approach when formulating its own recommended terminology.

Table I. Probability estimates of adverse event occurrence in a UK student sample compared with frequencies assigned by the European Commission (EC) according to the terminology used (EC guideline-recommended descriptors) to describe frequency of occurrence of adverse events[8]

EC guideline-recommended verbal frequency descriptor	Mean (SD) probability estimate of frequency assigned by study participants (n = 268)	Frequency assigned by the EC guideline
Very common	65% (24.2%)	>10%
Common	45% (22.2%)	1–10%
Uncommon	18% (13.0%)	0.1–1%
Rare	8% (7.5%)	0.01–0.1%
Very rare	4% (6.7%)	<0.01%

The fact that the EC did not base its selection of the five verbal descriptors on systematic empirical evidence is a concern for two reasons. Firstly, substantial research in the medical field, as well as in other areas, has shown that there are large individual differences in people's interpretation of terms that are commonly used to describe risk probability, even in relatively restricted domains.[9-13] Bryant and Norman,[9] for example, found that physicians' interpretations of the term 'likely' (with regard to risk probability) ranged from 25 to 75%. More recently, Timmermans[12] reported that interpretations of the term 'very likely' ranged from 30 to 90%, even when presented in a restricted medical context. The second reason for concern is that other studies[14-17] have shown that people's perception of risk and their stated likelihood of taking a particular medicine are influenced by the information that they are given about the medication's adverse effects (and the way that it is presented). Thus, if patients incorrectly interpret the EC guideline-recommended verbal descriptors, then this may well result in their misper-

ceiving the risk associated with taking a particular medicine and affect their decision whether or not to follow the prescribed course of treatment.

This current paper provides an overview of a number of studies that we have carried out to evaluate the utility of the EC guideline-recommended descriptors. Section 1.1 describes four experiments that investigated members of the general public's interpretation of the descriptors in relation to mild and severe adverse effects. In order to increase the validity of our conclusions, section 1.2 reports a study carried out with patients actually taking medicines described in the patient information leaflet (and to which the adverse effect information refers). Given the considerable increase in sales of over-the-counter (OTC) medicines, section 1.3 summarises a study which assessed how people interpret one of the EC-recommended descriptors in the context of a patient information leaflet describing a common anti-inflammatory medicine. The study also evaluates a second aspect of the EC guideline concerning how best to inform people about what to do should they experience adverse effects. The remaining sections of the paper consider the advantages and limitations of other risk scales, as well as addressing the role that the patients' experiences and other contextual factors might have in influencing their perception of risks and benefits of medicine taking. The paper ends with a discussion of some of the broader issues involved in communicating risk of medication adverse effects.

1. An Evaluation of the European Commission (EC) Guideline-Recommended Descriptors

Given the concerns outlined in the section above, we decided to carry out a series of experiments to investigate people's interpretations of the EC guideline-recommended descriptors and to assess how use of the descriptors impacts on their perception of risk and views about medicine taking.[8,18,19]

1.1 The Hypothetical Medicines Studies

1.1.1 Experiment 1

In an initial experiment,[8] 268 students from Reading University, UK, were presented with a simple but realistic scenario about visiting the doctor and being diagnosed with either a throat or ear infection. They were prescribed a (hypothetical) short-course antibacterial agent, which was said to be associated with five adverse effects. The probability of occurrence of each adverse effect was described using one of the five EC guideline-recommended descriptors (in a fully counterbalanced design), and participants were asked to estimate the percentage of the population (or the number of people out of 10 000) who would experience the adverse effect if they took the medicine.

The results showed no effect of the type of response (percentage or frequency). The combined means are shown in table I, together with the EC guideline-designated probability ranges for each descriptor. In this table, it can be seen that there is a considerable discrepancy between our participants' interpretation of the five probability terms and the corresponding assigned frequencies. For example, the EC guideline recommends that the term 'common' should be used to describe probabilities between 1 and 10%.[5] Our participants produced a mean probability estimate for this descriptor of 45%.

