RESEARCH ARTICLE JSLS

Words Representing Numeric Probabilities in Medical Writing Are Ambiguous and Misinterpreted

Douglas E. Ott, MD, MBA

ABSTRACT

Background and objective: Words of estimative probability are used in medical writing. Authors know the number they intend the word to mean, readers do not. The objective of this study is to assess the variability of words having numeric meaning to medical doctors.

Design: A survey of 131 American trained MD's and DO's was done regarding their interpretation of 27 commonly used words of the estimative probability they placed on each word.

Methods: Statistical assessments were done to evaluate specific word meaning as a number or range to physicians and compared to each other's interpretation.

Results: For 19 of the 27 words had a 30% +/− (60% variance) of interpretation. Twenty-five of the 27 had less than 38% agreement on numeric meaning. Only two words had more than 74% numeric agreement.

Conclusion: Words of estimative probability have widely varied interpretations to physicians. This makes interpretation of those words incomprehensible for scientific meaning resulting in communication without comprehension.

Key Words: Probability, Communication, Bias, Ambiguous, Estimate.

School of Engineering, Mercer University, Macon, GA 31210, USA.

Acknowledgements: The author thanks and acknowledges Christopher Brydges, Adila Esaak and Alice Tansill for graphic design and Lynda Davis for editing.

Disclosure: none.

Funding/Financial support: none.

Conflicts of interest: The author declares no conflict of interest.

Address correspondence to: Douglas E. Ott, MD, MBA, 109 Preston Court, Macon, GA 31210. Telephone: 1-478-477-8996, E-mail: gabiomed@mindspring.com

DOI: 10.4293/JSLS.2021.00034

© 2021 by SLS, Society of Laparoscopic & Robotic Surgeons. Published by the Society of Laparoscopic & Robotic Surgeons.

INTRODUCTION

Humpty Dumpty smiled contemptuously. . . . "When I use a word," Humpty Dumpty said, in rather a scornful tone, "it means just what I choose it to mean—neither more nor less." "The question is," said Alice, "whether you can make words mean so many different things."[1]

Use and interpretation of words conveying linguistic probabilities in science writing are inconsistent and ambiguous. Author(s) have knowledge and intent knowing what they mean with the word(s) they chose. Readers interpret the word(s) influenced by context, biases, culture, experience, and prior knowledge. Variability in understanding and application of words of estimate probability (WOEP) vary from person to person. To have clear and unambiguous message transmission medical authors must declare the numeric value they place on WOEP used. Readers do not know what number or range a writer means when are WOEP is used.[2] Each person has a number or fuzzy range of probability for what a word represents as a number. Without author(s) disclosure of a word's numeric meaning the result is loss of clarity, perpetuation of imprecise information, and a false understanding of known facts. This gap creates unclear and misunderstood statements. Science and medical research depend on a number driven interpretations to recognize and critically weigh alternatives and risks, compare outcomes, and prescribe treatments or medications. Medical knowledge is based on findings that over time and with replication have a known specific or range value. Data are assessed by statistical inferences and expressed as numbers not words.

Words implying numeric meaning should be expressed as a frequency, proportion, probability or percent (%) if it is known for the WOEP proposition they are stating. When the number or range is known for the WOEP used the number should be specifically disclosed at the beginning of the article so readers know what the author(s) means probabilistically each time that word is used. The value should appear in the text after the word in a bracket, i.e. X is very unlikely (1 per 1,000) or Y occurs (5 – 12%) of the time. This increases clarity and precision avoiding misinterpretation and miscommunication. Authors are given little guidance regarding use of WOEP from journals or editors. Outlets of medical information have an obligation

for authors to include numeric equivalence of probability expressions. This reduces uncertainty, vagueness, and ambiguity and enhances clarity and interpretation. An explicit declaration of accepted fact numbers should not be hidden by vague words. Without this simple but important measure, assumed knowledge that is or may be wrong will be perpetuated and assumed correct.

