Review

MDMP&P Policy & Practice

MDM Policy & Practice
2025, Vol. 10(1) 1–40

DOI: 10.1177/23814683241255337
journals.sagepub.com/home/mpp
Sage

How Point (Single-Probability) Tasks Are Affected by Probability Format, Part 2: A Making Numbers Meaningful Systematic Review

Jessica S. Ancker, Natalie C. Benda, Mohit M. Sharma, Stephen B. Johnson, Michelle Demetres, Diana Delgado, and Brian J. Zikmund-Fisher

Abstract

Background. The Making Numbers Meaningful review is intended to create guidance on the effect of data presentation format on comprehension of numbers in health. **Purpose.** This article (one of a series) covers research studying so-called "point tasks," in which a reader examines materials to obtain information about single probabilities. The current article presents evidence on the effects of data presentation format on probability perceptions and feelings, health behaviors and behavioral intentions, trust, preference, and discrimination outcomes. **Data Sources.** MEDLINE, Embase, CINAHL, the Cochrane Library, PsycINFO, ERIC, ACM Digital Library; hand search of 4 journals. **Study Selection.** Manual pairwise screening to identify experimental and quasi-experimental research that compared 2 or more formats for presenting quantitative health information to patients or other lay audiences. This article reports 466 findings of probability communication from 161 articles. **Data Extraction.** Pairwise extraction of information on stimulus (data in a data presentation format), task, and outcomes. **Data Synthesis.** Moderate to strong evidence is available on the effects of several format interventions to influence probability perceptions and feelings, including the 1-in-X number format, foreground-only (numerator-only) icon arrays, bar charts, anecdotes, framing, and verbal probabilities. However, only 3 (the 1-in-X effect, anecdotes, and framing) had moderate to strong evidence of influencing health behaviors and behavioral intentions. Research on patient preferences for numerical, graphical, and verbal formats yielded only weak evidence. **Conclusions.** The link between probability perceptions/feelings and health behaviors is not strongly reflected in the evidence about communicating numbers because many communication-focused studies measure short-term response rather than longer-term behaviors. Also, research into patient preferences for numerical, graphical, and verbal formats has not yielded strong evidence suggesting stable and predictable preferences.

Corresponding Author
Jessica S. Ancker, Department of Biomedical Informatics, Vanderbilt University Medical Center, 2525 West End Ave, Suite 1475, Nashville, TN 37203, USA; (jessica.s.ancker@vumc.org).

Highlights

- Formatting a probability as 1 in X, using a foreground-only icon array, adding anecdotes to numbers, and gain-loss framing all affect probability perceptions and feelings.
- The evidence on communicating numbers to influence perceptions is far stronger than the evidence on using it to change health behavior or behavioral intention.
- Only weak evidence is available on patient preferences for verbal, graphical, and numerical probability formats.

Keywords
numeracy, health literacy, health communication, risk communication, risk perception, data graphics

Date received: December 17, 2022; accepted: January 4, 2024

As people make decisions about their health, they frequently require information about health probabilities such as the likelihood of disease or the chance of side effects. This information may be conveyed as numbers (e.g., percentages, rates per 10^n, 1 in X, or other formats), graphics (such as icon arrays, bar charts, number lines, or others), and verbal descriptions of probabilities (such as "rare" or "unlikely"). We synthesized the evidence on how to communicate health-related numbers in medical and health domains through a large systematic review (Prospero registration number CRD42018086270).[1,2] We applied a conceptual model of communication in which a stimulus (containing data presented in a data presentation format) is available to a reader, who performs a cognitive task upon it to extract meaning, resulting in a cognitive, affective, perceptual, or behavioral response captured with an outcome measure.[1,2]

This article focuses on research on communicating health probabilities, that is, the chances of health events occurring. Throughout, we use the term *probability* rather than *risk* because of the sometimes ambiguous meaning of *risk* (for example, in epidemiology, *risk* is the probability of an event, whereas in environmental science, *risk* is the hazard times its probability).

In the systematic review, we grouped the research literature by the task performed by the research participant while looking at the stimulus. The current article includes point tasks, in which readers examine a stimulus for information about single probabilities, such as the chance of cancer recurrence. Future articles (see Table A) will cover difference tasks, those in which readers seek information about the differences between probabilities, such as the effect of a risk factor upon chance of disease. Other future articles will cover synthesis tasks, in which the reader integrates several probabilities, such as the set of risks of a medication or a list of risks and benefits. As shown in Table A, a subset of synthesis task research involves interpreting probability information to estimate Bayesian posterior probabilities. Another article will cover time-trend tasks, in which readers examine stimuli to evaluate patterns over time.

To keep article length manageable, the point research has been divided into a pair of articles. The current article presents point task evidence on commonly assessed outcomes: 1) probability perceptions (perceived likelihood of an event; sometimes called risk perception) or probability feelings (self-reported emotional response to

Department of Biomedical Informatics, Vanderbilt University Medical Center, Nashville, TN, USA (JSA); Columbia University School of Nursing (NCB), New York, NY, USA; Department of Population Health Sciences, Weill Cornell Medical College, New York, NY, USA (MMS); Department of Population Health, New York University Langone Health, New York, NY, USA (SBJ); Samuel J. Wood Medical Library, Weill Cornell Medical College, New York, NY, USA (MD, DD); Department of Health Behavior and Health Education, University of Michigan, Ann Arbor, MI, USA (BJZ-F); Department of Internal Medicine, University of Michigan, Ann Arbor, MI, USA (BJZ-F); Center for Bioethics and Social Sciences in Medicine, University of Michigan, Ann Arbor, MI, USA (BJZ-F). The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article. The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: Financial support for this study was provided entirely by a grant from the National Library of Medicine (R01 LM012964, Ancker PI). The funding agreement ensured the authors' independence in designing the study, interpreting the data, writing, and publishing the reports.

Table A Current Article's Scope within the Making Numbers Meaningful Systematic Review

Outcomes	Section Number[a]	Probability Tasks				
		Point	Difference	Trend	Synthesis	Synthesis Bayesian
Identification or recall	1					
Contrast	2					
Categorization	3					
Computation	4					
Probability perceptions or feelings	5	This article				
Effectiveness perceptions or feelings	6					
Behavioral intention	7	This article				
Behavior	7	This article				
Trust	8	This article				
Preference	9	This article				
Discrimination	10	This article				

[a]This standardized numbering system is used within the current article and across all Making Numbers Meaningful articles to allow the reader to find comparable information in different articles. Gray cells represent combinations that are not possible according to the definitions presented in Ancker et al.[1]

the message, including worry or concern; also sometimes called risk perception), 2) behavioral intentions (self-reported decision or plan to take some action) and health behaviors (self-reported or objectively assessed action), 3) trust in the information (self-reported credibility of the communication or the information), and 4) preference for or perception of the data presentation format (including perceived helpfulness of the format and how much the participant liked a format). Finally, we also include 5) discrimination, that is, the ability to make distinctions between probability levels on the basis of differences in the stimuli. This article includes the evidence on the effects of all different probability presentation formats—numerical, graphical, and verbal—on these outcomes. (The companion article[3] presents the parallel evidence for 5 additional outcomes.)

Methods

The systematic review and classification methods are reported in our companion methodology article.[1] In brief, we conducted a broad literature search through 2020 to find experimental (randomized) and quasi-experimental (nonrandomized) research comparing 2 or more formats for presenting quantitative health-related data to patients or other lay audiences. We searched MEDLINE, Embase, CINAHL, the Cochrane Library, PsycINFO, ERIC, and ACM Digital Library and conducted hand searches of the tables of contents of *Medical Decision Making, Patient Education and Counseling, Risk Analysis*, and *Journal of Health Communication*. The current article presents only research about probabilities.

We defined a data presentation format as a method of communicating a probability: the data presentation formats of interest were numbers, graphics, and verbal descriptions of probabilities. Numbers used to express single probabilities of health events included percentages, frequencies, counts, and age-based estimates such as "heart age" or "additional years of life." We distinguished between 2 types of frequencies: 1 in X (examples include "1 in 5" and "1 of every 25") and rate per 10^n (such as "12 in 100" or "2.5 per 1,000"). Unlike some others, we use the term *natural frequencies* only for presentations of a series of probabilities and joint probabilities computed from the same pool of patients in the context of Bayes' theorem.[4,5] (This definition is congruent with the original formulation of the term,[5] and using the term only for this purpose helps clarify otherwise contradictory findings about these numbers.[4]) Graphical formats for single probabilities included icon arrays, number lines and risk ladders, bar charts, pie charts, and flow charts as well as novel creations such as animated icon arrays and slide shows. Verbal formats for single-probability information included probability terms such as *rare* or *likely*. In addition, we assessed manipulations of or variations on the formats that appeared frequently in the literature (such as gain-loss framing, addition of or variation of contextual information [including order effects], addition of anecdotes, and manipulation of denominators for frequencies).

We identified 316 articles published through 2020 focused on health probability communication. Each included article is covered in 1 or more of our review articles (listed in the mapping file at the Making Numbers Meaningful Project at OSF; https://osf.io/rvxf2/). Of these, 161 articles (reporting 466 distinct findings) involved point tasks related to the 7 outcomes described above. The risk of bias of each study (S-ROB) was assessed according to a standard rubric developed for this project, which considered sample representativeness, randomization, protocol deviations, presence/absence of demographic and covariate information, missing data, and other potential biases.[1] From each study, we extracted information about task, stimulus (data and data presentation format), and outcome to produce multiple findings per article. We selected outcomes to track informed by behavioral and risk communication theory (behavior or behavioral intention, probability perceptions or feelings, recall) or empirically on the basis of what was frequently measured by the research included in our review (trust, preference for a format), particularly measures used to measure comprehension (identification, contrast, computation, categorization, discrimination).

Each finding was assessed by a panel of expert reviewers (J.S.A., B.J.Z.-F., with N.C.B. substituting in case of conflict of interest), who weighed sample size, statistical methods, effect size, face validity and comparability of the stimuli being compared, face validity or criterion validity of the outcome measures and covariates at the finding level, as well as the S-ROB for the study from which the finding came. Credibility was assigned holistically on a scale from 1 to 10 on the basis of the expert team's evaluation of these factors. We grouped findings by task type and outcome type and developed evidence statements for each group. We used the finding's risk of bias, the credibility of each finding, and the consistency of findings grouped together in a standard rubric to assess the strength of evidence for each guidance statement. Consistency was considered moderate if findings showed a mix of significant effects in one direction and lack of significant effects; consistency was considered low if the findings showed significant effects in different directions.

- **Strong:** High consistency within a group of 2 or more high-credibility findings or a mix of high- and moderate-credibility findings.
- **Moderate:** a) High consistency within a group of 2 or more moderate-credibility findings or b) moderate consistency within 2 or more moderate-to-high-credibility findings.
- **Weak:** Moderate consistency within a group of 2 or more moderate-credibility findings or only a single high-credibility finding.
- **Insufficient evidence—too few findings:** a) Only low-credibility findings available or b) only 1 moderate-credibility finding.
- **Insufficient evidence—conflicting findings:** Any case in which evidence consistency was low.

Findings with high credibility (7 or higher on a scale of 1 to 10) and moderate credibility (4.5–6.5) are discussed below. Findings with credibility of 4 or lower are mentioned in the text below, counted in Table B, and cited in the Findings files, but they do not contribute to the evidence summaries or the evidence tables.

The Making Numbers Meaningful Project at OSF (https://osf.io/rvxf2/) contains all supplemental files, including a Methodology Files folder containing the search strategy, the data extraction instrument, and the S-ROB rubric).

Results

Each subsection below includes evidence on the following comparisons in order: comparisons among number formats, among graphics formats, between number and graphic formats, between number and verbal formats, between different types of contextual elements, effect of framing, effect of representations of uncertainty, effect of manipulations of denominators, effect of animation or interactivity, and manipulations of time period.

Within each subsection (listed in Table B), evidence is arranged from strongest to weakest. Each subsection concludes with a table of the evidence-based guidance, arranged in the same order; each paragraph in the subsection corresponds to a row of that subsection's evidence table. The full spreadsheet of point task findings is available in the Probability Findings folder in the Making Numbers Meaningful Project at OSF (https://osf.io/rvxf2/).

Effects of Different Formats on Probability Perceptions and Feelings (Probability Perception and Probability Feelings Outcomes): Section 5

Researchers presenting the probability of an event to a patient sometimes measured perceptions on a quantitative or ordinal scale of size (e.g., how large the probability seems on a scale of 1 to 10) and sometimes on an affective scale (e.g., how concerned the patient feels about the event). We abstracted the first of these as "probability perceptions" and the second as "probability feelings." Both are presented in this section.