1.1.2 Experiment 2

The second study[8] used a broader participant sample (112 members of the UK general population), and compared two versions of a short explanation about a medication and its adverse effects. Similar to the methodology of experiment 1, the explanation was presented in the context of a scenario about visiting the doctor and being prescribed some medicine. In one version, people were told that the adverse effects occur in 15% of people who take the medicine; in the other version they were told that the adverse effects were very common, the term designated by the EC guideline for that probability. We chose 15% for the numerical condition on the grounds that it was sufficiently above the lower bound of the assigned

Table II. Probability estimates of adverse event occurrence and mean ratings for Likert-scale variables in a UK general population sample according to the terminology used to describe frequency of occurrence of adverse events (results of 'Experiment 2')[8]

	Terminology used[a]	
	'very common' (n = 56)	'15%' (n = 56)
Mean (SD) probability estimate for the occurrence of an adverse event	64.4% (20.2%)	20.0% (11.2%)
Likert scale variables[b]		
Satisfaction with information presented	3.1	3.7
Perceived severity of adverse effects	4.5	3.7
Perceived likelihood that an adverse event will occur	4.4	2.5
Perceived risk to health posed by an adverse event	3.8	2.8
Perceived likelihood of taking the medication	3.1	4.2

a The likelihood of the occurrence of an adverse event was described to participants either by using the description suggested by the European Commission guideline terminology of 'very common' or by using a corresponding numerical value of 15%.

b Variables were rated on a scale from 1–6, with 1 being the lowest rating and 6 being the highest rating.

probability range for 'very common' in the EC guideline framework (i.e. 10%) but would nevertheless be considered to be realistic, given that few adverse effects occur in much more than 10% of people who take a medicine. People were asked to rate (on a 6-point Likert scale) how satisfied they were with the information presented, how likely it was that they would experience the adverse effects if they took the medicine, the risk of doing so to their health, and their likelihood of taking the medication.

The main findings of experiment 2 are shown in table II. Analyses of variance confirmed that there was a significant difference between the verbal ('very common') and numerical (15%) conditions on all six measures. Thus, those who were told that the adverse effects were 'very common' were less satisfied with the information, perceived the adverse effects to be more severe and more likely to occur, and the perceived risk to their health to be greater than those who were told that the effects occurred in 15% of people who took the medication. The verbal descriptor also resulted in significantly lower ratings of intention to comply. Importantly, the average probability value assigned to the 'very common' term was very similar to that found in experiment 1 (64%).

1.1.3 Experiment 3

One limitation of experiment 2 was that the sample included few people who were aged >60 years. However, people in this age group, both in Europe and the US, are more likely to be taking prescribed medicines than their younger counterparts.[20-23] The UK National Service Framework for older people[24] indicates that 80% of people aged >75 years take at least one prescribed medicine, and 36% take four or more. We therefore decided to replicate experiment 2 using a sample of the general population (120 participants) that was balanced across three different age groups (18–40, 41–60, and 61–80 years).[19]

The results confirmed the findings of experiment 2, in that people's mean probability estimate when given the descriptor 'very common' was 69%. Furthermore, there were again significant differences between the verbal and numerical conditions on participants' judgements of satisfaction, adverse effect severity, likelihood of adverse effect occurrence, perceived risk to health, and intention to comply. Importantly, these effects occurred across all three age groups.

1.1.4 Experiment 4

Given that within the EC guideline framework the descriptor 'very common' is associated with the widest band of incidence rates (>10–100%), the final experiment[19] examined people's interpreta-

tion of two terms, 'common' and 'rare', whose bands have upper and lower limits, making it possible to identify over-estimations as well as underestimations in judgements. The experiment also examined whether people's judgements were affected by whether the scenarios described either relatively severe adverse effects (e.g. chest pain, convulsions) or mild effects (e.g. increased thirst, tiredness). 360 members of the UK general population took part in the investigation. Similar scenarios were used to those employed in experiments 2 and 3. The adverse effects were described as being 'common' (or 'rare') in people who took the medicine, or occurring in 2% (or 0.02%) of people who took the medication.

