Verbal Expressions of Probability in Informed Consent Litigation

JON F. MERZ, MBA, JD, MAREK J. DRUZDZEL, MS, MS, DENNIS J. MAZUR, MD, PhD

Informed consent litigation provides a forum in which probabilistic evidence is elicited from physicians as parties or as expert witnesses. The authors reviewed over 450 medical informed consent opinions reported by both trial and appellate courts in all 50 states over 40 years to determine 1) the extent to which verbal expressions of probability were used by testifying physicians to characterize the risks of medical procedures; 2) when such expressions were used, whether consistent numeric interpretations of the terms were being applied by the physicians; 3) whether the choice of expression was influenced by the severity of the consequences associated with the particular risk; and 4) whether the use of such terms was correlated with trial outcomes, inasmuch as the duty to disclose a risk is said to increase with the magnitude of the risk and probability is one measure of such magnitude. It was found that subjective verbal expressions of probability are used in the litigation setting, and that such expressions represent broad ranges of numeric probabilities. There was some correlation between the expression and the represented numeric probabilities. In general, expressions such as "extremely low" and "low" corresponded to probabilities lower than those represented by terms such as "high" and "very high." Further, verbal expressions appeared to be influenced by the severity of the consequences associated with the risks, but whether this increases or decreases the ambiguity of verbal expressions in the communication process warrants further research. The authors suggest a syntax of verbal expressions of probability as a means to reduce the numeric ambiguity of these terms. Key words: verbal expressions; probability; syntax; communication; medicolegal; litigation; informed consent. (Med Decis Making 1991;11:273–281)

Kong et al.[1] and Wallsten et al.[2] have found some consistency in the usage of verbal expressions of probability by physicians and other health professionals. Other researchers, however, have found that such consistency is limited.[3-7] There is a continuing debate in the medical community over the use of verbal expressions of probability by physicians and an ongoing attempt by the medical-decision-making community to encourage physicians to begin the process of providing patients with numeric estimates of the probabilities of risks and benefits of medical procedures.

Medical informed consent (as well as other malpractice) litigation influences medical decision making and regulates physician behavior in securing patient informed consent by providing feedback on what disclosures are legally acceptable. Due to the potential importance of the use of verbal expressions of prob-

ability to both the medical and the legal communities involved with medical informed consent, we undertook an analysis of the existing body of published court opinions. Our study focused on four questions: 1) the extent to which verbal expressions of probability were used to characterize the risks of medical procedures by physicians testifying at trial either as defendant or as an expert witness for either the plaintiff or the defendant; 2) whether consistent numeric interpretations—i.e., agreement among physicians about the quantitative value represented by a given word—of the terms were being applied by the physicians when using such expressions; 3) whether the choice of expression was influenced by the severity of the consequences associated with the particular risk; and 4) whether the use of such terms was correlated with trial outcomes, inasmuch as the duty to disclose a risk is said to increase with the magnitude of the risk and probability is one measure of such magnitude.

This analysis studied the use of verbal expressions in wide probability ranges covering several orders of magnitude. Apparently, the application of verbal expressions to probabilities spanning several orders of magnitude has not been studied before. Furthermore, the use of these terms occurs in a natural setting (if trial can be considered such) as opposed to a laboratory situation, perhaps more accurately reflecting the users' subjective interpretations.

273

Method

We reviewed over 450 appellate and trial court opinions involving informed consent that were issued in the 40-year period from 1951 through 1989. We estimate that this review included at least 90% of the published opinions dealing with informed consent issues—uncertainty arises from incomplete cross-citations in the legal literature and in Shephard's index (the primary means for searching for subsequent cases and articles that cite a particular case).

The review yielded 110 cases where verbal expressions of the probabilities of risks were recited by the court, often explicitly referring to the testimony of a physician at trial. Physicians may provide testimony at trial as defendants or as expert witnesses for either side—no distinction was made between which witness used the term. Inasmuch as testimony on the risks must be elicited from experts, we might guess that verbal expressions may have been used in other cases, but these data reflect only those cases where the reported (predominantly appellate court) opinion recited an expression. To limit the analysis to expressions used by physicians at trial, cases were omitted where the judges or other parties (plaintiffs or attorneys for either side) themselves used an expression to characterize a probability. Also excluded were instances where verbal expressions were used to quantify the likelihood of causation, such as "more probable than not." This is an opinion offered to prove legal cause. The expert's opinion may have little or no objective basis, and experts are taught the terminology so that the plaintiff will meet his or her burden of proof in establishing a causal link between the defendant's negligence and the plaintiff's injury.

Expressions only of the probability or frequency of occurrence of the risk were of interest, while expressions reflecting the magnitude of the risk, such as "substantial" or "minimal" *risks*, which implicitly include the speaker's subjective valuation of the consequences, were excluded from the analysis. We also attempted to avoid situations where a verbal expression was used in a relative sense—e.g., the probability of a particular risk might be described as "rare" in relation to another, more highly likely, risk of the same treatment—we excluded any expressions stated in an explicit comparison of one risk with others. Further, six terms implicitly indicative of a comparison (e.g., "most common," "relatively low," and "comparatively small") also were dropped. This resulted in a total of 95 cases where verbal expressions of probability were available for our study.