Importantly, in this section, we do not cover research into the accuracy of the probability perception or feeling. The accuracy of identification of a probability (when the

Table B Numbers of Findings by Outcome and Data Presentation Format Comparison

Subsection		Section (n Findings)					
		Outcomes					
Data Presentation Format Comparison		Probability Perceptions/ Probability Feelings	Behavioral Intention/ Health Behavior	Trust	Preference for Format	Discrimination	Total Findings per Data Presentation Format Comparison
	Section Number/ Subsection Letter[a]	5	7	8	9	10	
Comparisons between numerical formats	A	5A (n = 51)	7A (n = 16)	8A (n = 3)	9A (n = 14)	10A (n = 0)	84
Comparisons between graphical formats	B	5B (n = 47)	7B (n = 5)	8B (n = 2)	9B (n = 17)	10B (n = 2)	72
Comparisons between numerical and graphical formats	C	5C (n = 52)	7C (n = 7)	8C (n = 3)	9C (n = 13)	10C (n = 4)	74
Comparisons between numerical and verbal probabilities	D	5D (n = 24)	7D (n = 13)	8D (n = 0)	9D (n = 20)	10D (n = 1)	58
Comparisons of elements added for context	E	5E (n = 42)	7E (n = 11)	8E (n = 2)	9E (n = 8)	10E (n = 6)	66
Comparisons of frames (gain, loss, combination)	F	5F (n = 21)	7F (n = 9)	8F (n = 1)	9F (n = 6)	10F (n = 0)	37
Comparisons of methods for representing uncertainty	G	5G (n = 19)	7G (n = 1)	8G (n = 5)	9G (n = 4)	10G (n = 0)	29
Comparisons of larger v. smaller denominators	H	5H (n = 9)	7H (n = 0)	8H (n = 0)	9H (n = 0)	10H (n = 0)	9
Comparisons animation or interactivity	I	5I (n = 16)	7I (n = 5)	8I (n = 0)	9I (n = 3)	10I (n = 1)	21
Comparisons of shorter or longer time periods	J	5J (n = 1)	7J (n = 0)	8J (n = 0)	9J (n = 1)	10J (n = 0)	2
Total findings per outcome		282	67	16	86	14	465

[a]This standardized numbering system used in the rows and columns of this table has been used for results subheadings across all Making Numbers Meaningful (MNM) results articles. The standard numbers ensure that, for example, studies comparing graphical formats for their effects on behavior are always placed in a subhead labeled subsection 7B (whether or not there are sections labeled 1A though 6J or 7A in that particular article). Our goal is to ensure that readers can use this subhead system to more easily locate similar sections across articles. The full list of section headers is available in the Methodology Files folder at the MNN Project at https://osf.io/rvxf2/.

stimulus is visible) or recall of it (when the stimulus is not visible) is covered in our companion Part 1 article under section 1, the so-called "identification-recall" outcome.

Comparisons between numerical formats in effects on probability perceptions and feelings (subsection 5A). 1-IN-X VERSUS OTHER FORMATS: A total of 16 findings demonstrated a 1-in-X effect. Thirteen moderate and high-credibility findings showed that 1-in-X formats led to greater perceived probability and/or feelings than rate per 10^n or percentage formats (Fair et al. [perceptions and feelings][6]; Chapman et al.[7]; Sirota et al.[8] substudy 1 [perceptions and feelings] and 3; Pighin et al.[9] findings from substudies 1 [both perceived probability and feelings], 2, 3, and 6; Sirota et al.[10]; Pighin et al.[11]). Knapp et al.[12] showed the 1-in-X effect when compared to a "probability band," which was a range of probabilities in 1-in-X format (e.g., a chance of "up to 1 in 100" or "1 in 10 to 1 in 100"). A moderate-quality finding[12] showed that the 1-in-X probability band format was associated with higher perceived probability than a rate per 10^n value within that probability band (probability feeling not assessed). Two findings from the same Wu et al.[13]

Table 5A Evidence-Based Guidance for Effects of Numerical Formats on Probability Perceptions and Feelings

Comparison	Evidence Strength	Applied Example of Evidence-Based Communication	General Guidance
1 in X v. other formats	Strong (*n* = 22)	People may be more concerned about a 1 in 10 chance than a 1% chance or a 10 in 100 chance.	Probability perceptions and feelings are both higher when chances are shown as 1 in X than when shown as rate per 10^n or percentages.
Percentages v. rate per 10^n	Strong (*n* = 6)	People are equally concerned about a 1% chance and a 10 in 100 chance.	Probability perceptions and feelings do not appear to change when chances are shown as 1 in X or as percentages.
Arithmetic difference added to individual probabilities	Weak (*n* = 7)	If the chance of a bad outcome with drug A is 4% and with drug B is 6%, people may perceive the chance from drug B to be lower if they are also told that the increased chance with drug B is 2 percentage points.	Perceived probability of an event with a risk factor may be lower when the arithmetic difference between the probability of an event with and without the risk factor is added to the 2 probabilities.
Text v. table	Weak (*n* = 5)	It may not affect perceived probability or probability feelings whether a chance is presented in text or in a table.	Probability perceptions and feelings may not be affected by whether rates or percentages are presented in text or in tables.
More v. fewer significant digits	Weak (*n* = 1)	People may be more concerned about a 10.455% chance than a 11% chance.	Probability perceptions and feelings may be higher when a chance has more significant digits.
Heart age v. probability	Weak (*n* = 4)	People may be more concerned to learn that their heart age is 10 y older than their chronological age than they are to learn that they have a 10% chance of heart disease.	Probability perceptions and feelings may be higher when the chance of a cardiovascular event is shown as "heart age" than when it is shown as a percent chance.
Order effects	Weak (*n* = 2)	People may be more concerned when told, "The chance that the drug will help is 80%, but the chance of a side effect is 1%," than when told, "The chance of a side effect is 1%, but the chance that the drug will help is 80%."	In risk-benefit communication, feelings about side effects may be higher when benefits are presented before harms than the reverse order.
Verbalizing numbers v. numerals	Insufficient (too few findings; *n* = 1)	It is not clear whether expressing a 1-in-X frequency in terms of a large time interval (e.g., 1 case per 3,500 y) affects probability perceptions and feelings.	
Time interval variations	Insufficient (too few findings; *n* = 1)	It is not clear whether verbalizing 1 in X as "one out of every X" affects probability perceptions and feelings.	

publication (substudies 2 and 3) produced moderate-credibility findings demonstrating the 1-in-X effect (compared to percentage) but only for some probability levels and some emotional priming conditions.

A total of 6 moderate- to high-credibility findings did not find a 1-in-X effect (Gurmankin et al,[14] Knapp et al.,[15] Sirota et al.[8] substudies 2, 4, and 5 [perceptions and feelings]).

PERCENTAGE VERSUS RATE PER 10^n: While 1 high-credibility finding (Hill and Brase[16] substudy 1) and 4 moderate-credibility findings (Peters et al.,[17] Halpern et al.,[18] Peters et al.[19] substudy 1, Hill and Brase[16] substudy 2) compared percentage with rate per 10^n (such as 10 in 100), none showed a significant difference in perceived probability. One moderate-credibility finding[20] did find lower perceived probability with rates per 10^n embedded in text versus percentages in a table format, but confounding of format and table limits the ability to draw conclusions from this finding. Probability feeling was not assessed in these findings.

ARITHMETIC DIFFERENCE: Five findings examined the impact of providing the arithmetic difference between probabilities (compared with providing pairs of probabilities alone) on perceived probability of individual probabilities. (Arithmetic difference between probabilities may be called absolute risk difference, absolute risk reduction, or absolute risk increase.) Zikmund-Fisher et al.[21] substudy 2 (high credibility) and Shepperd et al.[22] (moderate credibility) demonstrated lower worry or concern about a negative outcome when the arithmetic difference between those with and without the risk factor was presented versus pairs of rates per 10^n alone. Similarly, both Zikmund-Fisher et al.[21] substudy 2 (high credibility) and Zikmund-Fisher et al.[23] (moderate credibility) showed lower perceived probability of the post-treatment probability in a similar comparison, and Berry et al.[24] (moderate credibility) showed lower perceived probability with harm when arithmetic difference was shown together with baseline probability than for the arithmetic difference alone. However, Sullivan et al.[25] (high credibility) showed no differences in perceived probability when medication side effect rates (percentage + rate per 10^n) were supplemented by the arithmetic difference between medication and no medication, and Shepperd et al.[22] (moderate credibility) similarly showed no differences in perceived probability of cancer when arithmetic difference was added to rate per 10^n.

TEXT VERSUS TABLE: A high-credibility finding and a moderate-credibility finding from the same author group demonstrated no differences in either perceived probability[26,27] or emotional concern[26] when medication side effects were presented as rate per 10^n in text versus in table format. However, in 2 high-credibility findings, Schwartz et al.[28] substudy 1 and 2 showed higher perceived probability when percentages were embedded in dense text than presented with rate per 10^n rates and arithmetic difference in drug facts box table, although the multiple differences between arms limits ability to attribute the effect to the table format.

MORE VERSUS FEWER DIGITS TO THE RIGHT OF THE DECIMAL POINT: One high-credibility finding from a study of the number of significant digits showed that an increase in the number of significant digits led to higher perceived probability.[29] However, there was not a linear relationship between increasing perceived probability and increasing number of digits.

HEART AGE VERSUS PROBABILITY: In 2 high-credibility findings, Bonner et al.[30] showed both higher perceived probability and stronger probability feelings when a personalized estimate of cardiovascular risk was shown as "heart age" plus the difference from actual age versus as pair of percentages plus arithmetic difference between them. However, 2 moderate-credibility findings comparing percentages with heart age showed no difference in perceived probability or probability feeling.[31] Translating these findings into practice may be challenging in the absence of a single universally accepted "Heart Age" calculator.

ORDER EFFECTS: In a high-credibility finding, Ubel et al.[32] demonstrated higher perceived probability and higher concern about side effect chances when benefits were presented before risks.

VERBALIZING NUMBERS: In a moderate-credibility finding, Miron-Shatz et al.[33] showed that writing 1-in-X statistics as "1:X" resulted in higher perceived probability than when the same ratio was written out as "one in every X."

TIME INTERVAL: One moderate-credibility finding[34] demonstrated that adding a large time interval (1 case per 3,500 y) to a 1-in-X chance reduced perceived threat, but a smaller time interval had no effect. Perceived probability was not assessed.

NOT SUMMARIZED: Two lower-credibility findings are not summarized due to insufficient information in the stimulus[35] and lack of statistical power and hypothesis testing.[36]

Comparisons between graphical formats in effects on probability perceptions and feelings (subsection 5B). PART-WHOLE VERSUS FOREGROUND-ONLY GRAPHICS: Studies comparing foreground-only/numerator-only graphics versus part-to-whole graphics produced 10 findings. Five high-credibility findings demonstrated higher perceived probability with numerator-only bars[37,38] or icon arrays (Okan et al.[39]

Table 5B Evidence-Based Guidance for Effects of Graphical Formats on Probability Perceptions and Feelings

Comparison	Evidence Strength	Applied Example of Evidence-Based Communication	General Guidance
Part-whole v. foreground-only graphics	Strong (n = 10)	People will perceive a probability as larger when shown a line of 4 icons alone than when shown an array of 96 white icons and 4 colored ones.	People perceive a probability as larger with a numerator-only graphic (icon arrays or bar charts) than with part-to-whole graphics. There is inconsistent evidence about whether the effect extends to probability feelings as well.
Icon shape variation	Strong (n = 3)	People will NOT perceive probability differently in an icon array of stick figures and an icon array of blocks.	Using different shapes in icon arrays (abstract v. anthropomorphic) does not affect perceived probability. Probability feelings were not assessed in these findings.
Single outcome v. multioutcome icon arrays	Strong (n = 5)	People will NOT perceive a probability as different when it is shown in side-by-side icon arrays each showing probability of a single negative outcome or as multioutcome icon arrays showing probability of several outcomes.	Presenting chances in side-by-side single-outcome icon arrays versus multioutcome integrated icon arrays does not affect perceived probability. There is weak evidence to support the same noneffect finding on probability feelings.
Visual representation of treatment effect	Strong (n = 5)	People will perceive and feel differently about a chance of a health outcome when they can see how it is affected by a protective factor (such as exercise) or a risk factor (such as a medication that raises the chance of side effects).	Illustrating the incremental change in probability of an outcome on the basis of a risk factor or therapy will alter perceptions and feelings about the probability of the outcome. Direction of the alteration may differ on the basis of context. Type of visual illustration does not appear to make an important difference.
Icon arrays v. bar charts	Moderate (n = 2)	People may perceive the chance to be larger in a bar chart depicting a 25% chance than in an icon array with 25 of 100 icons colored.	People perceive a probability as larger when it is in a bar chart than when it is in an icon array.
Random v. grouped icon arrangement	Moderate (n = 8)	People will NOT perceive or feel a probability to be different with an icon array with grouped icons than one with random icons.	Grouping versus randomly distributing icons in icon arrays does NOT affect probability perceptions or feelings.
Icon arrays v. linear scale number lines	Weak (n = 2)	If we show that the chance of heart disease without smoking is 4% and with smoking is 8%, the chance with smoking will appear higher when it is shown on a number line rather than with icon arrays.	When showing a pair of pre-post chances of an event, perceived probability of the post chance may be higher with a number line than with icon arrays.
Log-scale number lines v. icon arrays	Weak (n = 2)	People may perceive a 1% chance placed on a log-scale number line, with probabilities of other events graphed on the same line, to be larger than 1% shown in an icon array.	Using a log-scale number line with the probability placed in context of comparison probabilities, instead of a part-whole icon array, may increase probability perceptions and/or feelings.
Other graphics	Weak (n = 2)	People may NOT perceive a probability to be different with a bar graph versus a box-and-whisker plot.	Using bar graphs versus box-and-whisker plots may not affect probability perceptions and/or feelings.
Number labels v. no labels	Insufficient (inconsistent findings; n = 3)	It is not clear whether graphics with verbal-only labels versus verbal + numerical labels affect probability perceptions or feelings.	
Social comparisons	Insufficient—too few findings (n = 1)	It is not clear whether labeling quantitative graphics with "higher than average" or "lower than average" alters probability feelings.	

substudies 1 and 2, Stone et al.[40] substudy 2) than with part-to-whole displays. Stone et al.[37] showed the same pattern in probability feelings (fear). Five additional findings, all moderate credibility, showed no differences in perceptions or feelings. These included 2 moderate-credibility findings[22] showing no differences in either perceived probability or probability feelings in a small sample. In addition, 1 moderate-credibility finding compared a denominator-only icon array with a part-to-whole icon array (using very large numbers of typed dots rather than icons), finding no difference in perceived probability (Weinstein et al.[41] substudy 1). A moderate-to-high-credibility finding (Stone et al.[40] substudy 2) found no differences in probability feelings between denominator-only and part-to-whole icon arrays.