Similar to the results of experiments 1, 2 and 3, analyses of variance measures confirmed that people who received the EC guideline-recommended verbal descriptors reported significantly higher probability estimates, were significantly less satisfied with the information, rated adverse effect severity and risk to health to be greater, and said that they would be significantly less likely to take the medicine than those given the comparative numerical conditions (see table III). A particularly striking finding was that only 7 out of 180 participants who were given a verbal descriptor provided probability estimates that fell within the EC guideline-assigned range (1–10% for 'common', and 0.01–0.1% for 'rare'). Virtually all remaining participants produced values that were higher than the upper bound limits of the assigned range. As in our recently published work,[16] scenarios that described mild adverse effects, compared with those that described severe effects, resulted in participants reporting greater satisfaction with the information, less perceived risk to health, and higher intention to comply with taking the medication. In addition, people who received verbal descriptors ('common' or 'rare') reported significantly higher probability estimates when the scenarios described mild adverse effects (56% for 'common' and 25% for 'rare') compared with severe adverse effects (45% for 'common' and 19% for 'rare').

1.1.5 Conclusions from the Hypothetical Medicines Studies

The results were remarkably consistent across the four experiments. The effects (whether the scenarios referred to relatively high or low probabilities of adverse event occurrence) held for all EC guideline descriptors tested, were present in a student population and three general population samples, and were consistent across three different age groups (a total of nearly 850 participants).

In all four experiments, the use of the qualitative descriptors recommended by the EC guideline was not only associated with a considerable variability in responses, but also led to a gross overestimation of risk. Taken together, the findings provide a clear message for the producers of patient information leaflets that accompany medicines. Compliance with the EC guideline is likely to lead

Table III. Probability estimates of adverse event occurrence and mean ratings for Likert-scale variables in a UK general population sample according to the terminology used to describe the frequency of the occurrence of adverse events (results of 'Experiment 4')[19]

	Terminology used[a]			
	'common' (n = 90)	'2%' (n = 90)	'rare' (n = 90)	'0.02%' (n = 90)
Mean (SD) probability estimate for the occurrence of an adverse event	50.5% (24.4%)	9.5% (14.2%)	21.5% (17.7%)	6.8% (15.4%)
Likert scale variables[b]				
Satisfaction with information presented	3.73	3.98	3.71	4.12
Perceived severity of adverse effects	3.60	3.11	3.22	2.76
Perceived risk to health posed by an adverse event	2.77	2.08	2.38	2.15
Perceived likelihood of taking the medication	4.29	5.07	4.71	5.10

a The likelihood of the occurrence of an adverse event was described to participants either by using descriptions suggested by the European Commission guideline definition of 'common' or 'rare' or by using corresponding numerical values of 2% or 0.02%.

b Variables were rated on a scale from 1–6, with 1 being the lowest rating and 6 being the highest rating.

to a significant overestimation of the probability that adverse effects will occur. This, in turn, is likely to affect patient judgements of perceived risk to health and decisions about whether or not to adhere to the prescribed course of treatment.

The findings are also in line with two earlier studies carried out in The Netherlands[25] and Germany,[26] which showed that people may not interpret verbal probability labels in the way that is intended by the communicator. Ours are the first experiments, however, to provide evidence that these interpretational difficulties can have a negative impact on people's perceptions of risk and intended health behaviours.