Science can be subtle and nuanced, but it is also measured and under constraints predictable and expressed by numbers. WOEP used in scientific writing create situations of numeric intent known to the author(s) and guessed at by the reader. Revealing the specific or range value of these words the author(s) intends improves understanding of the message. The precision in science, its measurements, outcomes, and self-correcting methodological approach demands specifics not generalities or suppositions. There are suggestions for defined nomenclature of word numeric value for informing patients regarding aspects and frequencies of drug side effects and medical/surgical outcomes but none for physician authors communicating with other physicians in medical articles.[3,4] Readers depend on context, luck, intuition, experience, biases (all 235 of them), and anecdote to interpret every WOEP, hedging, weasel word, and truthiness statement not knowing the author(s) intended parameter(s).[5,6]

METHODS

The structure of this research is a blinded survey of physicians regarding their interpretation of numeric meaning for 27 WOEP used as numeric descriptors to understand the meaning medical physicians place on them. Responses were obtained from 131 respondents who are English speaking American-trained randomly chosen MD/DO physicians. The survey request is shown in **Table 1** and the survey words in **Table 2**. Descriptive statistics, calculations of mean, mode, standard deviation (S.D.), and 10th and 90th percentiles (P10 – P90) (the middle 80% as an index of variability eliminating the effects of outlier values) were calculated using known accepted formulas. Responses given as a range were analyzed as the mean of that range.

RESULTS

The mean, S.D., and P10 – P90 for the 27 words are arranged by decreasing magnitude of the mean (see **Figures 1** and **2**). Responses were as a single or range of numbers. **Table 3** shows percent disagreement for each word compared to the mode.

Single Number Interpretation

Single number answers were given 610.7% of the time versus 290.9% as a range and 80.4% as a written comment or left blank. Words with the largest S.D. as a single number were "almost" with a mean of 74 and S.D. of 27 (range of 47 – 99), "nearly" with a mean of 76 and a S.D. of 25 (range of 51 – 99), and "possible" with a mean of 33 and S.D. of 23 (range of 10 – 56). Eight of the 27 words had S. D. greater than 20 (a 40% range).

Words responded to as a single number with the most common mode (highest agreement) were "average" (94%), "never" (85%), "impossible" (75%), and "always" (72%). Words with the least consensus were "slightly" (13%),

Table 1.
Survey Request and Instructions

This survey is to identify your interpretation of these words if they were a number. Please go to the link and fill in the numeric value you interpret each of these words to be. I plan to write an article that reveals what readers interpret these words to mean if they were a number. Thank you.

Please click here to start the survey (https://weazelwordzsurvey.wufoo.com/forms/probabilistic-meaning-of-words/).

Probabilistic Meaning of Words Used in Scientific Articles

Scientific articles use words that are not precise in their numeric estimative probabilistic meaning. The aim of this survey is to establish the understanding of the numeric probabilistic value of a word used in scientific writing.

Please be precise in your assessment of the numeric probabilistic meaning for each word and assign a percentage (ex. 25%) or percentage range (ex. 15% to 25%) to each word, based on your understanding of what a specific word means as a range or precise number.

Be precise in your assessment of the numeric probabilistic word meaning.

You and any information will be and will remain anonymous.

Table 2.
Survey Words

Almost	Improbable	Possible
Some	Few	Slightly
Many	Impossible	Probable
Could	Frequent	Typical
Always	Generally	Unlikely
Barely	Most	Meaningful
Average	Nearly	Probably
Certain	Often	Never

"often," and "possible" (18%). Eighty-five percent (23/27) of the words had less than 50% agreement on its interpreted value.

Range Number Interpretations

Words with the largest S.D. when interpreted as a range were "almost" (mean of 77 and S.D. of 29) for a range of 47 – 97% and "possible" (mean of 40 and S.D. of 23) having a range of 17 – 63%. Six of the 27 words had a S.D. more than 20 (range of 40% interpretation).

Combined Interpretation

The combined findings (adding single number to average range response) had seven words with a S.D. greater than 20 (almost, nearly, possible, impossible, certain, meaningful, and uncertain). These words had the most varied interpretation. Only two words had S.D. below 10 (never and average) and were the least misunderstood. Words with the lowest P10 – P90 evaluation represent words closest to the mean interpretation. They were "never," "impossible," and "always" and were the only words with a single digit variance. Words with the largest P10 – P90 evaluation represent words with the largest difference from the mean. They were "almost," "possible," "meaningful," and "nearly." Based on mode there was only one word above 77%, "average" with 93% agreement. Nine of the 27 words had over 80% disagreement from their mode and 22/27 words had over 70% disagreement. Words with the most different individual number interpretations were "most," "nearly," and "often." Words with the fewest different discrete interpretations were "impossible" and "average."

