Archives of Disease in Childhood 1990; **65**: 520–523

How do parents of babies interpret qualitative expressions of probability?

N J Shaw, P R F Dear

Abstract

One hundred mothers and 50 medical students and doctors were interviewed to examine their understanding of eight common probability expressions as they are used in statements of the kind that doctors regularly make to mothers of babies. Respondents were asked to translate each term into a number on a scale from 0 to 10. There was a wide range of interpretation of the expressions, and significant differences were found between the mothers and medical group for seven of the expressions. Most mothers expressed a preference for receiving information in numerical terms.

We advise that verbal probability expressions should be restricted, and more information presented in numerical terms.

Effective communication between doctors and patients and patient's relatives is important, and the trend towards a greater respect for patients' autonomy makes it even more so. Qualitative expressions of probability such as 'probably' and 'likely' are often used in everyday language, but are also used by doctors to express views about diagnosis and prognosis. Sometimes these words are redundant, but at other times their meaning is critical to the correct interpretation of the information being transmitted. Previous studies in different professional groups have shown a wide variation of interpretation of such expressions,[1–15] but little research has been done into how non-professionals interpret them, and only one study has interviewed parents of child patients.[1]

The study that we report was designed to look at parents' understanding of probability expressions. There were two aims: first, to determine parents' understanding of eight common expressions of probability and to see if their understanding differed from that of doctors and, secondly, to find out if parents would prefer to receive information in numerical terms.

Subjects and methods

The first part of the study consisted of interviewing 100 mothers of babies in the postnatal wards over a six month period. They were all interviewed alone by the same person (NJS). After a preamble in which the purpose of the study was outlined, each mother was asked to consider eight statements given in the context of an imagined setting with her baby. Each statement contained a single expression of probability (table 1).

Table 1 Statements used to assess understanding of probability expressions

1 Your baby is jaundiced and will **probably** need treatment with phototherapy lights.
2 Your baby is premature and it is **likely** that he/she will need tube feeding.
3 Your baby is **very likely** to bring back some of his feeds in the first few days of life.
4 Premature babies such as yours **occasionally** need to be put in an incubator to keep them warm.
5 It is **unlikely** that your baby will be ready to go home before the weekend.
6 It is **possible** that your baby will need to be admitted to the neonatal unit.
7 Babies born as early as yours are **rarely** able to go home within one week of birth.
8 **Sometimes** babies need to have a blood test if they are not feeding very well.

After hearing each statement the mother was asked to translate the expression of probability into a numerical value on a scale from 0 to 10. It was indicated that 0 out of 10 was equivalent to no chance that their baby would require whatever measure was being discussed, and 10 out of 10 meant that the baby would definitely require this measure. A similar interview, using the same statements, was conducted with 22 paediatric doctors and 28 medical students.

The second part of the study was undertaken with 81 of the mothers. Each mother was read two statements about a theoretical problem with her baby. These problems were typical of those that are seen with newborn babies—for example, early onset of jaundice, a heart murmur, a clicky hip, or failure to pass meconium.

One version of the statement presented the information using qualitative expressions of probability, and the second statement presented the same information using numerical values in place of the expressions of probability. For example: 'Your baby has a heart murmur which will *probably* go away on its' own. It is *possible*, however, that this murmur may be of importance and so we will need to see him again in our clinic'. 'Your baby has a heart murmur and there is a 7 out of 10 chance that it will go away on its' own. There is a 3 out of 10 chance, however, that this murmur may be of import-

Table 2 Numerical rating of probability expressions by mothers

Expression	*Range*	*Mean (SD)*	*Interquartile range*
Very likely	2–10	8·6 (1·5)	2·0
Likely	2–10	6·6 (1·7)	3·0
Probably	2–10	6·6 (1·7)	3·0
Possible	2–10	6·2 (1·7)	2·0
Sometimes	2–10	5·0 (1·6)	2·0
Occasionally	2–10	4·7 (1·8)	2·0
Unlikely	0–8·5	2·7 (1·8)	3·0
Rarely	0–8	2·0 (1·9)	2·5

Department of Paediatrics and Child Health, St James's University Hospital, Leeds
N J Shaw
P R F Dear

Correspondence to:
Dr N J Shaw,
Institute of Child Health, Alder Hey Children's Hospital, Eaton Road, Liverpool L12 2AP.