One interesting and unexpected effect which appeared consistently in the experiments was that many people assigned to the numerical conditions did not respond with the percentage given in the scenario, but tended to provide higher estimates. For example, those who were told that the adverse effects occur in 15% of people who take the medicine provided mean estimates of around 20%. We are currently investigating this effect to determine its underlying cause. One possibility is that people do not completely trust the information that is provided in patient information leaflets.[27]

1.2 The Real Medicines Study

A potential criticism of our work described in section 1.1 is that the experiments were carried out with students and members of the general population, rather than with patients reading leaflets about actual medicines they have been prescribed. Although we tested a broad sample of the population and selected scenarios that would readily apply to them (i.e. visiting their general practitioner, being diagnosed as having a relatively minor infection, and being prescribed a short-course antibacterial agent), it is possible that different results would have been obtained with real patients. We therefore carried out three studies with patients in a hospital cardiac rehabilitation clinic and with adult patients who were visiting a community pharmacy to obtain their prescription medications.[28] In the largest of the studies we investigated patients' interpretations of two of the EC guideline-recommended descriptors in relation to adverse effects associated with two commonly prescribed HMG-CoA reductase inhibitors (simvastatin and atorvastatin). The incidence rates of each adverse effect were extracted from systematic reviews and trials published at the Cochrane Library.[29] 120 patients were given information about the risk of pancreatitis (either as a verbal descriptor, rare or as a numerical value, 0.04%) or constipation ('common' or 2.5%). Our findings were similar to those found in the hypothetical medicines studies, with people providing mean probability estimates of 34 and 18% for the effects described as common and rare, respectively. Patients assigned to the verbal conditions also reported significantly higher ratings of perceived risk to health than those assigned to the numerical conditions.

1.3 Over-the-Counter Medicines

The hypothetical and real medicines studies involved scenarios/situations where a general practitioner or hospital doctor had prescribed the medicines in question. However, a substantial and increasing proportion of medicine is bought over the counter, both in the UK and throughout the developed world. In the year 2000, sales of OTC medicines in the UK amounted to £1.6 billion and were equal to one-third of the NHS drugs bill.[30] In line with this, the EC directive[6] does not differentiate between prescription and OTC medicines; both require users to be given comprehensive and understandable information about the adverse effects associated with medications. We therefore thought it important to test whether our earlier findings, in relation to the interpretation of the EC guideline-recommended probability labels with prescribed medicines, also applied to situations involving OTC medicines.[31]

We also evaluated a second aspect of the EC guideline, which concerned actions to take should adverse effects be experienced.[5] Clearly, it is important that consumers are not just given information about adverse effects and their likelihood, but are also advised what to do should adverse effects

occur. The EC guideline recommends that, in such a situation, if the consumer needs to seek help urgently then the term 'immediately' should be used. In contrast, if the situation is less urgent, then the recommended term is 'as soon as possible'. Unfortunately, as with the advocated probability descriptors, the recommendation was not based on empirical evidence.

We therefore carried out a two-phase study (D.C. Berry, unpublished data) in which 188 members of the UK general population were given a scenario about visiting their local pharmacy with a stiff neck and purchasing ibuprofen (an anti-inflammatory medicine). In the first phase of the study the participants were told the patient information leaflet stated that the medicine was associated with an adverse effect (stomach discomfort or pain). The probability of occurrence was said to be either 6% or common (the EC-recommended descriptor for that incidence rate). In addition to estimating the probability that they would experience the adverse effect, people rated the information on a number of scales including perceived risk to health and intention to take the medicine. In the second phase, they were told that the medicine was also associated with a rarer adverse effect (wheezing or unexplained shortness of breath), and if they experienced this effect they should seek medical help either 'immediately' or 'as soon as possible' (depending on the experimental condition). They then had to select, from a given set, which action they would take, for example, return to the pharmacy or go to the hospital casualty department.

The results of the first phase reflected the findings of our earlier studies on prescribed medicines.[18] People who were told that the adverse effect was 'common' estimated the probability of occurrence as 57% compared with 19% for those who were given the numerical value. They also rated their perceived risk to health to be significantly higher and their intention to take the medicine to be significantly lower than those assigned the numerical presentation condition. Interestingly, only one of the 94 participants given the verbal descriptor produced a probability estimate that fell within the EC-designated band for common (i.e. 1–10%).