Figure 1 shows the mean, median, S.D., P10 – P90 and mode of the word's numeric percent interpretation.

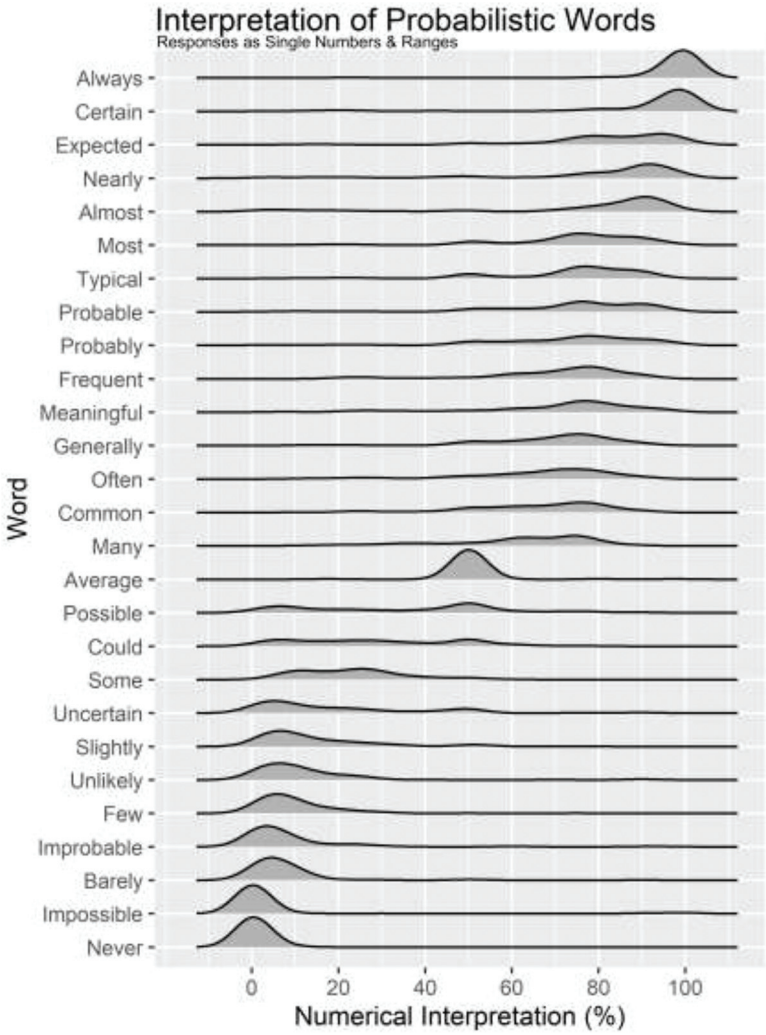
General Observations

Placing a numeric value on these words was objected to by 8% of responders saying there was no context or it was their opinion that the words had little to do with probability or frequency of occurrence. It has been shown that sentence word context modifies or influences the meaning of probability phrases into numerical equivalents by less than 2%.[7]

Seventy-eight percent of all evaluations ended in zero or five suggesting bias in numeric assessment by rounding up or down. Words with the greatest mode agreement were "average" (100%) and "typical" (96%). Words with the least mode frequency agreement and the largest S.D. are most likely (over 80%) to be misunderstood by the reader having the least agreed upon interpretation and were "probably," "nearly," "almost," and "most." **Table 3** shows percent ranking of respondents differing from the mode (the value occurring most often) of responses for each word. The three least understood words in this evaluation based on mode were "most" (85% disagreement), "often" (84% disagreement), and "uncertain" (84% disagreement). Of the 27 words 22 had disagreement from the mode over 62% of the time. Words with the largest percent agreement were "average" (93%) followed by impossible (77%) and "never" (74%). All other words had less than 38% agreement. The mode was agreed with only16% of the time meaning the word as a number was interpreted differently 84% of the time. Of the 27 words (15%) only four had over 40% agreement based on mode.