Fifty-five of these cases included a quantitative estimate of the probability, or a numeric estimate applicable to the facts of the case was applied from a similar case. Numeric estimates were sometimes given by the same witness who stated the verbal expression (e.g., "rarely, one-out-of-a-hundred")[8] and sometimes by different witnesses. We assumed that the reported numeric estimates were accurate; thus, no secondary sources from the medical literature were used to quantify the probabilities. In the cases where a quantitative estimate was available (excluding six "comparative" terms), 32 expressions ranging from "very very minimal" to "very high" were used, a richness of vocabulary comparable to that observed by Budescu et al.[9] These phrases represented a probability range of about 2×10^{-8} to 0.9. In some cases, a witness used a single verbal expression to represent a numeric range of probabilities (e.g., "one to three out of one thousand")[10] and, in others, several expressions, such as "rare" and "extremely rare," were used in reference to a single estimate of the numeric probability.

Each case was read by the attorney-investigator (JM). A subset of the 110 cases in which verbal expressions were recited (a total of 11 cases from California, Massachusetts, and New York) was reviewed independently by the physician-investigator (DM) to provide a reliability check on the analysis. This review was based on prior agreement between the investigators that only expressions relating to the probability or frequency of *risk occurrence* were to be included. There was 100% agreement in the coding of the verbal expression usage in this sample of cases.

Results

Figure 1 presents the data from the 55 cases. The horizontal axis shows the represented probability on a logarithmic scale. Each verbal expression takes one level on the vertical axis. For each expression used, we have plotted a point on this scale. The numeric values—as provided in testimony—represent either 1) a single value (42 cases) or 2) the upper and lower bounds of a numeric range (13 cases). We treated the latter situation as two observations, since the verbal expression applied equally to both values (and all in-between). The lines joining the points are indicative of the variability associated with each word, and do not reflect ranges of probabilities given in testimony.

Figure 1 shows large variation in the numeric values applicable to some verbal terms. Interestingly, we see from this figure that the numeric variability of the terms increases as the probabilities represented thereby decrease. Because we did not elicit values from subjects, we had no control over the number of observations within any single verbal term, which turned out to be small. Therefore, we decided not to order the terms by the average represented probability, instead relying upon English-language usage (e.g., uncommon < common < quite common) and the empirically derived orderings of the same terms presented by other investigators.[1,5,6,9,11–14] Based on this ordering, we saw a definite pattern separating what we have somewhat arbitrarily identified as "high" words and "low" words—

the variability appears smaller and the probabilities higher for the "high" words. This informal grouping of terms is discussed further in the following section.

CONSISTENCY OF INTERPRETATIONS

We wanted to test whether the verbal expressions were distinguishable based on the represented probabilities. Unfortunately, due to the small number of observations for any single word, we could not quantitatively distinguish between the values represented by different terms efficiently and in a statistically meaningful way. Therefore, we chose to group the terms in order to allow us to perform a quantitative assessment (one-way ANOVA) of whether they were applied consistently and whether they could be distinguished based on their represented numeric values.

To facilitate the following analysis and discussion, an ordered scale of verbal expressions is loosely defined. We choose as a central term "possible." From a Bayesian perspective, "possible" may be information-less (depending upon context), inasmuch as no insight into the probability of occurrence other than that a particular outcome is known to be associated with a diagnostic or therapeutic procedure is gained from the phrase itself. With this starting point, as shown in figure 1, "low" is less likely than "possible," and "high" is not as rare as "possible." Further, the adjectives "very" and "extremely" are applied to modify these
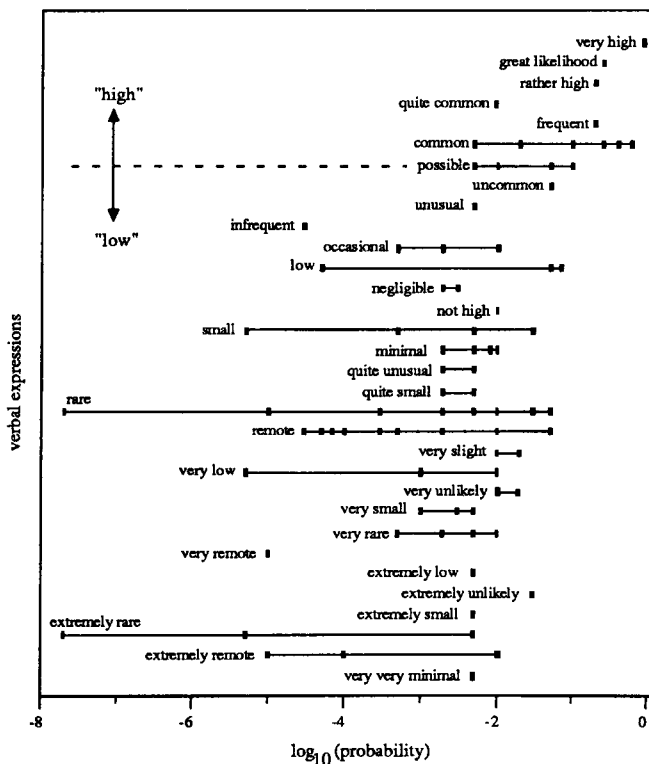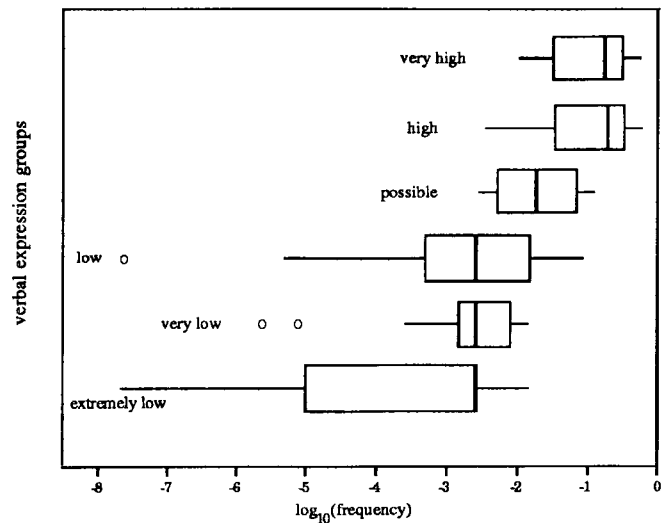
FIGURE 2. Boxplots by verbal expression groups.