In terms of the second phase of the study, there was no significant difference in people's interpretations of 'immediately' and 'as soon as possible'. Thus, from this study, there was no evidence to support the EC guideline's recommendation that the former term should be used when consumers need to seek medical help urgently and the latter when the need is less urgent.[5] Again, the EC guideline appears to be premature in recommending the use of particular terms or phrases, without assessing how people actually interpret their meaning.

2. Evaluating the Calman Risk Scale

Taken together, the studies of the hypothetical, real, and OTC medicines provide a clear and consistent message. Use of the EC guideline-recommended terms to inform patients about the incidence of adverse effects is likely to lead to a significant over-estimation of the probability that adverse effects will indeed occur. This, in turn, is likely to affect judgements of perceived risk to health and decisions about whether or not to take the medicine in question. Does this mean that verbal descriptors should never be used in such situations? Some researchers have argued that there are considerable advantages to using verbal labels, despite the variability in how they are interpreted.[32,33] Windschitl and Wells,[33] for example, suggested that verbal descriptions of uncertainty have the advantage that they allow for more associative and intuitive thinking. They pointed out that there are many everyday situations (e.g. deciding whether or not to walk home late at night) in which human decisions and behaviours are not based on deliberate and rule-based thinking, and that the use of numerical descriptions may misrepresent how individuals think about uncertainty in such situations. They reported three experiments showing that, relative to numerical values, verbal descriptors were more sensitive to manipulations of context and framing, and were better predictors of people's preferences and behavioural intentions.

Thus, it could be argued that verbal descriptors should be used, and that maybe the EC guideline[5] has simply attached inappropriate verbal labels to the specified probability ranges. Indeed, Calman[34] (former UK Government chief medical officer) advocated a different set of verbal labels (high, moderate, low, very low, minimal, and negligible) to describe the risk of occurrence of medication adverse effects. For example, he recommended that the term 'high risk' should be used for probabilities in the range to which the EC has assigned the terms 'very common' and 'common', and 'low risk' for the probability range to which the EC guideline has linked to 'rare'. Again, however, the recommendations were not based on empirical evidence, and a recent study by our research group has shown a significant divergence between Calman's recommendations and people's actual interpretations.[31] For example, our participants (160 university students) produced a mean probability estimation of 30% for the term 'moderate', whereas Calman recommended that the term be used for the probability range of 0.1–1%. Furthermore, our participants had particular difficulty in distinguishing the meaning of the three lowest frequency terms, producing mean probability estimates of 10, 9 and 8% for the terms 'very low', 'minimal' and 'negligible', respectively. There was also a fair amount of variability in people's interpretations, similar to our hypothetical and real medicines studies of EC descriptors,[8,19] suggesting that it may be difficult to produce a set of labels that will be interpreted in a consistent way by a diverse group of people.

3. Doctors' Interpretation of the EC Guideline and Calman Verbal Descriptors

The findings of the above studies show that members of the general public significantly overestimate the likelihood of adverse effects occurring, when interpreting the descriptors that were recommended by both the EC guideline[5] and Calman.[34] Is this because they have no accurate knowledge about typical incidence rates and believe that adverse effects in general occur with much greater frequency than is typically the case with most medicines?

This leads to the further question of whether doctors, who have a much better knowledge about typical incidence rates of adverse effects from medication, would interpret the EC guideline and Calman-recommended descriptors in a way that is more in line with the intentions of those who formulated the scales. Previous studies in the literature[12] have shown no effect of the level of expertise of doctors when interpreting verbal probability terms, but one might expect that the judgements of people with medical training would be more realistic than those of people with no training or experience. There is some suggestion in the literature that this might be the case,[9,35] but the evidence is inconclusive.