Accepted 19 December 1989

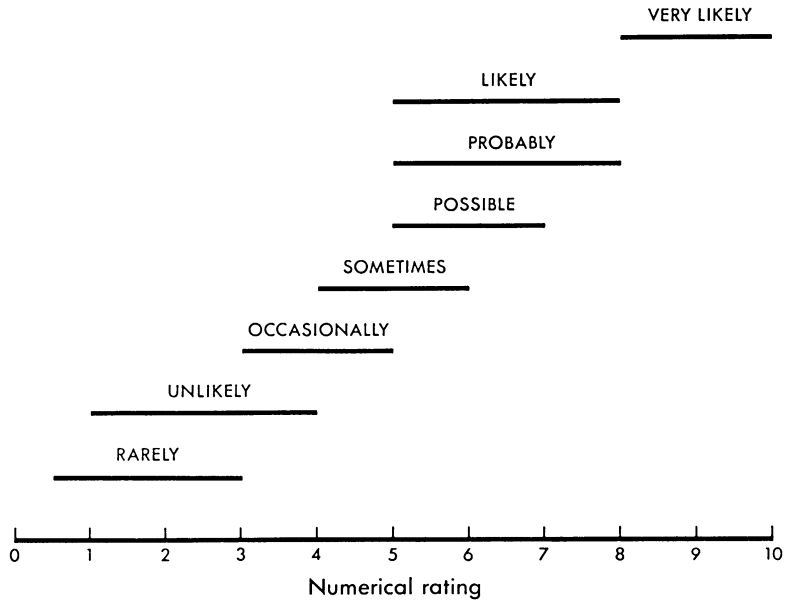

Figure 1 Numerical interpretation of probability expressions by mothers (interquartile ranges).

Table 3 Numerical ratings of probability expressions by doctors and students

Expression	*Range*	*Mean (SD)*	*Interquartile range*
Very likely	7·5–10	9·1 (0·5)	0·5
Likely	5–10	7·7 (1·3)	1·6
Probably	3–10	7·6 (1·4)	1·5
Possible	1–9	4·6 (1·9)	2·0
Sometimes	1·5–10	4·0 (1·5)	2·0
Occasionally	1–8	2·9 (1·4)	1·0
Unlikely	0–4	1·6 (1·0)	1·0
Rarely	0–2·5	0·8 (0·5)	0·5

ance and so we will need to see him again in our clinic'. Each mother was then asked in which form she would prefer to receive the information.

Statistical analysis of the results was done by calculating the mean for the numerical values and the SD and the interquartile range as an index of variation. The Mann-Whitney U test was used to compare the responses obtained from the mothers, and from the doctors and students.

A further analysis was done to see if the educational level of the mothers influenced the numerical rating of the expressions. They were divided into two groups. One group consisted of those mothers who had left school at the age of 16 and had not received any further education, and the other group of those mothers who had received education beyound the age of 16, whether at school or elsewhere. The Mann-Whitney U test again was used to look for significant differences in interpretation of the expressions.

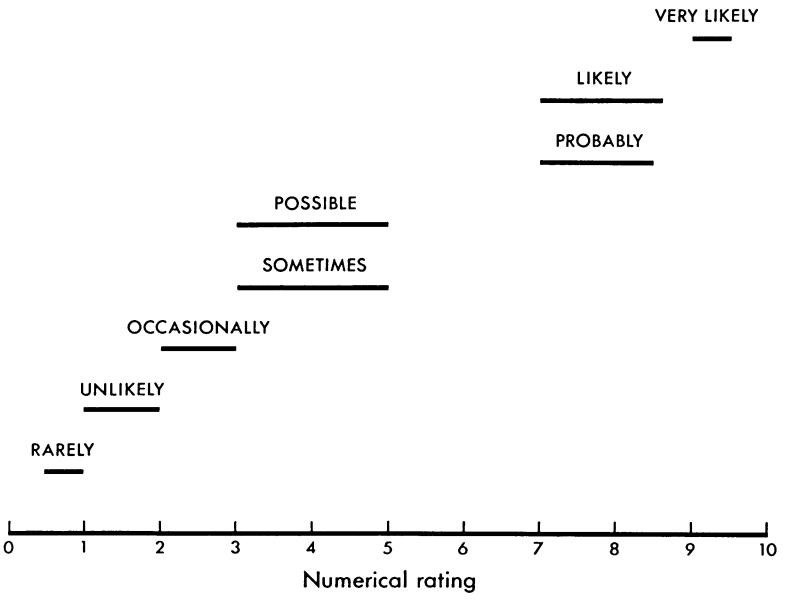

Figure 2 Numerical interpretation of probability expressions by doctors and students (interquartile ranges).