DISCUSSION

Physicians interpret WOEP as risk occurrence, frequency, or probability interpretation based on the word's positivity or negativity, context, and their understanding of the situation being discussed. This may be accurate, ill-informed, or biased. Responders to this survey translated WOEP as a point or range of personal possible interpretative meaning and had very little (7%) consensus with other physician understanding. This survey showed a wide range of understanding for 21 of the 27 (78%) words using a S.D. of 15 and a P10 – P90 difference above 20% as the criteria. Using mode concurrence as the criteria 22 of the 27 (81%) words had greater than 62% disagreement. Because the goal of medical communication is to optimize understanding written communication, using WOEP creates the opposite effect, i.e., a wide range of interpretation absent the intent of the author (s) compared to the understanding of the reader. Without definition of terms for numeric knowns, authors use of WOEP accurate communications between the author and reader is a guessing game.

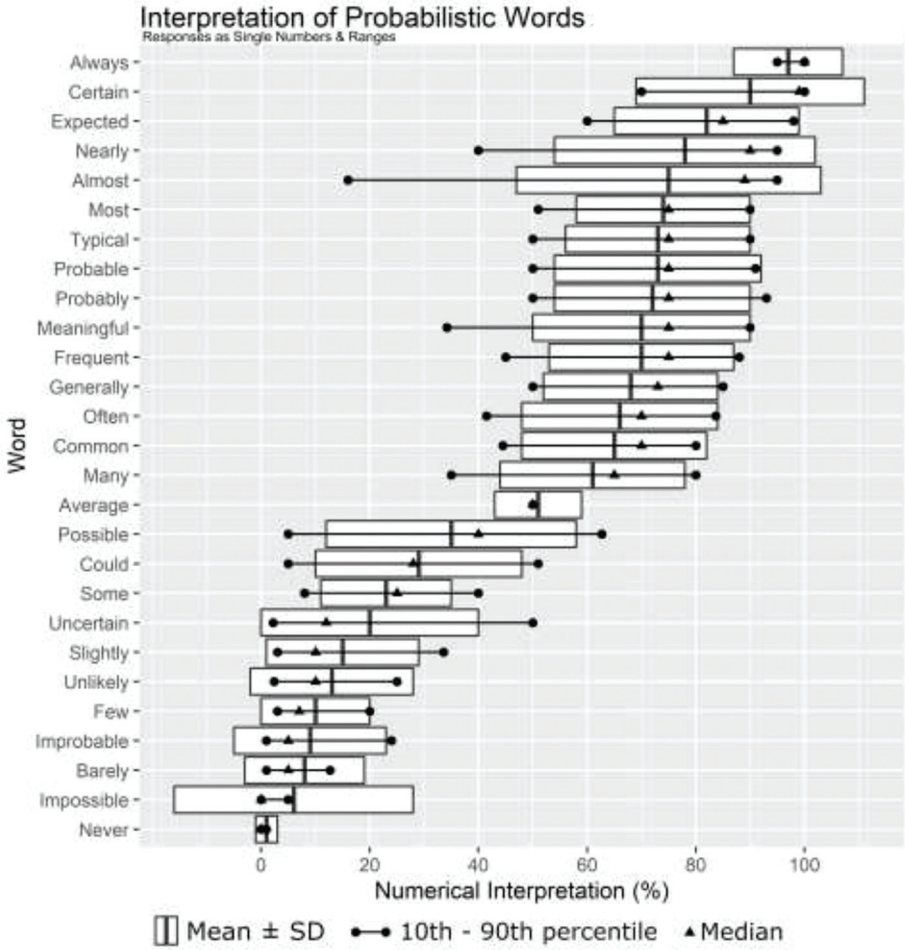


	Mean	Median	Mode	SD	10th	90th
Always	97.3	100.0	100.0	10.2	95.0	100.0
Certain	90.0	99.0	100.0	20.8	76.4	100.0
Expected	81.9	85.0	95.0	16.7	61.4	98.0
Nearly	77.6	90.0	90.0	24.4	41.5	95.0
Almost	74.9	89.0	90.0	28.1	20.0	95.0
Most	74.4	75.0	75.0	15.8	51.0	90.0
Typical	73.0	75.0	75.0	17.0	50.0	90.0
Probable	72.8	75.0	75.0	19.3	50.3	90.0
Probably	72.0	75.0	75.0	18.7	50.0	93.0
Frequent	69.8	75.0	80.0	16.7	47.5	87.5
Meaningful	69.6	75.0	75.0	19.9	37.0	90.0
Generally	68.2	73.0	75.0	16.3	50.0	85.0
Often	65.5	70.0	70.0	17.8	45.0	83.0
Common	65.1	70.0	75.0	16.6	48.5	80.0
Many	61.1	65.0	75.0	17.3	35.0	78.8
Average	50.9	50.0	50.0	8.3	50.0	50.0
Possible	35.2	40.0	50.0	23.1	5.0	60.3
Could	29.5	28.0	50.0	19.0	5.0	50.5
Some	23.3	25.0	25.0	12.3	9.0	40.0
Uncertain	20.4	12.0	5.0	20.3	3.0	50.0
Slightly	14.8	10.0	5.0	14.0	3.0	31.5
Unlikely	12.5	10.0	10.0	15.4	3.0	25.0
Few	10.1	7.0	10.0	9.8	3.0	20.0
Improbable	8.8	5.0	5.0	14.4	1.0	21.6
Barely	7.8	5.0	5.0	11.3	1.0	11.8
Impossible	5.5	0.0	0.0	21.2	0.0	5.0
Never	0.6	0.0	0.0	1.7	0.0	1.0

Figure 1. Mean, S.D., and 10th and 90th percentiles (P10 – P90) for the 27 words arranged by decreasing magnitude of the mean.