terms, where "very" is not as extreme as "extremely." Thus, our total scale of verbal expressions is: "extremely low"—"very low"—"low"—"possible"—"high"—"very high"—"extremely high."

Despite the broad collection of terms found in the cases, it was hypothesized that this scale could be adopted to explore whether the modifiers "very" and "extremely" indeed were applied in the assumed order. To accomplish this analysis, the terms were grouped by modifier. For example, "low" includes "rare," "remote," "small," "not high," "negligible," "occasional," "infrequent," "unusual," and "uncommon." The modifier "very" was assumed to subsume "quite," "rather," and "great." "Extremely" includes "very very."

Boxplots of the grouped data are presented in figure 2. To facilitate our analysis, we transformed the data by taking the 4th root of the negative of the log of the probability to stabilize the variance. This transformation has been used consistently throughout the following analyses. This extreme transformation was necessitated by the highly skewed nature of the data. The goal was to make the residuals from our ANOVA and regression analyses appear normally distributed on residual plots. No formal test of normality was performed.[15]

Performing a one-way analysis of variance, we found that the mean of at least one of the six groups presented differed from the others ($F^* = 12.67$ with 5, 85 df, $p = 0+$). Analysis of all pairwise contrasts indicated that each of the "low" terms was indeed different from both of the "high" terms at a 99% family (Tukey or Bonferroni) confidence level. Neither the "low" terms nor the "high" terms were significantly different from "possible," perhaps due to the small number ($n = 4$) of "possible" observations. Last, there was no significant difference in the probabilities represented by "extremely low," "very low," and "low," and likewise by "high" and "very high." Thus, as was informally rec-

FIGURE 1. Verbal expressions and numerical equivalents.

Table 1 • Severity Scale

Value	Physical State
1	Infection, pain
2	Wrongful conception
3	Scars, tissue damage
4	Damage to bodily part requiring further treatment
5	Partial or temporary paresthesia, loss, or loss of use of bodily part with no residual impairment
6	Paresthesia, loss, or loss of use of bodily organ with residual impairment
7	General paresthesia, loss of or loss of function of limb, eye, ear; minor brain damage
8	Paraplegia, hemiplegia, moderate brain damage, stroke, loss of two limbs, blindness, deafness, birth defects
9	Quadriplegia, severe brain damage
10	Death

ognized from assessing figure 1, we can conclude only that "high" words indicate higher probabilities than "low" words.

"Low" terms were found to have a corresponding mean probability of 0.009, and a range of 2×10^{-8} to 0.07. "High" terms had a mean of 0.3 with a range of 0.005 to 0.9. While "low" terms thus may be distinguished from "high" terms based on the foregoing analysis, the large variation in these terms provides further evidence that communication of probability information by use of verbal expressions is at best highly imprecise, and at worst misleading.

INCLUSION OF UTILITY IN VERBAL PHRASES

It is possible that verbal phrases are influenced by the situation, or context, in which the terms are used.[16] We hypothesized that the speaker's subjective assessment of the severity of the probabilistic outcome may be reflected in the choice of an expression. If this is occurring, we might expect to find that the probability expressed by a "low chance of infection" is substantially higher than the probability expressed by a "low chance of death." We wanted to explore whether the context in which we should be assessing the use of phrases should be not "medical risks" generally, but, instead, particular consequences. To quantitatively assess the consequences associated with each verbal expression used by the testifying physician, we utilized a crude severity scale (SS) that was generated and applied to the present data as part of a larger data set of informed consent cases (n = 227) comprehensively analyzed in another study.[17] This ordinal ranking of physical consequences ranges from 1 (infection and pain) to 10 (death). The scale is presented in table 1.