Results

The numerical values given by the mothers to each of the eight expressions are shown in table 2; each expression generated a wide range of interpretation. Most expressions had large SDs, and none of the interquartile ranges was less than 2. Figure 1 shows the interquartile ranges represented on a linear scale that shows the considerable degree of overlap of the values for the expressions. Most noticeable is the same range for the two expressions 'probably' and 'likely', and the pronounced overlap of the expressions 'probably' and 'possible'. There was a clear separation only of the expressions 'likely' and 'very likely'. It must also be borne in mind that as the interquartile range eliminates half the responses it is actually an underestimate of the variation seen.

Table 3 shows the results obtained from the doctors and medical students. Although there were smaller overall ranges at the extremes of the scale, four expressions in the middle of the scale ('probably', 'possible', 'sometimes', and 'occasionally') had wide ranges and correspondingly large SDs. The interquartile ranges were smaller for this group with only two expressions having values equal to 2; fig 2 shows them on a linear scale, which shows a better separation of the expressions, although two sets of expressions had the same ranges.

Comparison of the numerical values for each of the expressions between the professional and non-professional group showed significant differences in interpretation of seven of the expressions (p values $<0{\cdot}001$), the only exception being 'very likely'. When the educational level of the mothers was analysed as a variable, three expressions ('rarely', 'likely', and 'unlikely') showed a significant difference (at the 1% level) between the group of less educated mothers and the medical group that was not seen when the more educated mothers were compared with the medical group. For these three expressions, therefore, educational level did seem to affect interpretation.

Of the 81 mothers interviewed to see if there was a preference for information in numerical terms, 43 opted for a statement using numerical values compared with 30 who preferred a statement containing verbal expressions of probability. Eight mothers stated that they had no preference. Several of the mothers commented that they preferred numerical expressions as

they were more definite. One mother commented that she preferred verbal expressions as she felt the information using numerical terms frightened her more. One mother, whose baby had been diagnosed before birth as having a diaphragmatic hernia, said she was quoted a 50% chance of the baby dying, which she regarded as preferable to any verbal expression of probability.

Discussion

This study shows a large variation in interpretation of common expressions of probability. This phenomenon has been shown in several previous studies but in all but one of these the subjects were professional groups.[1-15] The study by Brun and Tergen (published in Norwegian) used a questionnaire sent to parents of children attending a nursery and, in addition to showing variations in interpretation of expressions of probability, they found that 72% of the parents felt that they understood expressions with numbers best and 66% said they would prefer the doctor to use numbers instead of words.[1]

Our study shows significant variations in the interpretation of the expressions among mothers and a group of doctors and medical students. It also shows significant differences in interpretation between these two groups, which has not previously been reported. The mothers expressed an overall preference for receiving information using numerical values to express probability.

The first question to ask is: Do these wide variations in interpretation that we and others have shown have any practical importance—that is, do they impair communication? It could be argued that the similar mean ratings and rankings that many studies have shown for expressions of probability provide a sufficiently common basis of understanding.[2 3] Several features of our study that differ from other published studies, do not lend support to the notion that a common code of understanding of probability expressions can easily be derived. Although (as in other studies) the numerical ratings of the doctors and students varied less at the extremes of the range, this was not true of the mothers who showed a wide range of interpretation of all of the expressions including 'very likely' and 'rarely'. In addition, although the mean ratings of the expressions by the medical group in our study are similar to those from professional groups in other published studies, the mean ratings of five of the expressions by the mothers ('possible', 'sometimes', 'occasionally', 'rarely', and 'likely') were considerably higher than those obtained from these other studies. As in other studies there was consistency in the ranking of the mean values of the expressions by the two groups, but this was not obtained at an individual level. Our study shows that significant differences in interpretation of such expressions do exist between different groups, and so the use of expressions of probability is of practical importance to communication.

There are many external factors that will influence a person's rating of expressions of probability, so that even the notion of consistent interpretation by an individual subject is questionable. The most obvious is the context within which such an expression is used.[4] Mapes showed that a high frequency context will increase the value of an expression whereas a low frequency context will decrease its value.[5] In our study, however, the same contexts were used for both groups and cannot account for the observed differences. The amount of 'post-secondary' education will also influence an individual subject's ratings, as shown in the study by Nakao and Axelrod[6] and also a significant factor in our study. This is a feature that exists in real life, however, where there may well be a difference in education between doctors and their patients and it is undoubtedly one of the main factors contributing to a lack of satisfactory communication between them.

It can be argued that in low key contexts such as the chances of a jaundiced baby requiring phototherapy the variation in interpretation of a verbal expression of probability is not critical. In the context of a discussion about the chances of disability in a preterm baby with an intraventricular haemorrhage, however, the correct interpretation of an expression of probability would be vital to the information being conveyed.