Two further studies were carried out,[31] with hospital doctors from a range of different specialities.[1] Fifty-eight doctors were presented with the following scenario: they are reading an article about a new antibacterial agent, which is associated with a number of mild or severe (depending on the experimental condition) adverse effects. As with our studies with the general public, the doctors were required to produce numerical estimates for the given verbal descriptors.

The results showed that, contrary to our hypothesis, the doctors significantly over-estimated the probability of adverse effects occurring, given each verbal descriptor, for both the EC and Calman scales. However, the degree of over-estimation was not as great as in the studies with the non-medically trained participants. For example, the doctors provided a mean estimate of 46% for both 'very common' (EC descriptor) and 'high' (Calman descriptor), compared with means of 65 and 52%, respectively, for the participants who had no medical training. Similarly, the doctors provided a mean estimate of 2% for both 'rare' (EC) and 'very low' (Calman), compared with 8 and 10%, respec-

1 This work was carried out in collaboration with Dr Wendy Holden from the Royal Berks and Battle Hospitals Trust.

tively, reported by the non-medically trained participants.

4. Other Risk Scales

It has been suggested by some researchers and health professionals[4,36] that the communication of risks can be improved by relating potentially hazardous events to events or concepts with which people are familiar. Calman and Royston,[4] for example, proposed a community risk scale that uses the size of various communities as risk comparisons. In this way, the risk of a particular adverse effect can be described in terms of it happening to approximately one person in a street or in a town. Paling[36] used the risks of familiar events such as being in a car accident, a murder victim, or struck by lightening as comparators. Clearly, in order for such scales to be effective, people need to have relatively accurate and consistent understanding of the likelihood or size of the comparative items. In line with this, Barclay, Constigan and Davies[37] pointed out that, given the popularity of the lottery in the UK, many people (with a broad range of backgrounds and educational levels) would have experience of buying lottery tickets and/or of watching the weekly draw on television. They therefore suggested that a lottery-based scale might be an effective way of conveying risk information. Based on the figures supplied by the lottery organisers, Barclay et al.[37] designed a near-logarithmic scale showing the probability of matching a number of balls for a £5 stake. For example, a three-ball prize (£10) corresponds to a risk of between 1 in 10 and 1 in 100. They suggested that many people might be expected to have knowledge of how likely these winning events would be, which would enable them to make effective risk comparisons.

In an attempt to find empirical support for the suggestion of Barclay et al.,[37] we asked 132 members of the general population to estimate the probability, given a £5 stake, of their having three, four, five or six numbers match the winning lottery numbers. Half of the sample bought lottery tickets every, or nearly every, week. Participants indicated their judgements by selecting one of the six verbal descriptors from Calman's scale, and by selecting one of six given numerical frequency bands. We found that people were very poor at selecting the correct probabilities, when presented in either verbal or numerical format. This was particularly the case for the three- and four-ball matches (i.e. the higher probability events), with <5% of participants providing correct responses. The majority of participants greatly under-estimated the probabilities, believing that their chances of winning were far less than they were in practice. Interestingly, there was no difference at all between the responses of those who were regular lottery players and those who were not. Thus, we could find no evidence to support the use of a lottery-based scale for communicating risk probability.

5. Concluding Comments

The findings from the broad range of studies we have reviewed in this paper confirm the statement made at the outset, that the problems of communicating complex medical information involving risk and uncertainty are considerable. Some advances might be made by using appropriate combinations of verbal and numerical information, building on the advantages of the two methods,[38] or by using graphical methods of presentation,[39,40] but we are still a long way from achieving a standardised language of risk.[41] Moreover, as Edwards and Elwyn[42] have recently acknowledged, this may not, in fact, be a sensible goal for researchers and communicators. Standardising the language does not allow the flexibility required for dealing with people with different levels of literacy, numeracy, and attitudes. 'The nature of risk, its burden, the context, and the timeframe over which one has to live with it are all important determinants of how individuals interpret risk for themselves and of whether they choose to accept it.'[42] For example, one cannot easily separate consideration of the size of risk from the value attached to harm, i.e. its meaning and implications for a person's life. Similarly, one should not think about possible risks associated with taking particular medicines in iso-