The SS is not a utility scale, but only a subjectively assembled ordinal ranking of consequences. The scale was applied to the large data set mentioned above by one investigator (JM) and by an independent lay rater, with inter-rater reliability of 0.85.[17] This scale, which

was not empirically derived from elicitation of subjects, is extremely similar to a "severity scale" developed in the 1970s for analysis of medical malpractice issues.[18] The two quantitative measures applied to the large data set were highly correlated (Spearman rank correlation ρ = 0.938, p = 0+).[17]

Boxplots of the probability data for each SS value are presented in figure 3. For this analysis, we used multiple linear regression, treating the verbal expression groups with indicator variables (ANOVA) and the SS as a continuous predictor variable. While not evenly spaced, we do believe that the SS measure is a monotonically increasing function with severity of consequences, which provides some justification for this treatment. Using best-subsets regression, we identified a model (F^* = 18.03 with 6, 84 df, p = 0+) having significant parameter estimates for the verbal expression groups "very low" (t^* = 6.41, p = 0+) and "low" (t^* = 3.17, p = 0.002) and for the interaction terms between the SS and verbal expression groups "extremely low" (t^* = 7.24, p = 0+), "low" (t^* = 3.85, p = 0+), "possible" (t^* = 3.51, p = 0.001), and "high" (t^* = 2.17, p = 0.033). This regression model explains a substantial amount of the variance in the transformed probabilities (R^2_{adj} = 0.53), indicative of a fair fit with the data. Interestingly, the parameter estimate for SS is not significant when the interaction terms are included in the model. There is no significant relationship between our verbal expression groups and the severity scale in our data set (Spearman rank correlation ρ = -0.161, p = 0.067). However, the negative sign of the coefficient indicates that more severe consequences are weakly associated with lower-probability verbal expressions. We concluded that, as is evident in figure 3, more severe outcomes of medical treatment are associated with lower probabilities, and there does indeed appear to be an effect of consequences on verbal expressions.

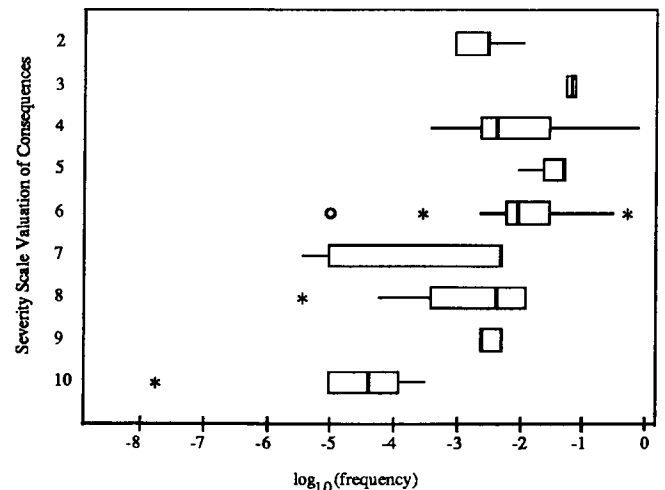To investigate this issue further, we chose the "low"

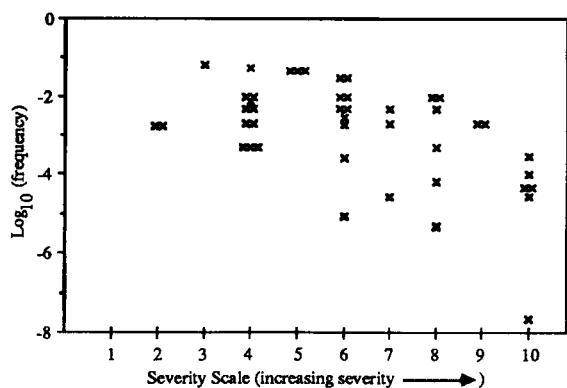FIGURE 3. Boxplots by severity scale groups.

FIGURE 4. "Low" expressions—probability vs severity scale values.

verbal expression group, which has the largest number of observations of these groups (n = 44). Figure 4 presents a plot of the probability versus the SS measures of the consequences for this subset of the data. Linear regression indicates that there is a significant relationship between probability and the SS (F^* = 16.0 with 1, 42 df, p = 0+). Thus, for the "low" grouped expressions, the probability represented by these terms will decrease as the severity of the consequences increases.
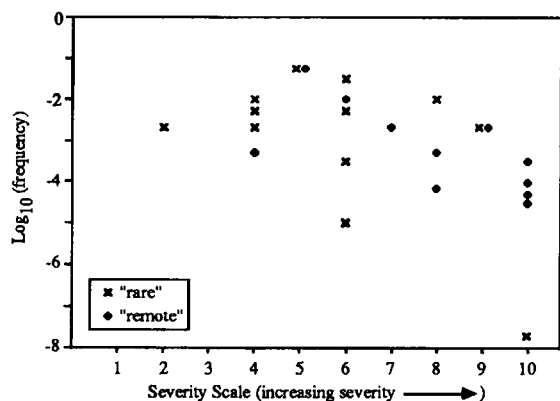

FIGURE 5. Remote" and "rare"—probability vs severity scale values.

Finally, we chose the two most frequently used expressions, "rare" (n = 12) and "remote" (n = 11), to see whether the same effects were evident in the usage of these single words. As can be seen in figure 5, an influence appears obvious. Regression for each term yields no significant relationship for "rare" (F^* = 2.50 with 1, 10 df, p = 0.145), but significant associations were found for "remote" (F^* = 7.34 with 1, 9 df, p = 0.024) and for both "rare" and "remote" combined (F^* = 8.68 with 1, 21 df, p = 0.008). Thus, the choice of verbal expression of probability appears to reflect the severity of the consequences associated with the risks.

Some of the variance in our observed numeric data is thus explained by the consequences associated with

the risks characterized by a verbal expression. *Verbal expressions of probability* may thus be a misnomer, inasmuch as it appears that more complex characteristics of risks are being defined by these terms.