We believe that it is useful to make a distinction between information used for decision making, and information 'for the record'. If subjects are being given information on the basis of which they will have to make a decision it is vitally important that verbal expressions of probability are avoided. One example is in antenatal counselling about the risks of genetic disorders and the procedures that can be carried out to detect them. Another example is when a decision has to be made about the continuation of neonatal intensive care on the basis of predicted prognosis. Under these circumstances it is essential that the information conveyed is clear and cannot be misinterpreted.

The results obtained from our study, particularly from the group of mothers, make verbal expressions of probability invalid for the purpose of decision making. There is also increasing emphasis on the importance of informed consent for the purpose of investigations, operations, and research. The use of verbal expressions of probability in these circumstances would lead to wide variations in interpretation of risk.

There are many reasons why doctors may deliberately use verbal expressions of probability when talking to patients (or their parents), even if they recognise the problem of interpretation. One is a sense of uncertainty, leading to a deliberate attempt to resist using numerical values so as to avoid having the quality of the forecast judged. Another reason is a deliberate desire not to be too precise in the interests of the patients—for example, in protecting a person from the full impact of bad news. Expressions of probability can at least leave some room for hope and, in these circumstances, can be justified. Doctors may also avoid using numerical estimates because they may feel compelled to specify a precise number, but an acceptable

alternative would be to specify a range. The hazards associated with verbal expressions of probability stem from variability in their interpretation, whereas with numerical estimates—even if a range is used—the 'message received' is the same as the 'message sent'.

Alternatives have been proposed to the use of verbal expressions of probability. Some authors have argued that their use should be eliminated altogether and only numerical estimates should be used.[6–8] A non-medical example is the National Weather Service in the United States, which now expresses its forecast numerically. In practice, however, it would be difficult to avoid verbal expressions in all conversations between doctors and patients. Another proposal is that a common scale of numerical values for probability expressions should be devised so that—for example, if the term 'probably' was used—people would know that this indicated a 60–70% chance.[9] Although this might be feasible to make communication easier between members of the same professional group, it would be much harder to introduce into the general population.

A better suggestion is that when qualitative terms are used because of lack of hard data, individual doctors could subsequently clarify, in numerical terms, their best estimate of the value or range that they are trying to convey.[10 11] This would certainly be easier to institute and would reduce the variation in interpretation of verbal expressions.

We believe that it is important that doctors become more aware of the hazards of using verbal expressions of probability, and should be prepared to restrict their use and provide more information using numerical values. If this is not possible then at least they should be prepared to clarify what they mean in numerical terms when they use such an expression.

1 Brun W, Tergen KM. Sprak og sannsynligheter. *Tidsskr Nor Laegeforen* 1986;**106**:1967–8.
2 Kong AM, Octo-Barnett G, Mosteller F, Youtz C. How medical professionals evaluate expressions of probability. *N Engl J Med* 1986;**315**:740–4.
3 O'Brien BJ. Words or numbers? The evaluation of probability expressions in general practice. *J R Coll Gen Pract* 1989;**39**:98–100.
4 Pepper S, Prytulak L. Sometimes frequently means seldom: context effects in the interpretation of quantitative expressions. *Journal of Research in Personality* 1974;**8**:95–101.
5 Mapes REA. Verbal and numerical estimates of probability in therapeutic contexts. *Soc Sci Med* 1979;**13A**:277–82.
6 Nakao MA, Axelrod S. Numbers are better than words: verbal specifications of frequency have no place in medicine. *Am J Med* 1983;**74**:1061–5.
7 Robertson WO. Quantifying the meaning of words. *JAMA* 1983;**249**:2631–2.
8 Bryant GD, Norman GR. Expressions of probability: words and numbers. *N Engl J Med* 1980;**302**:411.
9 Beyth-Marom R. How probable is probable? A numerical translation of verbal probability expressions. *Journal of Forecasting* 1982;**1**:257–69.
10 Kenney RM. Between never and always. *N Engl J Med* 1981; **305**:1097–8.
11 Cohen J, Dearnley EJ, Hansel CEM. A quantitative study of meaning. *Br J Educ Psychol* 1958;**28**:141–8.
12 Toogood JH. What do we mean by 'usually'? *Lancet* 1980;**i**: 1094.
13 Lichtenstein S, Newman JR. Empirical scaling of common verbal phrases associated with numerical probabilities. *Psychonomic Science* 1967;**9**:563–4.
14 Simpson RH. Stability in meanings for quantitative terms: a comparison over 20 years. *Quarterly Journal of Speech* 1963; **49**:146–51.
15 Hadek MD. How often is often? *Am Psychol* 1968;**23**:553–4.