Descriptions of known incidence must be told or explained rather than presumed by using WOEP. Generalizations are reasonable when accompanied by the accepted range. Assuming readers know the specifics or range of what is being described is an assumption fraught with peril. Readers should not be led to conclusions by conceptual generalizations caused by WOEP. Since reader experiences, biases, depth of knowledge, cultural or historic perspectives, and WOEP interpretation are unknown to the author(s), to convey explicit information the least confusing language should be used accompanied by the numeric value of the WOEP; a number, percent, probability, or range. To accommodate writer style and personal mode of expression WOEP may be preferred by the author(s) or editor(s) but to achieve the highest level of understanding these words must be accompanied by the numeric value or expression the author(s) places on them. This achieves message transmission with the least amount of misunderstanding.

Words in this survey and others like them have semantic meaning, i.e., different meanings for different people. **Table 4** shows the comparison of this survey with three other studies involving physician understanding of words of risk communication with patients but not in reading medical articles for five of the same words.[8–10] This study shows there is wide variation among physician interpretation of WOEP. For instance, in this survey the combined (single and range number) mean of "possible" was 35% with a S.D. of 23% resulting in a one S.D. range of 12 – 58%. This correlates well with risk managers' understanding for these words with a mean of 43% and a mean minimum/maximum of



	Mean	Median	Mode	SD	10th	90th
Always	97.3	100.0	100.0	10.2	95.0	100.0
Certain	90.0	99.0	100.0	20.8	76.4	100.0
Expected	81.9	85.0	95.0	16.7	61.4	98.0
Nearly	77.6	90.0	90.0	24.4	41.5	95.0
Almost	74.9	89.0	90.0	28.1	20.0	95.0
Most	74.4	75.0	75.0	15.8	51.0	90.0
Typical	73.0	75.0	75.0	17.0	50.0	90.0
Probable	72.8	75.0	75.0	19.3	50.3	90.0
Probably	72.0	75.0	75.0	18.7	50.0	93.0
Meaningful	69.8	75.0	80.0	16.7	47.5	87.5
Frequent	69.6	75.0	75.0	19.9	37.0	90.0
Generally	68.2	73.0	75.0	16.3	50.0	85.0
Often	65.5	70.0	70.0	17.8	45.0	83.0
Common	65.1	70.0	75.0	16.6	48.5	80.0
Many	61.1	65.0	75.0	17.3	35.0	78.8
Average	50.9	50.0	50.0	8.3	50.0	50.0
Possible	35.2	40.0	50.0	23.1	5.0	60.3
Could	29.5	28.0	50.0	19.0	5.0	50.5
Some	23.3	25.0	25.0	12.3	9.0	40.0
Uncertain	20.4	12.0	5.0	20.3	3.0	50.0
Slightly	14.8	10.0	5.0	14.0	3.0	31.5
Unlikely	12.5	10.0	10.0	15.4	3.0	25.0
Few	10.1	7.0	10.0	9.8	3.0	20.0
Improbable	8.8	5.0	5.0	14.4	1.0	21.6
Barely	7.8	5.0	5.0	11.3	1.0	11.8
Impossible	5.5	0.0	0.0	21.2	0.0	5.0
Never	0.6	0.0	0.0	1.7	0.0	1.0

Figure 2. Graph of mean, S.D., and 10th and 90th percentiles (P10 – P90) for the 27 words arranged by decreasing magnitude of the mean.