lation from the likely beneficial effects. For most people there will be only a single benefit that is sought but the potential risks are often multiple.[43] Finally, it is not clear that it is feasible, or justifiable, to apply estimates of risk derived from populations to individuals. People often have an 'all or nothing perspective to harm and risk',[44] as illustrated by the patient who questioned the doctor, 'what do you mean by saying that if I have the treatment, the risk of stroke is 5%. Surely, I will not have 5% of a stroke. I will either have a stroke, which is 100%, or I will not, which is 0%.'

Clearly, the biggest challenge for risk communicators and for future research will be to produce a standardised language of risk (and benefit) that is sufficiently flexible to take into account different perspectives, as well as changing circumstances and contexts of illness and its treatment. In the meantime, the EC and other legislative bodies and health professionals should stop advocating the use of particular verbal labels or phrases until there is a much more solid evidence base to support their use.

Acknowledgements

No sources of funding were used to assist in the preparation of this manuscript. The authors have no conflicts of interest that are directly relevant to the contents of this manuscript.

References

1. From compliance to concordance: achieving shared goals in medicine taking. London: Royal Pharmaceutical Society of Great Britain and Merck, Sharp & Dohme, 1997
2. National plan for the NHS. London: Department of Health, 2000
3. Doyal L. Informed consent: moral necessity or illusion? Qual Health Care 2001; 10: 29-33
4. Calman K, Royston G. Risk language and dialects. BMJ 1997; 315: 939-42
5. European Commission. A guideline on the readability of the label and package leaflet of medicinal products for human use. EC Pharmaceuticals Committee, 1998
6. European Commission. European Commission Council Directive 92/27/EEC (OJ NoL 113 of 30.4.1992, p.8) 1992
7. Dickinson D, Raynor DK Duman M. Patient information leaflets for medicines: using consumer testing to determine most effective design. Patient Educ Couns 2001; 43: 147-59
8. Berry DC, Knapp PR, Raynor DK. Is 15% very common: informing people about the risks of medication side effects. Int J Pharm Pract 2002; 10: 145-51
9. Bryant GD, Norman GR. Expressions of probability: words and numbers [letter]. N Engl J Med 1980; 302: 411a
10. Mazur DJ, Merz JF. Patients' interpretations of verbal expressions of probability: implications for securing informed consent to medical interventions. Behav Sci Law 1994; 12: 417-26
11. Reagan RT, Mosteller F, Youtz C. Quantitative meanings of verbal probability expressions. J Appl Psychol 1989; 74: 433-42
12. Timmermans D. The roles of experience and domain of expertise in using numerical and verbal probability terms in medical decisions. Med Decis Making 1994; 14: 146-56
13. Wallsten S, Budescu DV, Rapoport A, et al. Measuring the vague meaning of probability terms. J Exp Psychol Gen 1986; 115: 348-65
14. Berry DC, Michas IC, Gillie T, et al. What do patients want to know about their medicines and what do doctors want to tell them? A comparative study. Psych Health 1997; 12: 467-80
15. Berry DC, Michas IC, DeRosis F. Evaluating explanations about drug prescriptions: effects of varying the nature of information about side effects and its relative position in explanations. Psych Health 1998; 13: 767-84
16. Berry DC, Michas IC, Bersellini E. Communicating information about medication side effects: effects on satisfaction, perceived risk to health and intention to comply. Psych Health 2002; 17: 247-67
17. Berry DC, Michas IC, Bersellini E, Communicating information about medicine: the benefits of making it personal. Psych Health. In press
18. Berry DC, Knapp PR, Raynor DK. Provision of information about drug side effects to patients. Lancet 2002; 359: 853-4
19. Berry DC, Raynor DK, Knapp PR. Communicating risk of medication side effects: an empirical evaluation of EC recommended terminology. Psych Health Med. In press
20. Cartwright A. Medicine taking by people aged 65 or more. Br Med Bull 1990; 46: 63-76
21. Claesson S, Morrison A, Wertheimer AI, et al. Compliance with prescribed drugs: challenges for the elderly population. Pharm World Sci 1999; 21: 256-9
22. Kane RL, Ouslander JG, Abbrass IB. Essentials of clinical geriatrics. New York: McGraw Hill, 1994
23. McElnay JC, McCallion R. Adherence and the elderly. In: Myers LB, Midence K, editors. Adherence to treatment in medical conditions. London: Harwood, 1998
24. Department of Health. UK National Health Service Framework for Older People. 2001
25. Pander Maat H, Klaasan R. Side effects of side effect information in drug information leaflets. J Tech Writing Commun 1994; 24: 389-404
26. Fischer K, Jungermann H. Rarely occurring headaches and rarely occurring blindness: is rarely rarely? The meaning of verbal frequentistic labels in specific medical contexts. J Behav Dec Mak 1996; 9: 153-72
27. Bennett P. Communicating about risks to public health: pointers to good practice. London: Department of Health, 1998
28. Knapp PR, Berry DC, Raynor DK. Testing two methods of presenting side effect information about common medicines [abstract]. Int J Psychoanal Psychother 2001; 9: R6
29. Cochrane Library. Update Software [online]. Issue 3, 2000; Oxford. Available from URL: http://www.update-software.com/cochrane/cochrane-frame.html [Accessed 2002 Dec 10]
30. Annual Report, Proprietary Association of Great Britain. London: PAGB, 2001