ICON SHAPE VARIATION: Three moderate-to-high-credibility findings compared icon shapes (e.g., abstract versus anthropomorphic), finding no difference in perceived probability, although probability feelings were not assessed.[42–44]

SINGLE-OUTCOME VERSUS MULTIOUTCOME ICON ARRAYS: Several studies examined the problem of simultaneously conveying the chances of several mutually exclusive outcomes of an event. These studies compared 1 stimulus that contained several icon arrays, each portraying the chance of a single outcome, with a second stimulus containing 1 icon array portraying the chances of the different outcomes integrated into the same array. A high-credibility finding (Zikmund-Fisher et al.[21] substudy 2) and 1 moderate-credibility finding (McDowell et al.[45] substudy 2) all showed no difference in perceived probability between side-by-side icon arrays versus multioutcome integrated icon arrays. Zikmund-Fisher et al.,[21] substudies 1 and 2 (both high credibility), also showed no differences in probability feeling (worry). However, a moderate-credibility finding[23] showed lower perceived probability of side effects with multioutcome icon arrays that showed baseline probability plus arithmetic difference versus side-by-side icon arrays.

ICON ARRAYS VERSUS BAR CHARTS: One high-credibility finding[46] and one moderate-credibility finding[42] compared icon arrays to bar charts, both finding lower perceived probability of disease with icon arrays than with bar charts. Probability feelings were not assessed.

ICON ARRAYS VERSUS LINEAR SCALE NUMBER LINES: Both Adarkwah et al.[47] (high credibility) and Adarkwah et al.[48] (moderate credibility) demonstrated that perceived chance of disease was higher with pre/post numbers (percentages or rate per 10^n) placed on a number line than with an integrated icon array.

ICON ARRANGEMENT: Eight findings included a comparison of random versus grouped icons in icon arrays. One moderate-credibility finding demonstrated that perceived probability was higher with the random array,[42] but 4 moderate-credibility findings demonstrated no effect on perceived probability.[49–52] In addition, 3 moderate-credibility findings showed no difference in probability feeling.[49,50,52]

VISUAL REPRESENTATION OF TREATMENT EFFECT: In communicating the effect of medication on chance of side effects with rates per 10^n and icon arrays, 2 high-credibility findings from Zikmund-Fisher et al.[21] was that, in stimuli that illustrated the incremental risk, worry was not affected by whether the increment was illustrated with side-by-side icon arrays or a single icon array showing the increment. One of these showed the same pattern for perceived probability.[21] However, in a moderate-credibility finding, Janssen et al.[53] showed that in communicating "your chance" of several chronic diseases on a vertical risk ladder, perceived probability of the disease was slightly higher when the ladder also portrayed the (lower) postexercise probabilities showing how much exercise could reduce risk. Worry was not affected.[53]

LOG-SCALE NUMBER LINES VERSUS ICON ARRAYS: Two moderate-credibility findings (Siegrist et al.[54] substudy 1) showed that a log-scale number line that included probabilities of comparison events resulted in higher perceived probability and emotion than a part-whole icon array, although it is not clear how much the effect was a result of graphics versus the additional comparison probabilities.

OTHER GRAPHICS: Two moderate-credibility findings[55] compared several different graphics including bar graphs and box-and-whisker plots with no significant effect on probability perceptions or feelings.

NUMBER LABELS VERSUS NO LABELS: In a moderate-credibility finding, Janssen et al.[53] showed no differences in perceived probability when personal probability shown on a risk ladder was accompanied by verbal labels only or verbal labels plus rate per 10^n. Probability feelings were also not affected.[53] However, in a moderate-credibility finding, Okan et al.[38] showed that in communicating effect of drug therapy, adding rate per 10^n data labels to bar charts slightly reduced perceived probability (however, the measure of perceived probability was conflated with a measure of perceived effectiveness, somewhat reducing ability to draw conclusions about either alone).[38]

SOCIAL COMPARISONS: In a moderate-credibility finding by Emmons et al.,[56] worry about colorectal cancer did not differ significantly when risk ladders labeled with social comparisons (e.g., "higher than average") were added to either static or interactive risk ladders that used scales of rates per 10^n to communicate absolute probabilities.

NOT SUMMARIZED: Findings from lower-credibility studies were not summarized due to lack of power,[57] stimulus complexity (Damman et al.[58] for both perceptions and feelings), and lack of clarity regarding stimuli and outcome measures (Weinstein et al.[41] substudy 2).

Comparisons between numerical and graphical formats, and combinations of numerical and graphical formats, in effects on probability perceptions and feelings (subsection 5C). ICON ARRAYS VERSUS NUMBERS: Five high-to-moderate-credibility findings demonstrated that grouped icon arrays were associated with lower perceived probability than rate per 10^n or percentages (Siegrist et al.[54] substudy 1, Keller and Siegrist,[59] Galesic et al.[60] substudy 2, Navar et al.,[46] Tait et al.[26]). Both the first and the last of these (both high-credibility) demonstrated the same effect for concern about side effects. However, 2 high-credibility findings showed no difference between icon arrays and percentages,[52] for either perceived probability or probability feelings, and 3 moderate-credibility findings also showed no difference in perceived probability between icon arrays and frequency numbers (Weinstein et al.[41] substudy 1, Tait et al.,[27] Colome et al.[61]). The Weinstein et al. finding used sheets with large numbers of dots, rather than a traditional icon array.

ADDING AN ICON ARRAY TO A RATE PER 10^n: Two high-credibility findings (Fraenkel et al.[62] for perceptions and feelings) and 7 moderate-credibility findings (Cameron et al.,[63] McDowell et al.[45] substudy 2, Shepperd et al.[22] with 1 finding for feelings and another for perceptions, Zikmund-Fisher et al.,[23] Cozmuta et al.[64] for feelings and another finding for perceptions) examined the effect of adding a part-to-whole icon array to rate per 10^n numbers, consistently finding no differences in either perceived probability (6 findings) or concern (3 findings). However, when communicating effect of medication on chance of side effects with rates per 10^n with or without icon arrays, a moderate-credibility finding from Zikmund-Fisher et al.[21] was that worry about 1 out of the 4 side effects was reduced when the pre and post risk was explicitly stated in side-by-side icon arrays labeled with numbers versus numbers only.

FOREGROUND-ONLY ICON ARRAYS VERSUS NUMBERS: One high-credibility finding and 1 moderate-credibility finding demonstrated higher perceived probability with foreground-only icon arrays than with rate per 10^n (Stone et al.,[40] subsubstudies 1 and 2). Similarly, in a moderate-credibility finding, Stone et al.,[65] substudy 2, showed higher probability estimates with foreground-only icon arrays than with rate per 10^n, but concern was not affected. However, a high-credibility finding (Stone et al.[40] substudy 2) demonstrated no differences in probability feelings between rate per 10^n and foreground-only or part-to-whole icon arrays, and a moderate-credibility finding[22] showed no differences in either probability perceptions or feelings about cancer between rate per 10^n and foreground-only icon arrays.

LOG-SCALE NUMBER LINES VERSUS NUMBERS: Five findings examined a horizontal number line with a log-scale axis, all of which showed comparison probabilities or population probability as well as the target probability ("Paling" graphic). Four findings (Siegrist et al.[54] substudies 1 and 2) showed that probability perceptions (1 high and 1 moderate credibility) and feelings (both high credibility) were higher with the log-scale number line than with 1 in X, and 1 moderate-credibility finding[59] showed the same for perceived probability only.

ADDING AN ICON ARRAY TO 1 in X: Two moderate-credibility findings[33,66] demonstrated that estimated probability with 1 in X was reduced by adding an icon array, but only under certain conditions. In a related finding, a high-credibility finding (Pighin et al.[9] substudy 7) showed that supplementing frequency numbers with an icon array reduced the difference in perceived probability between 1 in X and rate per 10^n.

LINEAR-SCALE NUMBER LINE VERSUS NUMBERS: Two high-credibility findings showed no difference between percentages and number lines (Han et al.[67] substudy 1) for probability perceptions or feelings.

BAR CHARTS VERSUS NUMBERS: A high-credibility finding[46] did not find differences in perceived probability between a pair of percentages versus a bar chart. Similarly, a moderate-to-lower-credibility finding[20] showed no differences between percentages in a table and percentages plus a bar chart, although small sample and other factors limit confidence in this negative finding. However, 1 moderate-credibility finding demonstrated that the stacked bar chart was associated with lower perceived probability than rate per 10^n or more complex graphics.[55] The authors also assessed probability feelings but did not detect any differences.

PIE CHARTS VERSUS NUMBERS: One moderate-credibility finding compared pie charts versus percentages

Table 5C Evidence-Based Guidance for Contrasts between Numerical and Graphical Formats, and Combinations of Numerical and Graphical Formats, on Probability Perceptions and Feelings

Comparison	Evidence Strength	Applied Example of Evidence-Based Communication	General Guidance
Icon arrays v. numbers	Strong (n = 12)	People perceive a chance as smaller when it is portrayed as 4 colored icons with 96 white icons than when it is stated as a 4% chance or a 4 in 100 chance.	Perceived probability is lower with part-to-whole icon arrays than with rate per 10^n or percentages. There is insufficient evidence to determine whether this effect holds true for felt probabilities, and for both, the effect may depend on the size of the probability.
Adding an icon array to a rate per 10^n	Strong (n = 10)	If we tell people about a "4 in 100 chance," their perceptions and feelings will not be changed if we also add an icon array with 4 colored icons and 96 white ones.	Adding part-to-whole icon arrays to rate per 10^n does NOT affect probability perceptions or feelings.
Foreground-only icon arrays v. numbers: perceptions	Strong (n = 4)	People will perceive a chance as larger when it is portrayed as 4 colored icons alone than when it is stated as a 4 in 100 chance.	Perceived probability is higher with foreground-only icon arrays than with rate per 10^n.
Foreground-only icon arrays v. numbers: feelings	Strong (n = 3)	People's probability feelings are NOT affected by replacing a statement about a 4 in 100 chance with 4 colored icons.	Probability feelings are not affected by use of foreground-only icon arrays instead of rate per 10^n.
Log-scale number lines v. numbers	Strong (n = 5)	People will perceive the probability of health-related events as higher, and will be more concerned, with a horizontal log-scale number line, together with comparison probabilities, than with a statement that there is a 1 in 100 chance.	Horizontal number lines with a log-scale axis that also include comparison or population probabilities (the so-called "Paling" graphic) lead to higher-probability perceptions and feelings than 1 in X, although it is not clear whether it is due to the graphic or the presence of comparison risks.
Adding an icon array to 1 in X	Moderate (n = 3)	If we tell people about a "1 in 12 chance," they will perceive the chance to be smaller if we also add an icon array showing 1 colored icon and 11 white ones.	Perceived probability may be lower with part-to-whole 1-in-X icon arrays than with 1-in-X numbers alone; the effect reduces the differences between 1-in-X and rate per 10^n representations. Probability feelings have not been assessed.
Linear-scale number lines v. numbers	Weak (n = 2)	People will perceive the chance to be similar size whether it is shown as 10% on a linear number line or as a 10% chance. Probability feelings are also similar.	Probability perceptions and feelings may not differ when chances are shown with a linear-scale number line or with a percentage.
Bar charts v. numbers	Insufficient (inconsistent findings; n = 4)	It is not clear whether replacing percentages with bar charts affects perceived probability or probability feelings.	
Pie charts v. numbers	Insufficient (too few findings; n = 1)	It is not clear whether replacing percentages with pie charts affects perceived probability or probability feelings.	

(both with uncertainty), finding no difference in perceived probability.[68] This combination has not been assessed for probability feeling.

NOT SUMMARIZED: Some lower-credibility findings from relevant findings are not synthesized due to confounding with contextual manipulations (Damman

Table 5D Evidence-Based Guidance for Contrasts between Numerical and Verbal Probability Formats on Probability Perceptions and Feelings

Comparison	Evidence Strength	Applied Example of Evidence-Based Communication	General Guidance
Verbal v. numeric: European Commission (EC) terms	Strong ($n = 12$)	People consider a side effect described as "common" to be more likely than one that has a chance of "between 1 in 100 and 1 in 10," even though the EC considers the 2 phrasings equivalent.	Describing probabilities in the EC-approved verbal terms results in significantly higher probability perceptions and feelings than describing probabilities with the EC-approved corresponding numbers.
Verbal v. numeric: other verbal terms	Strong ($n = 6$)	People feel differently about a chance described as a 0.1% probability and one described as "rare," but it may be difficult to predict which they will perceive as larger or more concerning.	Describing probabilities in verbal terms results in significantly different probability perceptions and feelings than describing probabilities as numbers, but the direction of the difference may depend on the terms used.
Verbal plus numeric probability	Insufficient (inconsistent findings; $n = 3$)	It is not clear whether combining verbal and numerical labels, as compared to conveying them separately, affects perceived probability or probability feelings.	

et al.,[58] perceptions and feelings findings), modest sample and multiple differences between formats (Lee and Mehta,[69] perceptions and feelings findings), lack of power,[57] unclear statistical testing,[70] or combination of the communication with a counseling intervention (Henneman et al.,[71] perceptions and feelings findings).