28 – 58%.[11] Even words that should not have had ambiguous uncertainty: "always" (mean of 97), "average" (mean of 51), and "never" (mean of 1), expecting these to be 100, 50, and zero percent, respectively, have uncertain meaning interpretations although in a tight range.

Conclusions expressed by ambiguous words (WOEP) are interpreted by the reader as what they already believe.[10,11] A raw number or probability explanation is without bias unless it is false because words have implied preconceptions and ranges of meaning. Truth is more inherent and revealed in a number than by a word describing that number.

Words substituting for a number or probability require numeric disclosure to reduce ambiguity of meaning.[12, 13] This survey supports that position. As decision makers physicians need to know commonalities and disagreements of scientific studies expressed as numbers. Authors use WOEP to convey what they know; readers are left guessing the meaning of those WOEP. This study shows that there are more chances, over 90%, to misinterpret the numeric meaning of WOEP creating meaning ambiguity. Trying to impart numeric value without a number, when one is known, is an assumption. Readers need clear concise precise information.

Using WOEP is a failure to communicate, a shotgun approach to guessing what am I thinking? Without a definition of WOEP the reader inserts their own range or number creating an avoidable communication failure. Readers should not be left with a perception of what was meant but know what was meant. Style manuals suggested by journals as author guides do not address scientific communication with WOEP. These writing handbooks are not specific for scientific writing and say nothing about numeric probabilistic communications.

Table 3.
Percent Agreement/Disagreement for Each Word Compared to Mode

	Mode	% Disagreement
Most	75	85
Often	70	84
Uncertain	5	84
Probably	75	83
Common	75	83
Some	25	83
Could	50	82
Expected	95	81
Few	5	80
Probable	5	79
Meaningful	75	79
Generally	75	79
Slightly	5	79
Frequent	80	78
Improbable	5	78
Almost	90	77
Many	75	77
Possible	50	77
Unlikely	10	77
Typical	75	76
Nearly	90	75
Barely	5	73
Certain	100	62
Always	100	39
Never	0	26
Impossible	0	23
Average	50	7

Table 4.
Comparison of Five Words Assessed as a Percent from Three Studies Showing Strong Agreement with This Survey

	O'Brien	Kong	Bryant	Ott
Certain	95	99	95	99
Probable	75	70	77	75
Possible	25	2	43	30
Unlikely	13	11	20	10
Never	0	0		0

Table 5.
Pro and Cons for Numeric vs WOEP Use

	Pros	Cons
Clear	X	
Precise	X	
Unambiguous	X	
Reduced bias	X	
Easy for the writer to do	X	
Easy for a journal to do	X	
Transparency	X	
Reduces uncertainty of interpretation	X	
Ensures accurate transfer of information	X	

Readers need information to draw conclusions from data, which is a numeric function not a linguistic function. Information should mean the same thing to all persons reading it especially in medicine. The optimal method of communication when words of probabilistic meaning are used is for the author(s) to define what their WOEP means as a number, percent, or probability. This establishes standardization of information communication and enhances the value of a report.[14,15] WOEP in scientific use as interpreted by doctors in this survey are misunderstood over 90% of the time. Authors use WOEP because a number or range is not known, they don't know, they are deliberately trying to skew understanding from the accepted known, or are lazy writers. Expressing scientific quantitative findings as words rather than numbers creates a discrepancy between writer meaning and reader understanding, the two are not in sync (see **Table 3**).

When a number generated by research known by testing multiple datasets or accepted as fact using WOEP in its place leads to misunderstanding and is black box lingo shielding statements from scrutiny, this asks the reader to fill in the blank. This is charlatan linguistics; words masquerading as numbers but never revealing their true identity. The pros versus cons of using numbers rather than WOEP show irrefutable evidence for using a number rather than a word (see **Table 5**).