COURT OUTCOMES

The obligation to disclose information regarding risks and alternatives does not extend to "full" disclosure of all known risks and all possible alternatives. Indeed, many courts have opined that the duty to disclose is a function of the risks of and alternatives to medical treatment; that is, the larger the risk, the greater the physician's duty to disclose. Inasmuch as probability is one component of risk (the other being consequences), it may be assumed that as the probability of a risk increases, the duty to disclose the risk likewise increases. Since verbal probability information has been presented in the cases here, and the results above indicate that (only) two verbal phrase groups are distinguishable by the numeric probabilities represented by the words, it is possible to test whether the duty to disclose a risk, represented by the proxy variable of court outcome, is indeed a function of the probability of the risk.

Looking back at the 95 cases in which verbal expressions were used, we find that, for cases where the risk was characterized by "low" probability words, plaintiff won (or simply received a favorable appellate decision in) 19 of 81 cases (23%). Where "high" terms were used, plaintiff won five of 14 cases (36%). While plaintiff won a slightly higher percentage of the latter cases, it is a nonsignificant difference (χ^2 = 0.95 with 1 df, p $\gg$ 0.1). The large variations in fact situations, applications of the law by the courts, subjective characteristics of the parties, differences in competence of legal counsel, and the fact that the data set consists mainly of appellate opinions rather than jury determinations obfuscate the importance of these results. Nonetheless, this evidence does not support the tenet that the duty to disclose a risk increases with increasing probability of the risk. Indeed, a more comprehensive analysis of the larger data set discussed in the previous section (which includes the cases analyzed here), using numeric probabilities where available and monetary values and the severity scale as measures of consequences, likewise failed to find a relationship between decisions favoring plaintiffs and the probabilities of the undisclosed risks. That study, however, did find that the duty to disclose a risk increased with common measures of the consequences.[17]

Discussion

The wide intersubject variability of numeric interpretations of probabilistic phrases has been attributed to contextual effects[3,4,16] and subjective differences, such

as the experience or technical knowledge of the individuals.[5-7] Decision analysts working with physicians report the difficulty some physicians find in expressing their subjective estimates of probability as an odds ratio or decimal figure.[1,9] The hypothesized reasons why physicians prefer using verbal expressions over quantitative expressions include: 1) the fact that it is more natural to think and talk about probabilities as words; 2) the keen desire, especially in the medicolegal context, of avoiding the appearance of unwarranted precision in the estimates given to patients; 3) the fact that, in many types of medical interventions and with physicians having varying degrees of expertise in performing the procedures, there may be a lack of detailed knowledge of the numeric values of the particular physician in the particular patient's case[1,9]; and 4) a simple lack of physician knowledge about or interest in securing numbers from the medical literature.

Medical practitioners have been urged by decision analysts and statisticians to quantify probabilities whenever possible,[19-22] or at least to use words and numeric estimates together,[23] to avoid the possibility of confusion or miscommunication. There is also pressure within the medical community for institutions to audit physician performance quantitatively to enable informed decisions to be made regarding the clinical efficacy of medical care.[24] In addition, patients have shown a desire for numeric estimates,[25] consistent with the findings that people prefer knowledge of the probabilities of outcomes when choosing between lotteries involving losses.[9]

The law of medical informed consent requires physicians to communicate sufficient information, including the known risks inherent in the proposed and alternative diagnostic or therapeutic procedures (and refusal of the procedure), to enable the patient to make a knowledgeable decision.[26] In the terms of decision theory, such a decision is a choice among options, each having several potential outcomes. These outcomes have particular degrees of valuation and likelihoods of occurrence. A fully "rational" decision-maker would choose the option that maximizes his or her subjective expected utility.[27] People cannot do this if they do not understand the risks. If a physician relies upon verbal expressions of probability, then patients may lack that understanding. They may be confused or may select one of the many possible interpretations of the probability phrases giving the illusion of mutual understanding.[25] In such cases, the decision process of the physician and patient would be compromised.[28] Subsequently, if a collateral risk of treatment transpired, the unhappy patient might claim that, although a risk was mentioned, disclosure was inadequate or misleading because the probability was characterized by a verbal expression having a wide interpretation.[26]

One case has been found where the plaintiff raised this issue, but was prevented from trying the case because he failed to offer expert testimony to prove physical causation of the subsequent injury. In Maliszewski v. Rendon, plaintiff experienced foot drop from damage to the L4–5 disc. Defendant excised the disc, which improved the condition temporarily, but foot drop recurred. Plaintiff contended that he had been told, "contrary to the truth," that there was a "high likelihood" of success, while in fact defendant "perceived the probability of regaining use of the foot as remote."[29]

Other cases indicate that the way in which disclosure takes place may receive more legal attention in the future. For example, an appellate court in Florida empowered the jury (on retrial) to determine whether the method of communicating risk information to the patient was adequate to enable the patient to make an informed consent.[30] This case has been subsequently interpreted on the narrow facts there, where the patient's inability to understand English hindered physician–patient communication. In a New York case, plaintiff's counsel convinced the trial judge that simple disclosure of a risk may be inadequate to allow the patient to make a knowledgeable choice, and the judge held that, in some instances, the physician must disclose the "mechanics" of the risks of treatment.[31] These cases indicate that judges may let juries decide whether the *form* and *content* of disclosure are adequate to allow the patient to make an informed choice.[26] Astute advocacy could extend the law to situations where the consent was based on ambiguous information.