Comparisons between numerical and verbal probabilities in effects on probability perceptions and feelings (subsection 5D). VERBAL VERSUS NUMERIC PROBABILITY: Twelve high-credibility findings examined the mapping between verbal probabilities and numbers endorsed by the European Commission (EC).[72] Overall, these findings demonstrated that probability perceptions and feelings with verbal descriptors were higher than with the corresponding EC probability numbers, whether in percentage, 1 in X, or rate per 10^n formats. Three of the findings detected this pattern in perceived probability only (Berry et al.[73] substudy 2, Berry and Hochhauser,[74] Peters et al.[75]), 1 in probability feelings only (Young and Oppenheimer[76] substudy 3), and 8 in both (Knapp et al.[77] substudy 1, Knapp et al.,[12] Berry et al.,[78] Berry et al.[73] substudy 1).

However, 1 high-credibility finding that did not focus on the EC terms[79] demonstrated that perceived probability was lower with verbal probability than with percentage plus rate per 100 or 1,000. Similarly, a moderate-credibility finding[80] showed that verbal descriptions evoked higher risk perceptions for low probabilities but lower risk perceptions for higher probabilities. Also, 2 moderate-credibility findings[81] showed that perceived probability (but not probability feelings) was higher with raw numbers of people affected than with verbal probabilities, and 2 moderate-credibility findings[82,83] showed no difference between verbal probabilities and rates per 100/1,000 or percentages.

VERBAL PLUS NUMERIC PROBABILITY: Several findings assessed the effect of combining verbal probabilities and numeric probabilities but did not find any consistent results. Knapp et al.[84] (high credibility) showed that perceived probability was increased by adding verbal probability to 1 in X, but this change did not affect probability feelings (moderate credibility). A moderate-credibility finding[85] suggests that when genetic test results are labeled "screen negative," adding verbal or numeric probabilities does not affect probability feelings such as anxiety and worry.

NOT SUMMARIZED: A lower-credibility finding was limited by insufficient power (Young et al.[76] substudy 2). A finding from Timmermans and Oudhoff[70] was not summarized due to floor effects, and a finding from de Wit et al.[86] was not summarized due to small sample size limiting confidence in the exact ranking of the different interventions studied.

Comparisons of elements added for context on probability perceptions and feelings (subsection 5E). ANECDOTES: Several studies examined the effect of anecdotes, or short

Table 5E Evidence-Based Guidance for Effect of Adding Contextual Information on Probability Perceptions and Feelings

Comparison	Evidence Strength	Applied Example of Evidence-Based Communication	General Guidance
Anecdotes v. no anecdotes, or different numbers of anecdotes	Moderate ($n = 4$)	A side effects chance of 5% may appear larger if it is accompanied by anecdotes about Tom and Jane who got headaches from the drug.	Adding anecdotes to a numerical probability—and specifically increasing the proportion of anecdotes referring to negative adverse events—increases perceived probability of the event. The effect has not been assessed for probability feelings.
Chance of comparison events v. no comparison events	Moderate ($n = 6$)	People may not perceive a disease chance of 1% to be different if the message also mentions that the lifetime chance of a car accident is 3%.	Adding chances of comparison events (e.g., of getting into a car accident) to a message about a particular health event does not affect perceived probability or probability feelings about that event.
Risk-reducing actions added to probability	Weak ($n = 1$)	People may perceive a 5% chance of heart disease to be lower if they are also told about specific lifestyle changes they can take to reduce their probability.	Perceived probability of disease may be lower when numerical probability information is accompanied by information about so-called risk-action links to risk-reducing actions.
Explanation of uncertainty v. no explanation	Moderate ($n = 3$)	People may not respond any differently to a chance of 5% with the addition of a statement such as, "We cannot predict the future of any one person."	Probability perceptions and feelings may be similar whether or not an explanation of uncertainty is provided with probability information.
Interpretive labels v. no labels	Weak ($n = 2$)	People may respond more strongly to a test result indicating a chance of 5% if they are told that the test result is "abnormal."	Probability perceptions and feelings may be amplified when qualitative interpretive labels (e.g., "normal" or "positive") are added to probabilities.
Descriptive labels v. no labels	Weak ($n = 1$)	When given a series of chances such as 5%, 6%, and 7%, people's perception of the probability may not be affected by the addition of a statement saying that the third option has the highest chance.	Perceived probability may not be affected when numerical probabilities are accompanied by a label indicating which of several options has the higher probability.
Denominator salience variation	Weak ($n = 2$)	When given a chance of 5 in 1,000, adding bold or coloring to the 1,000 may not affect perceived probability.	Using bolding and color to highlight the size of the denominator in rate per 10^n may not affect perceived probability.
Population average v. no population average	Insufficient (inconsistent findings; $n = 15$)	It is not clear whether adding a population value to a personal probability value will affect probability perceptions and feelings.	
Conceptual illustration of tradeoffs	Insufficient (too few findings; $n = 1$)	It is not clear whether providing conceptual illustrations of risk-benefit tradeoffs change perceived probability or probability feelings.	
Social norm manipulation	Insufficient (too few findings; $n = 1$)	It is not clear whether presenting screening behaviors for a probability as a social norm affects perceived probability.	

text narratives about individuals who experienced an event. Three moderate-to high-credibility findings all showed that anecdotes about adverse events increased perceived probability of the adverse event (Gutierrez and Cohn,[87] Betsch et al.[88] substudies 1 and 2; concern not assessed). However, in a more complicated moderate-credibility finding, Sheridan et al.[82] showed perceived probability of disease did not vary when anecdotes were provided in addition to X-in-N numbers.

CHANCE OF COMPARISON EVENTS: Six findings (2 high credibility, 4 moderate credibility) examined the effect of adding comparison chances of other events. Ubel et al.[32] showed that adding comparison events did not affect perceived probability but did eliminate risk versus benefit order effects on probability feelings. Similarly, Schapira et al.[42] showed that perceived probability of breast cancer did not differ when probabilities of other cancers were provided. Siegrist et al.[54] substudy 2 (assessing only perceived probability) showed no difference between showing comparison probabilities as 1 in X or depicted on a log scale number line. A similar finding by the same author (Siegrist et al.[54] substudy 1) assessed both perceived probability and probability feelings but combined the presence of comparison probabilities with the log scale number line format, reducing ability to draw conclusions about the role of either alone.

RISK-REDUCING ACTIONS: A high-credibility finding[63] showed perceived probability of disease was lower when numerical probabilities were accompanied by text describing the "risk-action link" to risk-reducing actions.

EXPLANATION OF UNCERTAINTY: Three moderate-credibility findings examined different ways of verbally explaining uncertainty, with none finding an effect on perceived probability[52,89] or probability feeling.[52]

INTERPRETIVE LABELS: In a high-credibility finding, Zikmund-Fisher et al.[90] showed that adding qualitative interpretive labels "normal" or "positive" to test results had an amplifying effect that either raised or lowered both probability perceptions and feelings.

DESCRIPTIVE LABELS: In a high-credibility finding, Sullivan et al.[25] demonstrated that perceived probability was not affected by whether a table of probabilities was accompanied by labels describing which option had the higher probability for each outcome.

DENOMINATOR SALIENCE: A high-credibility finding (Stone et al.[40] substudy 1) showed no difference in perception of probabilities when a foreground-only icon array accompanied by numerical denominator had denominator salience increased by making the denominator bold and red. A moderate-credibility finding[91] demonstrated a similar noneffect of highlighting either the number of events or the time period.

POPULATION AVERAGE: Of the 15 findings studying the effect of adding or showing a population average, 7 demonstrated an effect on probability perceptions or feelings. Three high-credibility findings (Fair et al.[6] [probability perceptions and feelings], Weinstein et al.[92]) showed that showing or explaining that the person was higher than the population average increased probability perceptions and/or feelings. One moderate-credibility finding[93] found that showing that the person had higher probability than the lowest-probability group also increased perceived probability. Two moderate-credibility findings (Han et al.[67] substudy 1) demonstrated increased probability feeling but not probability perceptions with the addition of population averages. Lastly, Janssen et al.[53] (moderate credibility) showed that in communicating "your chance" of several diseases with a vertical risk ladder, perceived probability was lower when accompanied by verbal social comparison labels such as "much higher than average." However, 7 moderate-quality findings showed no effect of showing the population average on perceived probability (Hess et al.[94] finding 1, Gibson et al.,[55] Harris and Smith[95]) or negative emotion.[53,55,93,95] In a separate moderate-credibility finding (Raghubir,[96] finding 1) shifting the reference group so that counts of deaths represented a larger fraction of a smaller denominator lead to higher perceived probability.

CONCEPTUAL ILLUSTRATION OF TRADEOFFS: A moderate-credibility finding[64] demonstrated no differences in probability feelings when probability information was accompanied by balance beam graphics conceptually illustrating the risk-benefit tradeoff.

SOCIAL NORM MANIPULATION: A moderate-credibility finding[97] demonstrated higher perceived probability of cancer when screening was described as the default social norm.

NOT SUMMARIZED: Several lower-credibility findings are not synthesized due to questions of generalizability (Dillard et al.[98] substudy 2), lack of power,[57,86] and confounding with graphical manipulations.[58,59]

Effects of gain-loss framing on probability perceptions and feelings (subsection 5F). GAIN VERSUS LOSS FRAMING OF NEGATIVE EVENTS—PERCEPTIONS: Eight high- and moderate-credibility findings looked at the impact on perceived probability of

Table 5F Evidence-Based Guidance for Effects of Gain-Loss Framing on Probability Perceptions and Feelings

Comparison	Evidence Strength	Applied Example of Evidence-Based Communication	General Guidance
Gain v. loss framing of negative events—probability perceptions	Moderate ($n = 10$)	People will perceive the probability of disease as larger when stated as 10% chance of disease and smaller when it is stated as a 90% chance of no disease.	The chance of a negative event is associated with higher perceived probability when it is loss framed (presented as the chance of the negative event occurring). The effect on probability feelings is unclear.
Gain v. loss framing of positive events—probability perceptions	Strong ($n = 11$)	People will perceive the probability of success as larger when it is stated as a 90% chance of success than when it is stated as a 10% chance of failure.	The chance of a positive event is associated with higher perceived probability when it is gain framed (presented as the chance of the positive event occurring). The effect on probability feelings is unclear.
Gain v. loss framing of both positive and negative events	Insufficient—too few findings ($n = 1$)	It is unclear how probability perceptions and feelings are affected when both the chances of harm and chances of benefit of an intervention are gain or loss framed.	

framing of negative events (such as mortality or chance of side effects), with 4 (including 2 high credibility) finding that perceived probability of the negative event was higher when loss-framed (Peters et al.,[17] Garcia-Retamero and Galesic,[99] Chapman et al.[7] substudy 2, Halpern et al.[18]) and 4 (including 1 high-credibility) finding no effect (Williams et al.,[100] Chapman et al.[7] substudy 4, Farrell et al.,[101] Kalluru et al.[102]). Two additional findings showed no similar effect on probability feelings.[101,103]

GAIN VERSUS LOSS FRAMING OF POSITIVE EVENTS—PERCEPTIONS: Eleven high- and moderate-credibility findings examined the effect of framing positive events (such as effectiveness of a vaccine), all finding that gain-framing increased the perceived probability of that event (Peng et al.[104] substudy 1; Peng et al.[105] substudy 1; Kreiner and Gamliel[106]; Bigman et al.[107]; Zamarian et al.[108]; Gamliel et al.[109]; Biswas and Pechmann[110] findings 1, 3, 4, 5; Levin et al.[111]).

GAIN VERSUS LOSS FRAMING OF BOTH POSITIVE AND NEGATIVE EVENTS—PERCEPTIONS: One examined the effect of framing both the harms and the benefits of an intervention, finding no effect on perceived probability of disease.[82]

Comparisons of methods of representing uncertainty on probability perceptions and feelings (subsection 5G). EXPLANATION OF UNCERTAINTY: Three moderate-credibility findings examined different ways of supplementing numbers with an explanation of uncertainty in words, none finding an effect on perceived probability[52,89] or probability feelings.[52]

POINT ESTIMATES VERSUS UNCERTAINTY NUMBERS: Twelve moderate-to-high-credibility findings compared point estimates (as numbers or graphs) to uncertainty displayed as range of numbers, range of points on a graph, or distribution graphs. Of these, 5 showed increased perceived probability (Johnson and Slovic[112] substudies 2 and 4, Gibson et al.[55]) or probability feeling (Han et al.[67] substudy 1 and Johnson and Slovic[112] substudy 2) when uncertainty was shown. Others showed no difference for perceived probability (Kuhn et al.,[89] Han et al.[67] substudy 1 and 2, Lipkus et al.,[68] Sladakovic et al.[113]) or probability feeling (Johnson and Slovic[112] substudy 4, Gibson et al.,[55] Han et al.[67] substudy 2). We noted inconsistencies in findings among the 2 outcomes measures in Han et al.[67] finding 1 and Johnson and Slovic[112] finding 4, which creates additional questions regarding the strength of this evidence.

NARROW VERSUS WIDE CONFIDENCE INTERVAL: A moderate-credibility finding[114] compared narrow versus wide confidence interval to a point estimate. Perceived probability was higher only with the wide confidence interval. Probability feelings were not assessed.