LIMITATIONS

A limitation of this survey was context free numeric judgement assessment of the selected words. No attempt was made to assess specialty or age of the physicians surveyed.

CONCLUSIONS

The degree of imprecision for 26 of the 27 words in this survey make their use unacceptable to communicate a numeric value when its value is known because of the broad range of misinterpretation. This creates ambiguity, imprecision, and poor communication. These data show that expressions of WOEP, as terms representing a number or frequency, are assumed differently by physicians resulting in little consensus. It should be the physician scientist that authors intend to educate and inform their audience with the least amount of misunderstanding.

Words of estimative probability create unfounded qualities and quantities creating assumption mistakes to reinforce a falsehood or establish one. This is a barrier to understanding. Therefore, it is recommended that medical writers and editors enhance communication and eliminate doubt in communication by requiring reporting the frequency of events or meaning of words when known as numbers, probability, or percent as a table under the abstract and in brackets after each word is used in the text. Without comprehension there is no understanding. Without understanding there is no knowledge.

References:

1. Westacott E. Humpty Dumpty's philosophy of language. ThoughtCo, 2021. https://www.thoughtco.com/humpty-dumpty-philosopher-of-language-2670315.

2. Kent S. "Words of estimative probability," Central Intelligence Agency, Tech. Rep. 1964 https://www.cia.gov/library/center-for-the-study-of-intelligence/csi-publications/books-and-monographs/sherman-kent-and-the-board-of-national-estimates-collected-essays/6words.html

3. Lipkus I. Numeric, verbal, and visual formats of conveying health risks: suggested best practices and future recommendations. *Med Decis Making.* 2007;27(5):696–714.

4. Ethics Committee of the American Society for Reproductive Medicine. Fertility treatment when the prognosis is very poor or futile: an Ethics Committee opinion. *Fertil. Steril.* 2019;111: 659–663.

5. Ott D. Hedging weasel words and truthiness in scientific writing. *JSLS.* 2018;22(4):e2018.00063.

6. Chavalarias D, Ioannidis J. Science mapping analysis characterizes 235 biases in biomedical research. *J Clin Epidemiol.* 2010;63(11):1205–1215.

7. Johnson E. "Numerical Encoding of Qualitative Expressions of Uncertainty," Army Research Institute for the Behavioral and Social Sciences, National Technical Information Service (December 1973), http://stinet.dtic.mil/cgibin/GetTRDoc?AD=AD780814&Location=U2&doc=GetTRDoc.pdf

8. O'Brien B. Words or numbers? The evaluation of probability expressions in general practice. *J R Coll. Gen. Pract.* 1989;39: 98–100.

9. Kong A, Barnett GO, Mosteller F, Youtz C. How medical professionals evaluate expressions of probability. *N. Engl. J. Med.* 1986;315(12):740–744.

10. Bryant G, Norman G. Expressions of probability: words and numbers. *N Engl J Med.* 1980;302(7):411.

11. Hillson D. *Describing probability: The limitations of natural language.* Project Management Institute Global Congress: 2005, EMEA Proceedings, Edinburgh, UK

12. Heuer RJ Jr., Richards J. Psychology of Intelligence Analysis (CIA Center for the Study of Intelligence 1999;153, https://www.cia.gov/library/center-for-the-study-ofintelligence/csi-publications/books-and-monographs/psychology-of-intelligenceanalysis/PsychofIntelNew.pdf

13. Windschitl P, Weber E. The interpretation of "likely" depends on the context, but "70%" is "70%"—right? The influence of associative processes on perceived certainty. *J. Exper. Psychol. Learning, Memory, Cognition.* 1999;25:1514–1533.

14. Wozny J. Optimizing communication between decision makers and intelligence analysts. Stopping "slam dunks" and avoiding "dead wrongs." Thesis, Mercyhurst College, May 2005. https://a8e.4c7.myftpupload.com/wp-content/uploads/2018/06/THESIS-JEN-WOZNY-Communicating-with-Decisionmakers-in-the-Intelligence-Community.pdf?time=1592746175

15. Silberman L, Robb C, Levin R, et al. (2005) The Commission on the Intelligence Capabilities of the United States Regarding Weapons of Mass Destruction: Report to the President of the United States, March 31, 2005, Washington, DC