If an informed consent case goes to trial, a jury may be presented with verbal expressions of the probability of a risk that has transpired in that case. The jury may be free to subjectively assess the magnitude of that probability as part of their determinations of the "materiality" of the risk (in states where this standard is applicable) and whether the reasonable patient would have agreed to the recommended treatment had the risk been disclosed.[32] The jury's decision undoubtedly would be aided, and perhaps would be more predictable, if based on a quantitative estimate of the probability. Unfortunately, simply providing a numeric figure may not yield a consistent or proper subjective valuation of a probability, and the jury's task might be aided further by an exhibit at trial, such as a visual representation of the probability of the risk. Kaplan et al. presented student subjects, who were faced with the hypothetical choice of receiving an influenza vaccination, with visual representations comprised of small dots printed on paper. The risks of death from both the flu and the vaccination were presented numerically to control group subjects and with the visual aid to the treatment groups. The presentation of the visual aid influenced the decisions made by the students: those subjects who received the graphic presentation were more likely to say that they would accept the

vaccine.[33]

The numeric interpretations of verbal phrases appear to vary greatly among physicians[1] and among patients.[34] These two sources of uncertainty may deplete of meaningful content communications via verbal expressions. The difficulties in achieving a solution to the questions we have raised are great. Physicians may find it difficult to specify numeric probabilities or to shift from verbal to numeric probability frameworks for clinical decision making and patient communication. Even in cases where physicians can specify probabilities, quantification in physician–patient communications may lead to patient perceptions that more accuracy exists than is realistically attainable within the clinical setting. If so, ranges of probabilities should be used to communicate applicable uncertainty.

It is reasonable to assume that some patients are numerically literate and will prefer, but perhaps not be able to request (and generally are not required to under the law),[26,35] numerical probabilistic information. Physicians may minimize their financial exposure from the informed consent laws by altering their standard disclosure practices to meet the needs of this most sophisticated subset of patients. Patients and physicians may both benefit from the reduction in ambiguity of their discussions: physicians will by necessity need to know the literature and their own experience with medical procedures, which may lead to better clinical decision making; patients may be more empowered to make choices of therapy and may not be as surprised if risks of treatment actually transpire.

CONTEXTUAL EFFECTS

One possible limitation to our analysis is that the context in which we have assessed the usage of the verbal expressions is not narrow enough. An important dimension of the context that we were able to assess in this study was the severity of the risk to which each probability estimate applied. Some researchers have found that the perceived utilities for outcomes influence subjective probability estimation.[36] Likewise, we hypothesized that the choice of a verbal expression may be influenced by the subjective valuation or utility for the consequences associated with the risk. In support of this hypothesis, we found that the magnitude of the particular consequences associated with a risk as reflected in a crude severity scale appears to influence the use of verbal expressions of probability. Thus, the variation of numeric probabilities associated with verbal expressions may be explained in part by our simple measure of consequence severity. Whether the severity scale is a valid index (or continuous measure) for health outcomes and whether alternative consequence measures would not yield similar results require empirical validation.

Further, we cannot be sure that the effects of other contexts, such as the nature of the procedure or the severity of the patient's presenting condition, are not implicit in our data. Moreover, drawing on trial testimony as we have done here, we cannot be sure that the disclosure of multiple risks—which could provide a relative scaling between expressions—in the actual informed consent setting would not decrease the ambiguity of these terms. We assert, however, that utilities and any other limiting contexts, unless clearly understood by both the disclosing and the receiving parties, cannot aid the communication process. Inasmuch as the disclosing and receiving parties may not recognize the implicit conditions on interpretation of the phrases (and if they do, they may place different values on those conditions), placing multiple duty on probability expressions can arguably increase the ambiguity of those words in communication. Whether such implicit contexts are understood between patients and physicians is an empirical question worthy of further research. Until we know more about how patients comprehend this type of information, such use of probability expressions should be avoided.

Such admonition alone does not solve the problem of ambiguity in verbal expressions. Suggesting quantification wherever possible also does not solve the problem, for it will be resisted for the (valid) reasons set forth above. Therefore, a possible solution lies in continuing to use verbal expressions, but providing a framework in which some consistency of terminology and range of applicable numeric probabilities can be attained. Therefore, we propose here a syntax of verbal probability expressions.

PROPOSED SYNTAX OF VERBAL EXPRESSIONS

Drawing a loose analogy to the electromagnetic (EM) spectrum,* a six-expression scale can be defined within the framework utilized in the analysis presented in this paper. The syntax is, in decreasing probability, "extremely high" (EH), "very high" (VH), "high" (H), "low" (L), "very low" (VL), and "extremely low" (EL). The major difference from common verbal characterizations of the EM spectrum is the substitution of the modifier "extremely," which has been observed in use by physicians, for "ultra," which has not. The suggested probability ranges to which these terms apply are presented in figure 6. Notice that, in this syntax, there is no "possible" or similar term. Such terms, although used by people, convey only a minimum amount of probabilistic information [i.e., that the probability does—or may—not equal zero, or, absent *any* diagnostic information, the Bayesian's prior is ~uniform(0, 1) with a mean of 0.5, and any alternative interpretation of the term may be seen to depend upon what is known about the probabilistic event and the
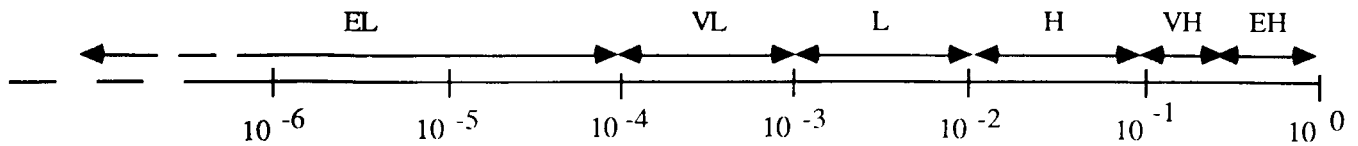

FIGURE 6. Suggested syntax of verbal probability expressions. EL = extremely low; VL = very low; L = low; H = high; VH = very high; EH = extremely high.

context in which the term is used], and their use should be avoided.