Table 5G Evidence-Based Guidance for Stating or Illustrating Numerical Uncertainty on Probability Perceptions and Feelings

Comparison	Evidence Strength	Applied Example of Evidence-Based Communication	General Guidance
Explanation of uncertainty v. no explanation	Weak ($n = 3$)	People may not feel any differently about a 7% chance with the addition of a statement such as, "but we cannot predict the future of any individual person."	Probability perceptions and feelings may be similar whether or not an explanation of uncertainty is provided with probability information.
Point estimates v. uncertainty	Insufficient (inconsistent findings; $n = 12$)	It is not clear whether replacing point estimates with uncertainty displayed numerically or graphically affects perceived probability or probability feelings.	
Narrow v. wide confidence interval	Insufficient (too few findings; $n = 1$)	It is not clear whether the width of the confidence intervals affects perceived probability.	
Verbal uncertainty v. no verbal uncertainty	Insufficient (too few findings; $n = 2$)	It is not clear whether verbal statements of uncertainty used to supplement numbers ("may occur" v. "will occur") affect perceived probability or probability feelings.	

VERBAL UNCERTAINTY: One moderate-credibility finding (Knapp et al.[84] substudy 1) showed no difference in perceived probability or probability feelings by whether side effects "will occur" or "may occur." However, in this finding, the number was a probability band ("up to 1 in 10") which may have given the impression of uncertainty as well.

Comparisons of larger and smaller denominators on probability perceptions and feelings (subsection 5H). DENOMINATOR MANIPULATIONS: Of 5 moderate- to high-quality findings comparing several rates with different numerators/denominators, 1 finding[30] showed a difference in perceived probability, while the remaining 4 did not (Pighin et al.[9] substudies 4 and 5, Raghubir[96] substudy 2, Zikmund-Fisher et al.[23]). A high-credibility finding by Zikmund-Fisher et al.[21] had mixed findings, suggested that icon arrays with larger denominators increased worry for 2 rare side effects but not for 2 more common ones. Conversely, a classic pair of moderate-to-high-quality findings (Yamagishi[115] substudies 1 and 2) suggests that probability feeling is influenced more by an increasing numerator than by changes to the denominator ("denominator neglect"). However, another moderate-credibility finding (Galesic et al.[60] substudy 2) demonstrated no differences in perceived seriousness of a probability with changing denominators, although small sample size may have limited ability to pick up an effect.

Comparisons of animation or interactivity on probability perceptions and feelings (subsection 5I). We defined "interaction" as an action such as clicking on or manipulating a graphic, inputting information into a form, drawing, or responding to a question; the opposite of interacting with a visual was passively viewing it. We defined "animation" as a visual that moved or changed, as opposed to a static visual.

INTERACTION WITH GRAPHICS: Three trials of different sorts of interactive graphics showed no impact on probability perceptions and feelings (Han et al. [high credibility],[52] Fraenkel et al.,[62] and Ancker et al.[116] [moderate credibility]). A moderate-credibility finding from Emmons et al.[56] also showed no effect on concern about cancer when interactivity (ability to toggle risk factors to change risk level) was added to vertical risk ladders.

PERSONALIZATION AND INTERACTIVITY: A high-credibility finding, Witteman et al.[117] showed that adding a personalized avatar to an animated icon array slightly increased probability feelings.

ABILITY TO PERSONALIZE PROBABILITY ASSESSMENT: One high-credibility finding[118] showed no impact on perceived probability depending on whether or not the participant could interact with the risk assessment instrument.

INTERACTION WITH INFORMATION OUTPUT: Four moderate-credibility findings of cognitive activities, assessing both probability perceptions and feelings, had mixed results: drawing a bar chart reduced perceived probability, but calculating the answer to a question about the

Table 5H Evidence-Based Guidance for Manipulating Denominators on Probability Perceptions and Feelings

Comparison	Evidence Strength	General Guidance
Denominator manipulation	Insufficient (inconsistent findings; $n = 9$)	It is unclear whether changing the denominator of a probability affects probability perceptions and feelings.

Table 5I Evidence-Based Guidance for Effects of Animation or Interactivity on Probability Perceptions and Feelings

Comparison	Evidence Strength	Applied Example of Evidence-Based Communication	General Guidance
Interaction with graphics v. no interaction	Moderate ($n = 7$)	People may have similar perceived probability regardless of whether they view an icon array with 4 colored icons out of 100 or interact with an icon array to move the icons.	Perceived probability or probability feelings are not affected by whether graphics are static or interactive.
Personalization and interactivity v. none	Weak ($n = 1$)	People may have higher probability feelings with an animated icon array when given the opportunity to personalize one of the icons as an avatar.	Including a personalized avatar in an animated icon array may increase probability feelings.
Ability to personalize probability assessment v. none	Weak ($n = 1$)	People may have similar probability perceptions and feelings regardless of whether they are given a personalized probability or whether they complete a risk assessment to produce the personalized probability.	Inviting readers to interact with the information input may not affect perceived probability or probability feelings.
Interaction with information output v. viewing only	Insufficient (inconsistent findings; $n = 4$)	It is not clear whether inviting readers to perform an interactive cognitive task with the output of the information (e.g., perform a calculation, draw a chart) affects perceived probability or probability feelings.	
Animated slide show v. static graphics	Insufficient (too few findings; $n = 2$)	It is not clear whether a static graphic, an animated spinner, or an animated video slide show results in higher perceived probability or probability feelings.	

probability did not (Natter and Berry[119] substudies 1 and 2).

ANIMATED SLIDE SHOW: In addition to the interactive graphic described above, Fraenkel et al.[62] (moderate credibility) also compared a static icon array with an animated but noninteractive slide show of people affected/not affected by the condition, finding no effect on perceived probability or concern.

NOT SUMMARIZED: A lower-credibility finding from Weinstein et al.[41] substudy 2 is not summarized due to lack of clarity about stimuli and outcomes.

Comparisons of shorter versus longer time periods on probability perceptions and feelings (subsection 5J). LONGER VERSUS SHORTER TIME PERIODS: One high-credibility finding[46] showed much higher perceived probability with chances presented as lifetime probability rather than 10 y. However, the fact that the lifetime probability number was higher than the 10-y probability number may have accounted for the finding.

Effects of Different Formats on Health Behaviors and Behavioral Intentions (Health Behavior/Behavioral Intention Outcome): Section 7

Findings of health behavior outcomes were relatively few, limiting our ability to draw strong conclusions about the effects of data presentation formats on health behaviors. Instead, we present health behavior findings

Table 5J Evidence-Based Guidance for Varying the Time Period on Probability Perceptions and Feelings

Comparison	Evidence Strength	Applied Example of Evidence-Based Communication	General Guidance
Longer v. shorter time periods	Weak ($n = 1$)	People may perceive the chance of heart disease as higher if they are told there is a 10% lifetime probability than if they are told there is a 5-y probability of 1%.	Presenting lifetime probabilities rather than short-term probabilities may increase perceived probability of the event.

Table 7A Evidence-Based Guidance for Effects of Probability Formats on Health Behavior and Behavioral Intention

Comparison	Evidence Strength	Applied Example of Evidence-Based Communication	General Guidance
1 in X v. other number formats	Moderate ($n = 3$)	People may be more likely to act when told that they have a 1 in 20 chance than when they are told they have a 5% chance or a 5 in 100 chance.	Probabilities in 1-in-X format have a stronger impact on behavioral intention than probabilities as percentage or rate per multiple of 100.
Heart age v. probability	Weak ($n = 2$)	People may be more likely to act when told that their "heart age" is 10 y older than their chronological age than when told that they have a 5% chance of heart disease.	Probabilities presented as "heart age" may have a stronger impact on behavioral intention than probabilities as percentage.
Round numbers v. nonround numbers	Weak ($n = 3$)	A person may be more likely to get vaccinated when told the chance of flu is 40.00% than when told it is 40.21%.	Percentages shown as round numbers may have a stronger effect on behavior than nonround numbers, when both are presented to the same degree of precision.
Rate per 10^n v. percentage	Insufficient (inconsistent findings; $n = 2$)	It is not clear whether presenting probabilities as rate per multiple of 100 v. percentage affects behavioral intention.	
Adding time interval to a probability	Insufficient (too few findings; $n = 1$)	It is not clear whether adding a lengthy time interval (e.g., "1 case every X years") to a 1-in-X format affects behavioral intention.	

grouped with behavioral intention findings. Most behaviors and behavioral intentions were assessed by self-report. As mentioned in the Methods section, face and criterion validity of measures contributed to the credibility assessments.

Comparisons between probability formats in effects on health behavior and behavioral intention (subsection 7A). 1 IN X: Three findings compared 1 in X to other formats. Two high-credibility findings[6,10] demonstrated that behavioral intention was more strongly affected by 1 in X than by rate per 10^n or percentages. One moderate-credibility finding (Sirota et al.[8] substudy 5) demonstrated no difference between 1 in X and other formats.

HEART AGE: Two findings compared "heart age" to percentage probability; 1 (high credibility, Lopez-Gonzalez[120]) finding a stronger effect from heart age and 1 (moderate credibility, Soureti et al.[31]) finding no difference. Applying this finding to practice may be challenging in the absence of a universally accepted Heart Age calculator.

ROUND NUMBERS: The effect of round versus nonround numbers was assessed in several substudies by Wadhwa and Zhang.[121] Both "round" numbers and unrounded numbers were presented to same number of significant digits (e.g., 3.00 and 3.14). Round numbers had a stronger effect on self-reported behavior in 2 moderate-credibility findings but not in a third that assessed effect on behavioral intention.

Table 7B Evidence-Based Guidance for Effects of Graphical Formats on Health Behavior and Behavioral Intention

Comparison	Evidence Strength	Applied Example of Evidence-Based Communication	General Guidance
Icon arrays v. other graphics	Moderate ($n = 3$)	People's intention to take health-related actions may not be affected by whether they see chances portrayed a bar charts, icon arrays, or other sorts of graphics.	Different graphic types tested to date do not appear to have different effects on behavioral intention, but this finding may not apply to all possible graphics.
Random v. grouped icon array	Weak ($n = 1$)	People, when viewing an icon array showing 4 colored icons and 96 white ones, may be more likely to take action if the colored icons are grouped than if they are scattered.	Grouped icon arrays may lead to stronger effects on behavioral intention than randomly arranged icon arrays.

Table 7C Evidence-Based Guidance for Comparisons of Numerical and Graphical Formats, and Combinations of Numerical and Graphical Formats, on Health Behavior and Behavioral Intention

Comparison	Evidence strength	General Guidance
Numbers v. graphics	Insufficient (inconsistent findings; $n = 6$)	It is not clear whether numbers or graphics on average have a stronger impact on behavioral intention

RATE PER 10^n: One moderate-credibility finding[122] demonstrated no differences in behavior between percentage and rate per 10^n. Another moderate-credibility finding[123] showed that intention to take a drug was influenced more strongly by rate per 10^n plus verbal probability than by rate per 10^n alone, percentages alone, or percentages plus verbal probability.

ADDING TIME INTERVAL: A moderate-credibility[34] finding showed that adding a very long time interval ("1 case every 3,500 y") to explain a 1-in-X probability reduced behavioral intention (but adding a shorter time interval had no effect).

NOT SUMMARIZED: Findings were from 2 very small studies that may have been underpowered.[124,125] Three additional findings by Wadhwa and Zhang[121] were not summarized because the small sample size reduced confidence in findings of no difference.

Comparisons between graphical formats in effects on health behavior and behavioral intention (subsection 7B). ICON ARRAYS VERSUS OTHER GRAPHICS: A high-credibility finding (Weinstein et al.[41] substudy 1) showed no impact on intentions by whether the graphic added to a frequency number (either 1 in X or rate per 10^n) was a denominator-only icon array or part-to-whole/numerator-denominator icon array. A high-credibility finding[62] demonstrated no behavioral intention differences by graphic format (static icon array, interactive "spinner" graphic, and sequence of slide images of people who experienced the different outcomes). Another moderate-credibility finding[58] showed no consistent differences between an icon array and a vertical bar chart as part of an infographic. (The effect of the animated graphics in Fraenkel are discussed in the animation/interactivity section below.)

RANDOM VERSUS GROUPED ICON ARRAY: A large high-credibility finding[117] compared several sorts of static and animated icon arrays, showing that intention to change behavior was higher when the array had grouped icons rather than randomly arranged ones.

NOT SUMMARIZED: A lower credibility finding[126] was not summarized due to moderate sample size and additional experimenter activities in one of the graphics conditions.

Comparisons between numerical and graphical formats, and combinations of numerical and graphical formats, in effects on health behavior and behavioral intention (subsection 7C). NUMBERS VERSUS GRAPHICS: Six findings in which numbers could be directly compared with graphics had inconsistent results. One high-credibility and 3 moderate-credibility findings showed no difference

Table 7D Evidence-Based Guidance for Comparisons of Numerical and Verbal Probabilities on Health Behavior and Behavioral Intention

Comparison	Evidence Strength	Applied Example of Evidence-Based Communication	General Guidance
Numbers v. verbal probabilities	Strong (n = 10)	People will be less likely to take a drug if its side effects are described as "rare" than if they are described as affecting 0.1% of people.	Verbal probabilities have a greater impact on behavioral intentions than numbers alone do.

Table 7E Evidence-Based Guidance for Context on Health Behavior and Behavioral Intention

Comparison	Evidence Strength	Applied Example of Evidence-Based Communication	General Guidance
Anecdotes v. no anecdotes, or different numbers of anecdotes	Moderate (n = 4)	People may be more likely to act on a 2 in 100 chance of a headache as a side effect if they are also told an anecdote about a man named Tom who developed a headache after taking the drug.	Adding anecdotes about positive or negative outcomes to a probability number influences behavioral intention, with the direction depending on whether the anecdote is positive or negative.
Population average v. no population average	Moderate (n = 3)	People may be no more likely to act on their 2 in 100 chance of heart disease if they are also told that the average chance for people of their age is 1 in 100.	Adding the population average probability to communication about individual probability probabilities may not alter behavioral intention.
Comparison events v. no comparison events	Weak (n = 1)	People may be no more likely to act on their 2 in 100 chance of heart disease if they are told that their chance of developing lung cancer is 4 in 100.	Adding chance of comparator events to numerical or graphical representations of probabilities does not appear to result in differences in behavioral intention.
Years of life lost v. probability	Insufficient (too few findings; n = 2)	It is not clear whether expressing a probability as number of years of life likely to be lost affects behavioral intention more than a standard percentage probability does.	