31. Berry DC, Bersellini E. Doctors' and patients' understanding of the Calman risk scale. Proceedings of the British Psychological Society. In press
32. Kuipers B, Moskowitz AJ, Kassierer JP. Critical decisions under uncertainty: representation and structure. Cognit Sci 1988; 12: 177-210
33. Windschitl PD, Wells GL. Measuring psychological uncertainty: verbal versus numeric methods. J Exp Psych App 1996; 2: 343-64
34. Calman KC. Cancer: science and society and the communication of risk. BMJ 1996; 313: 799-802
35. Edwards A, Elwyn GJ, Stott N. Researchers should present results with both relative and absolute risk. BMJ 2001; 318: 7183
36. Paling J. Up to your armpits in alligators: how to sort out what risks are worth worrying about. Gainesville (FL); Risk Communication and Environmental Institute, 1997
37. Barclay P, Costigan S, Davies M. Lottery can be used to show risk. BMJ 1998; 316: 1242-3
38. Weber EC, Hilton, DJ. Contextual effects in words: perceived base rate and severity of events. J Exp Psychol Hum Percept Perform 1990; 16: 781-9
39. Lipkus IM, Hollands JG. The visual communication of risk. J Nat Cancer Inst Monogr 1999; 25: 149-63
40. Edwards A, Elwyn GJ, Mulley A. Explaining risks: turning numerical data into meaningful pictures. BMJ 2002; 324: 827-30
41. Calman K. Issues of risk: this unique opportunity [The William Pickles lecture]. Br J Gen Pract 2001; 51: 47-51
42. Edwards A, Elwyn GJ. Risks: listen and don't mislead. Br J Gen Pract 2001; 51: 259-60
43. Edwards R, Wiholm BE, Martinez C. Concepts in risk benefit assessment: a simple merit analysis of a medicine. Drug Saf 1996; 15 (1): 1-7
44. Edwards A, Pill R, Stott N. Use of standard terms is unlikely to result in standard communication [letter]. BMJ 1996; 313: 1483

Correspondence and offprints: Professor *Dianne C. Berry*, Department of Psychology, University of Reading, Earley Gate, Whiteknights, Reading, RG6 6AL, UK.
E-mail: D.C.Berry@Reading.ac.uk