The probabilities represented by the terms in our syntax are loosely analogous to the ranges observed in the informed consent cases. The proposed scale mimics the increasing variation as probabilities decrease seen in the present data, probably also reflecting actual uncertainties in the quantitative estimates available for medical risks. The primary difference between the observed data and our syntax is the smaller and mutually exclusive ranges applied to the terms. These terms can, of course, be used to communicate ranges or uncertainty in probability estimates by using two terms (e.g., "between very low and low").

Similar to what we observed in our data, probabilities above 1 percent are represented in our syntax by "high" words. One might object that the use of the word "high" to represent a 1 to 10 percent probability of some outcome in unreasonable, and may have the undesired effect of scaring patients into rejecting medically beneficial treatments. This objection appears valid if a 6 percent probability of infection[37] is being disclosed, since a patient may not initially want to accept a treatment having a "high" chance of infection. But a 6 percent chance of death or major disability may seem naturally "high." This might of course depend further on what the alternative is—a 6 percent chance of death associated with certain surgery may not seem "high" if the short-term chance of death from no treatment is, say, 0.5 (which is "extremely high" in our syntax). Such comparative contexts also place excess baggage on expressions of probability. The purpose of the suggested syntax is to do away with implicit context effects: a 6 percent chance of infection may well be "high" to a patient, and it is inappropriate and confusing to weight a probability expression by one's valuation of the particular outcome. At the very least, disclosure of a "high" chance of an undesirable outcome may stimulate further, and possibly quantitative, discussion between physician and patient. To these authors, this result in itself would be highly beneficial to patient–physician relationships.

The suggested scale has several potential advantages: 1) the imprecision associated with a verbal estimate is retained, while the variation of the associated probability is decreased; 2) the number of words used is reduced to a manageable few (that already find usage in the medical setting); 3) the uniformity of the language of probability is promoted; and 4) a framework is provided in which physicians *and patients* can be

calibrated to increase the likelihood of effective communications.[22] The disadvantages include: 1) the difficulty in institutionalizing the calibration of physicians and patients; and 2) the possibility that users may become sloppy in using the expressions, without properly attending to their underlying probabilities.[5] The first problem may be addressed by the medical community in its schools, continuing education programs, and hospital administration. Patients, often offered a deluge of pamphlets in their physicians' waiting rooms, could find one describing the language of medical risks. The second problem can degrade the effectiveness of the scale, but the net result may be better than the present situation.

In forming a standardized scale for a professional forecasting organization in Israel, Beyth-Marom concluded: "A verbal scale of probability expressions is a compromise between people's resistance to the use of numbers and the necessity to have a common numeric scale. There is no doubt that this is the second best solution and should be implemented only after giving up hope to convince the organization to use numeric expressions explicitly." [5, p. 268]

POTENTIAL BIASES IN RESEARCH OF CASE LAW

There are significant limitations on how far we may extrapolate our findings from this study. The data was derived from published reports of trial and appellate court opinions, which may not represent an unbiased sample of the cases actually litigated. There may indeed be a bias in litigation toward higher dollar value injuries, or cases which present novel problems for judicial resolution.[17] Further, expert witnesses, who are the source of our data, may be biased, inasmuch as all but the defendant are hired to provide testimony to support one party's position. Defendant and defendant's witnesses may (perhaps unconsciously) downplay risks, and plaintiff's experts may overemphasize risks, which undoubtedly could affect the context in which the choice of a verbal expression was made. For example, defendant may testify that a 0.01 chance of paraplegia is "low," while plaintiff will try to establish that such a chance is "high." Unfortunately, we did not have enough information to assess this factor.

Further, we made no effort to verify that quantitative estimates provided in the cases were accurate. We generally did not have enough information to assess whether testimony reflected the physician's personal

experience, experience in the relevant community of physicians or institution in which the testifying physician practices, or the literature available at the time of treatment of the patient (or of trial, which should be irrelevant to the adequacy of disclosure at the time of treatment). The philosophy behind our adversarial system of trials is that each party is motivated to establish the truth, and if one party provided significantly biased estimates of risk probabilities, it is incumbent upon the other to rectify that assertion of fact. Therefore, we believe that independent verification of numeric probabilities for this analysis was neither possible nor necessary.

Conclusion

The medical informed consent cases reviewed here provide further evidence that verbal expressions of probability or frequency are assigned broad numeric interpretations within the medical community. Despite a rich lexicon of verbal expressions used in the court cases, only two groups, representing "high" and "low" probabilities, could be distinguished. The choice of verbal expressions appears to be influenced by the severity of the consequences associated with the risks, and whether this aids or detracts from the communication process requires further research. Nonetheless, based on these findings, it may be inferred that the ambiguity of these expressions will interfere with the informed consent communication and decision-making process. As a possible aid to the communication process, a syntax of verbal probability expressions, if adopted by medical practitioners, may be valuable.