Table 7F Evidence-Based Guidance for Framing on Health Behavior and Behavioral Intention

Comparison	Evidence Strength	Applied Example of Evidence-Based Communication	General Guidance
Gain v. loss framing of chance of benefit	Strong (n = 3)	People are more likely to choose a therapy if told that 90% of patients will survive with it than if they are told that 10% will die.	Positively framing chance of benefit, as compared to negatively framing it, increases intention to take an action.
Gain v. loss framing of chance of harm	Strong (n = 3)	People are more likely to avoid a therapy if they are told that 10% of people will experience side effects than if told that 90% of people will not experience them.	Negatively framing probability of side effects, as compared with positively framing it, reduces intention to take a therapy.
Gain-loss framing: behavior	Insufficient (n = 1)	It is not clear whether gain-loss framing of individual probabilities affects behaviors as well as intentions.	

Table 7G Evidence-Based Guidance for Uncertainty on Health Behavior and Behavioral Intention

Comparison	Evidence Strength	General Guidance
Uncertainty displayed v. not displayed	Insufficient (too few findings; $n = 1$)	It is unclear whether displaying uncertainty in probability graphics affects behavioral intentions.

Table 7I Evidence-Based Guidance on Effect of Animation or Interactivity on Health Behavior and Behavioral Intention

Comparison	Evidence Strength	General Guidance
Interactivity v. view-only images	Insufficient (too few findings; $n = 4$)	It is not clear whether adding different types of interactivity to probability information affects behavioral intentions.
Animation v. static graphics	Insufficient (too few findings; $n = 1$)	It is not clear whether adding different types of animation to probability information affects behavioral intentions.

Table 8A Evidence-Based Guidance on Effect of Numerical Probability Formats on Trust

Comparison	Evidence Strength	Applied Example of Evidence-Based Communication	General Guidance
More v. fewer significant digits	Weak ($n = 1$)	People may find probability information more credible when described as a 1% chance than as a 1.12% chance.	Percentages with 0 or 1 decimal points may be more trustworthy than percentages with 2 or more decimals.
Percentages v. other formats	Weak ($n = 2$)	People's trust in information about their cardiovascular disease probability may be the same whether it is described as a 5% chance, a 1 in 20 chance, or a heart age of 5 y older than their chronological age.	Different numeric formats (percentage, heart age, 1 in X) may not affect trust in the information.

between numerical and graphical formats (Weinstein et al.[41] substudy 1, Gibson et al.,[55] Henneman et al.,[71] Damman et al.[58]). However, 1 moderate-credibility finding[122] demonstrated that self-reported condom use was higher when the probability of disease was presented as a part-to-whole icon array than when it was presented as a number (either percentage or rate per 10^n). A high-credibility finding[62] demonstrated that in communicating probabilities of all the different outcomes of surgery, intention to get the surgery was higher with formats that included visuals than with rate per 10^n alone.

NOT SUMMARIZED: A lower-credibility finding[126] was not summarized due to moderate sample size and additional experimenter activities in a graphics condition.

Comparisons between numerical and verbal probabilities in effects on health behavior and behavioral intention (subsection 7D). NUMBERS VERSUS VERBAL PROBABILITIES: In comparing the effect of verbal and numerical probabilities on behavioral intention, 8 of 10 findings showed that verbal probabilities had stronger impact. Studies showing an impact were 3 high-credibility findings (Berry et al.[127] substudy 2, Berry et al.,[78] Barry et al.[73] substudy 1) and 5 moderate-credibility findings (Young and Oppenheimer substudy 3,[128] Young and Oppenheimer[76] substudy 3, Sinayev et al.,[123] Peters et al.,[75] Berry and Hochhauser[74]). However, 2 findings showed no difference (a high-credibility finding by Dahlstrom et al.,[81] a moderate-credibility finding by de Wit et al.[86]). However, several of these findings assessed the mapping between probabilities and specific verbal probability terms established by the EC to describe relatively small chances of drug side effects; for these, it is unclear whether other verbal terms, higher-probability events, or other verbal-numeric mappings would produce different results.

Table 8B Evidence-Based Guidance on Effect of Graphical Formats on Trust

Comparison	Evidence Strength	General Guidance
Icon array design features	Insufficient (too few findings; $n = 2$)	It is not clear whether specific icon array features (icon arrangement, number of icons, addition of rate per 10^n or verbal probabilities) affect trustworthiness.

Table 8C Evidence-Based Guidance on Comparisons between Numerical and Graphical Formats or Combinations of Numerical and Graphical Formats on Credibility

Comparison	Evidence Strength	General Guidance
Numbers v. graphics	Insufficient (too few findings; $n = 3$)	It is not clear whether numbers or graphics are trusted more.

Table 8E Evidence-Based Guidance on Effect of Contextual Information on Trust

Comparison	Evidence Strength	General Guidance
Elements added for context v. none	Insufficient (too few findings; $n = 2$)	It is not clear whether contextual elements, specifically population average or interpretive categories, affect the perceived credibility of information.

Table 8F Evidence-Based Guidance on Effects of Framing on Trust

Comparison	Evidence Strength	General Guidance
Gain-loss framing	Insufficient (too few findings; $n = 1$)	It is not clear whether trust is affected by gain or loss framing of a probability.

Table 8G Evidence-Based Guidance on Effect of Uncertainty on Trust

Comparison	Evidence Strength	General Guidance
Uncertainty displayed v. not displayed	Insufficient (too few findings; $n = 5$)	It is not clear whether showing confidence intervals or other forms of uncertainty affects perceived credibility of information.

NOT SUMMARIZED: Three low-credibility findings (Young and Oppenheimer[128] substudy 2, Young and Oppenheimer[76] substudy 2, Cheung et al.[125]) were not summarized due to small sample sizes.

Comparisons of elements added for context on health behavior and behavioral intention (subsection 7E). Four different types of contextual information were studied: anecdotes, population average, comparison events, years of life lost.

ANECDOTES: Three moderate-credibility findings examine the effect of adding narrative anecdotes to numerical information. Two showed that behavioral intention was strongly affected by the anecdotes (Betsch et al.,[88] substudies 1 and 2) and one finding[87] did not. The one that did not focused on intention to try recreational drugs, which may have been too strongly

Table 9A Evidence-Based Guidance on Preferences for Numerical Probability Formats

Comparison	Evidence Strength	General Guidance
Different number formats for probability	Insufficient (inconsistent findings; $n = 4$)	It is not clear whether people consistently prefer any particular probability number format (X in N, 1 in X, percentage, or combinations of these).

Table 9B Evidence-Based Guidance on Preferences for Graphical Probability Formats

Comparison	Evidence Strength	Applied Example of Evidence-Based Communication	General Guidance
Numerator-only v. part-to-whole icon arrays	Weak ($n = 1$)	People may like an icon array more if it shows 4 colored icons and 96 white ones rather than just showing the 4 colored ones.	Part-to-whole icon array displays may be preferred to numerator-only icon arrays.
Icon shape in icon arrays	Insufficient (inconsistent findings: $n = 2$)	It is not clear whether people prefer abstract icons (such as blocks) or human-shaped icons (such as stick figures) in icon arrays.	
Grouped v. random icon arrays	Insufficient (inconsistent findings: $n = 2$)	It is not clear whether people prefer icons in icon arrays to be randomly arranged or grouped.	
Types of graphics compared with each other	Insufficient (inconsistent findings; $n = 5$)	It is not clear whether a particular probability graphic (icon array, bar chart, number line) is generally preferred to others.	

influenced by preexisting attitudes toward drug use. A low-to-moderate-credibility finding[86] demonstrated no difference between anecdotes (alone or with verbal probability) and percentage with verbal probability.

POPULATION AVERAGE: Three moderate-credibility findings examined the effect of adding a population average or typical probability for context, and none showed that this addition affected behavioral intention.[6,93,95]

COMPARISON EVENTS: In a moderate-credibility finding examining the effect of adding a different probability for comparison (e.g., chance of other cancers, chance of injury and construction accident), Gibson et al.[55] did not find an effect.

YEARS OF LIFE LOST: One high-credibility finding[120] showed that showing heart age + numbers of years lost + additional explanation had a larger impact on behavioral intention than percentage probability. It is not clear whether the effect is due to the heart age, the explanation, or the number of years lost. A moderate-credibility finding[58] compared heart age plus years of life lost to standard probability percentage, but the contextual element (years of life lost) was added only in the heart age condition, making it unclear whether the effect was due to the heart age or the years of life lost.

NOT SUMMARIZED: A small study may have been underpowered to detect a difference but was generally consistent with the summaries above for population average and comparison events.[129]

Comparisons of frames (gain, loss, or comparison) on health behavior and behavioral intention (subsection 7F). GAIN VERSUS LOSS FRAMING OF CHANCE OF BENEFIT—BEHAVIORAL INTENTION: Two moderate- and high-credibility findings from Levin et al.[111] and Garcia-Retamero and Cokely[122] showed that gain framing the chance of benefit increased likelihood of choosing an option (behavioral intention). However, Bigman et al.[107] demonstrated no effect of framing success rates of a vaccine on intention to take the vaccine (but the same study did find a significant effect on perceived probability).

GAIN VERSUS LOSS FRAMING OF CHANCE OF BENEFIT—BEHAVIOR: An additional finding from one of these studies also showed a similar framing effect of the benefit on self-reported behavior.[122]

GAIN VERSUS LOSS FRAMING OF CHANCE OF HARM—BEHAVIORAL INTENTION: Moderate- and high-credibility findings from Gurm and Litaker[130]

Table 9C Evidence-Based Guidance on Preferences for Numerical versus Graphical Formats

Comparison	Evidence Strength	General Guidance
Numbers v. graphics	Insufficient (inconsistent findings; $n = 6$)	It is not clear whether numbers or graphics are generally preferred for communicating probabilities.

Table 9D Evidence-Based Guidance on Preferences for Numerical versus Verbal Probability Formats

Comparison	Evidence Strength	Applied Example of Evidence-Based Communication	Comparison
Numerical v. verbal probability	Moderate ($n = 13$)	People are likely to prefer to have a side effect labeled as having a 7% chance instead of described as "common."	Numbers or numbers combined with verbal probabilities are preferred to words alone to describe probability

Donovan and Jalleh,[131] Taylor et al.,[132] and Wilson and Kaplan[133] substudy 2 showed that loss framing the chance of harm from an option reduces intention to select the option.

NOT SUMMARIZED: A lower-credibility finding[134] is not summarized due to small sample size.

Comparisons of methods of representing uncertainty on health behavior and behavioral intention (subsection 7G). One moderate-credibility finding[55] did not find differences in behavioral intention between graphics that demonstrated uncertainty and graphics that did not show uncertainty.

Comparisons of animation or interactivity on health behavior and behavioral intention (subsection 7I). INTERACTIVITY: Four moderate- to high-credibility findings examined the effect of adding some form of interactivity, all showing no effect on behavioral intention (Harle et al.,[118] Natter and Berry[119] substudies 1 and 2, Witteman et al.[117]). However, the interactive features were extremely heterogeneous, reducing ability to compare and draw conclusions. The interactivity included answering questions for a personalized risk screening,[118] drawing a bar chart or computing an equivalent probability (Natter and Berry[119] substudies 1 and 2), and personalizing an avatar with a color.[117]

ANIMATION: The Witteman finding[117] additionally examined several forms of animation applied to icon array graphics including whether a personal avatar was displayed and whether the avatar moved, with no effect on behavioral intention. No other findings looked at behavioral impact of other animated graphics.

Within the health behavior and behavioral intention outcome, no relevant findings examined effect of manipulating denominators of probabilities or effect of varying the time period.

Effects of Different Formats on Trust in the Message (Trust Outcome): Section 8

Many researchers assessed trust in the information, credibility of the information, or related constructs. We summarize findings on the effects of different formats on trust.

Comparisons between numerical probability formats in effects on trust (subsection 8A). SIGNIFICANT DIGITS: In a high-credibility finding, Witteman et al.[29] showed that trustworthiness was highest when a percentage had 0 or 1 digits to the right of the decimal and decreased with more digits.

PERCENTAGES VERSUS OTHER FORMATS: In high-credibility findings, Soureti et al.[31] showed no trust difference between percentage probability and heart age, and Gurmankin et al.[14] showed no difference in trust in information presented as percentage or as 1 in X.

Comparisons between graphical formats in effects on trust (subsection 8B). Only 2 findings, both moderate credibility, examined effects of different graphic types on trust. Both assessed icon arrays.

ICON ARRAY DESIGN FEATURES: Schapira et al.[42] showed that "perceived truth" of an icon array with probability of breast cancer was higher for randomly arranged icons than for grouped ones and for

Table 9E Evidence-Based Guidance on Preferences for Added Contextual Information

Comparison	Evidence Strength	General Guidance
Elements added for context	Insufficient (too few findings; $n = 5$)	It is not clear whether different kinds of information provided as context (e.g., probabilities of comparison events, lifetime probability provided as context for 10-y probability) are preferred.

Table 9F Evidence-Based Guidance on Preferences for Framing Probabilities

Comparison	Evidence Strength	Applied Example of Evidence-Based Communication	General Guidance
Gain-loss framing	Moderate ($n = 4$)	People may not care whether a probability is described as a 90% chance of success or a 10% chance of failure.	People do not have a preference for whether probabilities are gain framed or loss framed.

Table 9G Evidence-Based Guidance on Preferences about Uncertainty

Comparison	Evidence Strength	General Guidance
Uncertainty displayed v. not displayed	Insufficient (too few findings; $n = 1$)	It is not clear whether presenting uncertainty influences preference for information.
Numerical v. verbal portrayals of uncertainty	Insufficient (too few findings; $n = 1$)	It is not clear whether presenting uncertainty as numerical ranges or verbal descriptors is preferred.

icon arrays with smaller number of icons. For both of these, an interaction with numeracy meant that the finding occurred only among those low in numeracy.

Brewer et al.[135] showed that trust in the information was lower for an icon array paired with rate per 10^n and verbal probability than for the original Oncotype Dx report and for several custom-designed alternative graphics.

Comparisons between numerical and graphical formats, and combinations of numerical and graphical formats, in effects on trust (subsection 8C). Only 3 moderate-credibility findings have compared numbers versus graphics in effect on trust. Both Lipkus et al.[68] and Han et al.[67] substudy 1 showed no trust differences between information in numbers or in graphics. Brewer et al.[135] showed trust to be lower when cancer recurrence probability was presented as an icon array plus rate per 10^n and verbal probability than with several other types of combined graphics such as the Oncotype Dx report or a number line.

Comparisons of elements added for context on trust (subsection 8E). One moderate-credibility finding (Han et al.[67] substudy 1) compared graphic depictions of probability with and without a reference (population average) value, finding no difference in trust. A second moderate-credibility finding[135] compared graphics with and without interpretive categories. However, so many factors were altered in the comparison that differences cannot be attributable to this factor.

Comparisons of frames (gain, loss, combination) on trust (subsection 8F). One moderate-credibility finding, Webster et al.,[103] showed that frame did not affect credibility, but small sample size and confounding of numerical frame with verbal label reduce confidence in negative finding.

Comparisons of methods of representing uncertainty on trust (subsection 8G). The effect of showing uncertainty on perceived credibility or trust has been studied in 5 moderate-credibility findings. Longman et al.[114] showed

Table 9I Evidence-Based Guidance on Preferences about Animation and Interactivity

Comparison	Evidence Strength	General Guidance
Interactivity or animation v. static or view only	Insufficient (too few findings; $n = 3$)	It is not clear whether people have preferences for specific types of animation or types of interactivity with the information.

Table 9J Evidence-Based Guidance on Preferences about Time Period

Comparison	Evidence Strength	General Guidance
Shorter v. longer time periods	Insufficient (too few findings; $n = 1$)	Whether people prefer to receive 10-y or lifetime probability versus both time frames.

Table 10B Evidence-Based Guidance on Effects of Graphical Formats on Discrimination

Comparison	Evidence Strength	General Guidance
Icon arrays v. risk ladders	Insufficient (too few findings; $n = 1$)	It is not clear whether icon arrays or risk ladders affect discrimination between different probabilities.

Table 10C Evidence-Based Guidance on Effects of Numerical and Graphical Formats on Discrimination

Comparison	Evidence Strength	Applied Example of Evidence-Based Communication	General Guidance
Numbers v. graphics	Weak ($n = 3$)	Ability to distinguish between similar probabilities may not be affected by whether the probabilities are shown as (for example) 1% and 2% alone or whether the percentages are supplemented by risk ladders.	There may be no overall difference between graphics and numbers in helping people be sensitive to differences in probabilities.

credibility was lower when estimates were accompanied by confidence intervals but only when the confidence intervals were wide, not narrow.

Both Han et al.[67] substudy 1 and Lipkus et al.[68] showed that adding confidence intervals did not affect credibility, but the Han finding altered multiple factors between stimuli, making it difficult to attribute differences to the confidence intervals. Han et al.[67] substudy 2 showed no differences in perceived credibility on the basis of how the uncertainty was shown (numerically or graphically).

Brewer et al.[135] showed that several information formats that did not include confidence intervals were perceived as less trustworthy than the original Oncotype report that included a confidence interval. However, the effect was confounded by a number of different format differences between the versions.

Within the trust outcome, no relevant findings examined the effects of manipulating probability denominators, animation or interactivity, or varying the time period. No findings met our criteria for comparisons of verbal and numerical probabilities.

Preferences about Formats (Preference Outcome): Section 9

It was common for researchers to assess how much participants liked or preferred certain formats, using measures such as ranking or choosing preferred formats. All

Table 10D Evidence-Based Guidance on Effects of Numerical and Verbal Probabilities on Discrimination

Comparison	Evidence Strength	Applied Example of Evidence-Based Communication	General Guidance
Numbers v. verbal	Weak (*n* = 1)	People may distinguish between probabilities better when they are described as "low, 5 per 1,000 and 4 per 1,000" than when they are both just described as "low."	Adding rate per 10^n to a verbal probability may increase people's sensitivity to probability differences.

Table 10E Evidence-Based Guidance on Effects of Contextual Information on Discrimination

Comparison	Evidence Strength	Applied Example of Evidence-Based Communication	General Guidance
Interpretive labels v. no labels	Weak (*n* = 2)	It may be easier for people to make distinctions between a 4% chance and a 7% chance if they are labeled as "low" and "moderate" chances.	Adding interpretive labels to probabilities may help people distinguish between probability levels.
Chances of comparison events v. no comparison events	Insufficient (inconsistent findings; *n* = 2)	It is not clear whether displaying chances of comparison events helps people distinguish between probability levels.	
Population average v. no population average	Insufficient (inconsistent findings; *n* = 2)	It is not clear whether showing a population average value helps people distinguish between probability levels.	

Table 10I Evidence-Based Guidance on Effects of Interactivity on Discrimination

Comparison	Evidence Strength	General Guidance
Prompting to count v. no prompting	Insufficient (too few findings; *n* = 1)	It is not clear whether prompting people to count with a probability graphic affects people's discrimination between different probabilities.

relationships between formats and preferences are summarized in this section.

Preferences for different numerical probability formats (subsection 9A). PERCENTAGES: Both a high-credibility finding[136] and a moderate-credibility finding[137] suggested percentage was preferred to 1 in X, but another high-credibility finding from Nagle et al.[138] showed the opposite. A high-credibility finding[15] showed percentage combined with 1 in X was preferred to either alone.

NUMBERS PLUS VERBAL: Two findings lend support to combinations of verbal interpretation or verbal probability and numerical information. A high-credibility finding[139] suggested people had more comfort with numbers combined with a verbal interpretation ("some chance of") than with the verbal description alone. A moderate-credibility finding[140] also showed that self-reported understanding was higher with percentage plus verbal probability rather than with rate per 10^n or percentage alone. However, Steiner et al.[141] (a moderate-credibility finding) showed that verbal labels alone were considered easier to understand than percentages or percentages plus verbal labels. (This evidence is summarized further in the numbers versus verbal category, so we have not added it to the evidence table below.)

NOT SUMMARIZED: Relevant findings produced several lower-credibility findings not included in the synthesis due to small sample size and/or lack of hypothesis testing (Lobb et al.,[142] Hamilton et al.,[57] Cheung et al.,[125] Knapp et al.,[35] Hovick et al.[143] substudy 1, Strathie et al.[36]) or insufficient information provided to participants.[144]

Preferences for different graphical probability formats (subsection 9B). Evidence about preference for different types of graphics is limited by the fact that the findings in this category have assessed very different types of graphics.

NUMERATOR-ONLY VERSUS PART-TO-WHOLE ICON ARRAYS: In a high-credibility finding, Okan et al.[39] substudy 1 showed part-to-whole icon arrays preferred over numerator-only icon arrays.

ICON SHAPE: Two moderate-credibility findings on icon shape were inconsistent: Kreuzmair et al.[43] showed no preference difference between block icons and human figure icons, but Schapira et al.[42] showed a preference for human figure ones.

GROUPED VERSUS RANDOM ICON ARRAYS: Ancker et al.[116] (moderate credibility) showed no preference difference between randomly arranged and grouped icons, but Schapira et al.[42] (moderate credibility) showed a preference for grouped over random ones.

TYPES OF GRAPHICS: Four moderate-credibility findings compared icon arrays to other graphics formats. Ghosh et al.[145] showed no overall preference difference between icon arrays and bar charts, but Edmonds et al.[146] showed that a bar chart was preferred to an icon array. Etnel et al.[147] reported that pie charts were preferred to icon arrays and/or bar charts. Brewer et al.[135] showed that an icon array was one of the multiple formats rated as more understandable than the Oncotype Dx graphic display.

Three studies examined number lines. A moderate-credibility finding[146] showed a preference for bar charts over number lines. Another moderate-credibility finding[135] showed that a number line was one of the formats rated more understandable than the Oncotype Dx report. A high-credibility finding[148] of a number line/risk ladder showed that the number lines that included categories shown in rate per 10^n were rated as more understandable and actionable than number lines with verbal labels alone.

NOT SUMMARIZED: Lower-credibility findings not synthesized here due to small sample size and/or lack of statistical testing (Clarke et al.,[149] Hamilton et al.,[57] Hovick et al.[143] substudy 1, Housseini et al.[150] substudy 1); complicated, multidimensional stimuli[58]; additional experimenter activities conflated with the display itself[126]; and contradictory findings between 2 measures of satisfaction.[56]

Preferences for numerical versus graphical formats (subsection 9C). Two studies examined icon arrays versus numbers. Siegrist et al.[54] substudy 1 showed that icon arrays were preferred to other graphics and number formats including 1 in X, rate per 10^n, and log scale number line. However, Gett et al.[151] showed that a pie chart + percentage was preferred to an icon array.

Siegrist et al.[54] substudy 2 (a high-credibility finding) showed no preference difference between 1 in X and a log-scale horizontal number line ("Paling").

A moderate-credibility finding, Brewer et al.,[140] showed numbers (with or without verbal probability) were preferred to pie charts.

A moderate-credibility finding, Brewer et al.,[135] showed that several alternate formats were all preferred to the original Oncotype Dx report.

A high-credibility finding, Kiely et al.,[152] showed a preference for a histogram graphic that showed best case, worst case, and typical cases over a numerical description of median survival, but it is unclear how much the graphic design contributed to this preference.

NOT SUMMARIZED: Some lower-credibility findings are not summarized due to small sample size and/or lack of statistical testing (Hamilton et al.,[57] Hovick et al.[143] substudy 1, Hagerty et al.,[153] Henneman et al.,[137] Hill et al.[154]); complicated, multidimensional stimuli[58]; or additional experimenter activities conflated with the display itself.[126]

Preferences for numerical versus verbal probability formats (subsection 9D). Thirteen moderate-to-high-credibility findings have elicited preferences for or satisfaction with numbers versus verbal probabilities.

NUMERICAL VERUS VERBAL PROBABILITY: Eight of these (4 high-credibility, 4 moderate) showed a preference for either numbers or numbers plus verbal labels over verbal probabilities alone. A caveat is that they used different sorts of numbers and verbal labels (Berry et al.[127] substudy 2, Berry et al.,[78] Knapp et al.,[12] Carey et al.,[155] Connolly et al.,[148] Freeman and Bass,[156] Nagle et al.,[138] Berry et al.[74]). Nagle et al.[138] also showed certain number formats (1 in X) preferred to verbal probabilities but not to other number formats (percentage alone or percentage + verbal combinations).

However, 3 moderate-credibility findings showed no difference in preference between numbers and verbal terms (Knapp et al.,[77] Knapp et al.[84] substudy 2, Damman et al.[58]). In addition, 1 high-credibility finding[136] and 1 moderate[141] showed a preference for verbal probabilities over numbers or numbers + verbal. (The latter finding was in a context in which multiple forms of birth control were being displayed simultaneously.)

NOT SUMMARIZED: Lower-credibility findings are not included in the synthesis due to small sample size and/or lack of statistical testing (Knapp et al.,[144] Lobb et al.,[142] Hallowell et al.,[157] Bloch et al.,[158] Cheung et al.,[125] Hovick et al.[143] substudy 1) or lack of clarity about outcome measures.[159]

Preferences for elements added for context (subsection 9E). Evidence for preferences about contextual elements is limited by the fact that the few findings in this category evaluated different sorts of context.

ELEMENTS FOR CONTEXT: Three moderate-to-high-credibility findings examined preferences for providing probabilities of other events as context/comparison for the probability of interest. In a high-credibility finding, Keller and Siegrist[59] showed no difference in preference between presenting chance of adverse outcomes associated with radon alone or pairing it with the equivalent probabilities from smoking. In a moderate-credibility finding, Siegrist et al.[54] substudy 1 showed no preference difference between a log-scale number line that provided chance of comparison events and other numerical and graphical formats that did not provide comparison event chances. In a moderate-credibility finding, Freeman and Bass[156] showed that single probabilities alone (numerical or verbal only) were preferred to expressing the probability as a verbal comparison probability ("as likely as being struck by lightning").

In a moderate-credibility finding, Henneman et al.[137] showed a preference for a combination of lifetime and 10-y probability versus either time interval separately.

In a moderate-credibility finding, Harris and Smith[95] showed no difference in perceived understandability by whether a personal probability was accompanied by the population average or by whether the personal probability was higher or lower than the average.