References

1. Kong A, Barnett GO, Mosteller, F, Youtz C. How medical professionals evaluate expressions of probability. N Engl J Med. 1986;315:740–4.
2. Wallsten TS, Forsyth BH, Budescu DV. Stability and coherence of health experts' upper and lower subjective probabilities about dose–response functions. Organ Behav Hum Perform. 1983;31: 277–302.
3. Mapes REA. Verbal and numerical estimates of probability in therapeutic contexts. Soc Sci Med. 1979;13A:277–82.
4. Parducci A. Often is often. Am Psychol. 1968;23:828.
5. Beyth-Marom R. How probable is probable? A numerical translation of verbal probability expressions. J Forecasting. 1982;1: 257–69.
6. Budescu DV, Wallsten TS. Consistency in interpretation of probabilistic phrases. Organ Behav Hum Decis Proc. 1985;36:391–405.
7. Wallsten TS, Budescu DV, Rapoport A, Zwick R, Forsyth B. Measuring the vague meanings of probability terms. J Exper Psychol Gen. 1986;115:348–65.
8. Contreras v. St. Luke's Hosp., 78 Cal. App. 3d 919, 923, 144 Cal. Rptr. 647, 650 (1978).
9. Budescu DV, Weinberg S, Wallsten TS. Decisions based on numerically and verbally expressed uncertainties. J. Exp Psychol Hum Percept Perform. 1988;14:281–94.
10. Hartke v. McKelway, 707 F.2d 1544, 1547 (D.C. Cir. 1983).
11. Stone DR, Johnson RJ. A study of words indicating frequency. J Educ Psychol. 1959;50:224–7.
12. Lichtenstein S, Neuman JR. Empirical scaling of common verbal phrases associated with numerical probabilities. Psychonomic Sci. 1967;9:563–4.
13. Simpson RH. The specific meaning of certain terms indicating differing degrees of frequency. Q J Speech. 1944;30:328–30.
14. Simpson RH. Stability in meanings for quantitative terms: a comparison over 20 years. Q J Speech. 1963;49:146–51.
15. Neter J, Wasserman W, Kutner MH. Applied linear statistical models. 2nd ed. Homewood, IL, Irwin, 1985.
16. Cohen J, Dearnley EJ, Hansel CEM. A quantitative study of meaning. Br J Educ Psychol. 1958;28:141–8.
17. Merz JF. An empirical analysis of the medical informed consent doctrine: the search for a "standard" of disclosure. Risk: Issues in Health & Safety 1991;2: 27–76.
18. Daniels S, Andrews LB. The shadow of the law: jury decisions in obstetrics and gynecology cases. In: Rostow VP, Bulger RJ, eds. Medical professional liability and the delivery of obstetrical care. Volume II: An interdisciplinary review. Washington, DC: National Academy Press, 1989;161–93.
19. Eddy DM. The challenge. JAMA. 1990;263:287–90.
20. Hakel MD. How often is often? Am Psychol. 1968;23:533–4.
21. 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.
22. Fryback DG. Decision maker, quantify thyself! Med Decis Making. 1985;5:51–60.
23. Kenney RM. Between never and always. N Engl J Med. 1981;305:1097–8.
24. Beebe HG, Clagett GP, DeWeese JA, et al. Assessing risk associated with carotid endarterectomy: a statement for health professionals by an ad hoc committee on carotid surgery standards of the Stroke Council, American Heart Association. Stroke. 1989;20:314–5.
25. Brun W, Teigen KH. Verbal probabilities: ambiguous, context-dependent, or both? Organ Behav Hum Decis Proc. 1988;41:390–404.
26. Merz JF, Fischhoff B. Informed consent does not mean rational consent: cognitive limitations on decision-making. J Legal Med. 1990;11:321–50.
27. von Winterfeldt D, Edwards W. Decision analysis and behavioral research. New York: Cambridge University Press, 1986.
28. Mazur DJ. Informed consent: court viewpoints and medical decision making. Med Decis Making. 1986;6:224–30.
29. 374 Pa. Super. 109, 542 A.2d 170 (1988).
30. Dandashi v. Fine, 397 So. 2d 442 (Fla. App. 1981).
31. Nisenholtz v. Mt. Sinai Hosp., 126 N.Y. Misc. 2d 658, 483 N.Y.S.2d 568 (1984).
32. Kissinger v. Lofgren, 836 F.2d 678 (1st Cir. 1988) (applying Massachusetts law).
33. Kaplan RM, Hammel B, Schimmel LE. Patient information processing and the decision to accept treatment. J Soc Behav Pers. 1985;1:113–20.
34. Mazur DJ, Hickam DH. Patients' interpretations of probability terms. J Gen Intern Med. 1991;6:237–40.
35. Canterbury v. Spence, 464 F.2d 772, 783 (D.C. Cir. 1972).
36. Zakay D. The relationship between the probability assessor and the outcomes of an event as a determiner of subjective probability. Acta Psychol. 1983;53:271–80.
37. Larson E, Oram LF, Hedrick E. Nosocomial infection rates as an indicator of quality. Med Care. 1988;26:676–84.