NOT SUMMARIZED: Three lower-credibility findings are not included here due small sample size (Gett et al.,[151] Hovick et al.[143] substudy 1) or complicated, multidimensional stimuli.[58]

Preferences for frames (gain, loss, or combination) (subsection 9F). GAIN VERSUS LOSS FRAMING: Four moderate- to high-credibility findings suggest that whether a probability is gain framed or loss framed does not strongly affect preference for the information or related constructs such as satisfaction and perceived usefulness (Biswas and Pechmann[110] finding 1, Garcia-Retamero and Cokely,[160] Garcia-Retamero and Cokely,[122] Webster et al.[103]). Most of these findings are small, somewhat reducing the confidence in negative findings.

NOT SUMMARIZED: Two lower-credibility findings are not summarized here due to small sample size.[132,161]

Preferences for methods of representing uncertainty (subsection 9G). DISPLAYING VERSUS NOT DISPLAYING UNCERTAINTY: A moderate-credibility finding[135] showed that a number of different formats without uncertainty were preferred to the Oncotype Dx report, which includes confidence intervals.

NUMERICAL VERSUS VERBAL UNCERTAINTY: A second moderate-credibility finding[148] showed that in showing probability of harm from eating fish on a vertical number line/risk ladder graphic, participants preferred a version in which the scale was shown in units of rate per 10^n plus uncertainty to a verbal only scale.

NOT SUMMARIZED: Two additional lower-credibility findings are not summarized here due to small sample size.[134,158]

Preferences for animation or interactivity (subsection 9I). INTERACTIVITY OR ANIMATION: The 3 available moderate-credibility findings have examined very different sorts of animation/interactivity. Natter and Berry[119] substudy 2 showed that undergraduates were more satisfied with information when they had to answer a reflective/computational question about it than when they did not. Natter and Berry[119] substudy 1 showed there were no differences in satisfaction when undergrads were given a bar chart or invited to draw one on the basis of information provided. Ancker et al.[116] showed no preference differences between interactive and static icon array graphics.

Preferences for shorter versus longer time periods (subsection 9J). SHORTER VERSUS LONGER TIME PERIODS: A moderate-credibility finding[137] showed a preference for presenting both lifetime probability and 10-y probability rather than either alone, but the general preference for more over less information reduces confidence in this finding.

Within the preference outcome, no relevant findings examined the effect of manipulating probability denominators.

Ability to Distinguish between Quantities (Discrimination Outcome): Section 10

Metrics that assessed differences in responses to different levels of probability on the basis of differences in the stimulus were termed *discrimination* outcomes. An assessment of whether respondents estimated 2 probabilities to be the same or different on the basis of presence or absence of a graphic, or design of the graphic, was considered a discrimination outcome.

Comparisons between graphical formats in effects on discrimination (subsection 10B). ICON ARRAYS VERSUS RISK LADDERS: A moderate-credibility finding[162] in a contingent valuation context showed greater discrimination between probability levels when numbers (1 in X and rate per 10^n) were accompanied by a large icon array than with either of 2 risk ladder formats.

NOT SUMMARIZED: A lower-credibility finding from Schonlau and Peters[163] finding 2 was not synthesized due to poor-quality stimuli and lack of clarity in reporting.

Comparisons between numerical and graphical formats in effects on discrimination (subsection 10C). NUMBERS VERSUS GRAPHICS: Two findings, 1 high credibility and 1 moderate, find no difference in direct comparison of numbers versus graphics for helping readers discriminate between probability levels.[59,164] A moderate-credibility finding[162] showed greater sensitivity to probability differences when numbers (1 in X and rate per 10^n) were accompanied by a large denominator icon array but showed no effect of 2 types of risk ladders.

NOT SUMMARIZED: A lower-credibility finding from Schonlau and Peters[163] finding 2 is not synthesized due to poor-quality stimuli and lack of clarity in reporting.

Comparisons between numerical and verbal probabilities in effects on discrimination (subsection 10D). A high-credibility finding[80] showed that discrimination between different probabilities was higher when probabilities were explained with verbal probabilities plus rate per 10^n than with verbal probabilities alone.

Comparisons of elements added for context on discrimination (subsection 10E). INTERPRETIVE LABELS: A high-credibility finding (Pighin et al.[165] study 2) showed that adding verbal interpretive labels to a number assisted in discrimination between probability levels. A moderate-credibility finding[166] showed that ability to discriminate between levels was not affected by slight differences in the labeling of the interpretive probability categories.

CHANCE OF COMPARISON EVENTS: Keller et al.[167] showed that showing a chance of comparison events increased discrimination among less numerate but not more numerate readers. Also, a moderate-credibility finding[162] showed no effect of adding risk ladders (containing comparison risk information) on sensitivity to probability variations, although the conflation of context with graphic type undermines confidence in this negative finding.

POPULATION AVERAGE: Two findings examined the effect of showing a population average value for comparison. In a high-credibility finding, Hess et al.[94] finding 1 showed that adding a second icon array with the population average did not improve discrimination and may have reduced it among low-numeracy respondents. A moderate-credibility finding by Keller and Siegrist[59] showed that a horizontal log-scale number line with population probability shown for comparison improved discrimination among high-numeracy respondents only but not low-numeracy ones; conflation of graphic type and the population value reduces confidence in finding.

Comparisons of animation or interactivity on discrimination (subsection 10I). Only 1 moderate-credibility finding (Hess et al.[94] substudy 3) examined the effect of interactivity on ability to discriminate between probabilities, finding that encouraging participants to count icons in an icon array did not make a difference.

Within the discrimination outcome, no relevant findings examined the impact of different numerical probability formats, framing manipulations, expressing uncertainty, manipulating the denominator, or changing the time period.

Summary of Evidence

Evidence is **moderate or strong** that probability perceptions and/or feelings are higher:

- with 1 in X than with rate per 10^n or percentage (subsection 5A: numerical formats comparison),
- with 1 in X alone than with 1 in X plus a part-to-whole icon array (subsection 5C: numerical and graphical format comparison),
- with foreground-only (numerator-only) icon arrays than with part-to-whole (numerator-denominator)

icon arrays (subsection 5B: graphical format comparison),

- with foreground-only (numerator-only) icon arrays than with rates per 10^n (subsection 5C: numerical and graphical format comparison),
- with a bar chart than with a part-to-whole icon array (subsection 5C: numerical and graphical format comparison),
- with rate per 10^n or percentage than with part-to-whole icon arrays (subsection 5C: numerical and graphical format comparison),
- With log-scale number lines (that displays risks of comparison events) than with 1 in X (subsection 5C: numerical and graphical format comparison),
- with verbal probabilities (such as “rare”) than with the corresponding probability numbers designated by the EC (subsection 5C: numerical and graphical format comparison),
- with numbers supplemented by anecdotes about people who have experienced the event than with numbers alone (subsection 5E: contextual information comparison), and
- when negative events are loss framed or positive events are gain framed (subsection 5F: framing comparison).

There is also **strong** evidence that explicitly illustrating the size of the effect of a treatment or risk factor on the chance of a health outcome will alter risk perceptions and feelings about the likelihood of the health outcome itself, although the direction of the effect may differ by context (subsection 5B: graphical format comparison).

Furthermore, evidence is **moderate to strong** that the following modifications do **not** affect probability perceptions and/or feelings:

- replacing a percentage with a rate per 10^n (subsection 5A: numerical formats comparison),
- replacing human-shaped icons in an icon array with abstract icons (subsection 5C: graphical formats comparison),
- replacing a series of side-by-side icon arrays with a single integrated multioutcome icon array (subsection 5B: graphical formats comparison),
- grouping versus randomly distributing icons in icon arrays (subsection 5B: graphical formats comparison),
- supplementing a rate per 10^n with a part-to-whole icon array (subsection 5C: numerical and graphical format comparison),
- supplementing a rate per 10^n with a foreground-only (numerator-only) icon array (subsection 5C: numerical and graphical format comparison),
- supplementing a probability of an event with the chances of comparison events (subsection 5E: contextual information comparison), and
- replacing a static graphic with an interactive one (subsection 5I: animation or interactivity comparison).

Specifically, we found only **moderate to strong evidence** that the following manipulations increase behavior and/or behavioral intentions:

- using numbers instead of verbal probability terms to describe negative outcomes such as side effects (subsection 7D: numerical versus verbal probabilities comparison),
- presenting probabilities as 1 in X versus other number formats (subsection 7A: numerical formats comparison),
- adding anecdotes showing an outcome—the effect direction depends on whether the outcome described is positive or negative (subsection 7E: contextual information comparison),
- presenting positive outcomes such as efficacy using gain framing (subsection 7F: framing comparison), and
- presenting negative outcomes using gain framing (e.g., chance of avoiding a side effect) (subsection 7F: framing comparison).

In addition, evidence is **moderate to strong** that providing a population average probability for comparison (subsection 7E: contextual information comparison) does not affect behavior or behavioral intention.

There is **moderate-strength** evidence regarding preferences such that people prefer numbers alone or in combination with verbal interpretations to verbal probabilities alone (subsection 9D: numerical versus verbal probabilities comparison) and that people do **not** have strong preferences between gain-framed and loss-framed probabilities (subsection 9F: framing comparison).

Few studies examined discrimination outcomes, and as a result, all findings were either weak or insufficient.

Discussion

The outcomes discussed in this article (see Table A) focus on the interpretation of probability information and subsequent perceptions and actions. They include both cognitive and emotional interpretations of probability magnitudes (probability perceptions, probability feelings) and measures of the degree that people adjust their perceptions to variations in those magnitudes (discrimination), which is a particularly important measure of the degree that a probability presentation format enables people to derive accurate and meaningful information from the probability data. They also include the behavioral intentions or actual behaviors after exposure to probability information. We also report evidence related to both preference and trust outcomes, which conceptually moderate peoples' attention to information and the degree that they are influenced by that information.[168–170] Put another way, these outcomes measure what people think, feel, or do about probabilities, whether or not they can remember, categorize, contrast, or perform computations with them. (See our companion article for the evidence regarding these additional outcomes.)

Probability perceptions and feelings are the most frequently studied outcomes in probability communication, and evidence points to a large number of ways that data presentation formats can influence these outcomes. For the most part, perceptions and feelings had similar strength of evidence, but in a few cases, evidence was weaker for feelings. Broadly speaking, these findings are a reminder of the persuasive power inherent in the selection of format for probability communications. The effects identified here are effective tools for intentionally increasing or decreasing reactions to probability information if persuasion is desired and appropriate. Yet, it is worth remembering that no "gold standard" level of probability perceptions and/or feelings exists, and thus it is always up to the communication designer to assess whether using a particular format will support or inhibit their communication goal. (Importantly, the accuracy of probability perceptions is not covered in this article. Whether participants accurately identified a probability while viewing a stimulus, or recalled it after the stimulus was removed, is considered an identification-recall outcome and is covered in the companion Part 1 article in this series.)

Although probability perceptions and feelings are precursors to behavior in well-established health behavior models such as the health belief model[172] and the extended parallel process model,[172,173] the evidence about behavior and behavioral intention is far weaker in our review. We did not find evidence against a link between perceptions and behaviors but rather found that both behaviors and intentions were less frequently studied than perceptions and feelings, and the findings that are available for the distal outcome are mixed.

Similarly, we found that preferences for communication formats are frequently studied, but the resulting findings are extremely mixed. Other evidence is limited by the fact that the data presentation formats studied, especially the graphics, had important structural differences. Overall, we conclude that preferences may not be reliably or stably linked to data presentation format. In fact, it appears that preferences may be influenced by factors such as format familiarity and the differing meanings that readers need to derive from the communication. We found that trust in the communication also has similarly weak and mixed evidence.

Limitations of this set of reviews includes the possibility of missed and omitted studies, the use of a small expert group to evaluate risk of bias and credibility, and the highly granular data extraction process that slowed work and may make it less replicable. Types of studies varied so that the numbers of high-credibility articles within any category was small, and this limited the strength of the evidence that could be derived from this literature. We did not subset the studies by participant or population characteristics such as education, culture, or numeracy because of the small numbers of comparable articles for each relevant characteristic and because of significant heterogeneity in how these characteristics were measured. These potential confounding factors could contribute to the heterogeneity of findings when studies are grouped and may also limit generalizability of the findings to diverse settings and populations.

Another key limitation is that this article examines only the research evidence related to peoples' ability to perform what we have termed *point tasks*, that is, situations in which the audience is asked to focus on single probabilities (presented either alone or in larger sets). It does not touch upon the evidence for communicating probability differences (difference tasks), which include, for example, evaluations of the effectiveness of health interventions or the magnitude of differences in risk across populations. Nor does this article consider situations in which the audience is asked to consider time trends (trend tasks) or multiple types of probability information simultaneously (synthesis tasks), such as evaluations of or choices about treatment options that affect people's chances of experiencing multiple complication risks and/or benefits.

In summary, many factors appear to modify probability perceptions and feelings in the context of point tasks, while other factors appear to have few or no effects. In addition, the link between interpretation of probability information and health behaviors, despite being supported by theory, is not strongly reflected in the evidence about communicating probabilities. Also, the large amount of research into patient preferences for numerical, graphical, and verbal formats has yielded relatively little evidence that would suggest stable and predictable preferences, other than a general preference for numbers alone or combined with words rather than words alone. While these findings do provide clear evidence that the choice of formats used to present probabilities will influence people's perceptions of individual probabilities under certain circumstances, we remind readers that the findings presented here represent only one part of the larger evidence base needed to guide practice. We encourage future efforts to use our findings to clearly define the ways that different methods of communicating probabilities have both strengths and weaknesses when used to achieve various communication goals.

Acknowledgment

We thank the Numeracy Expert Panel for contributions to conceptualizing the MNM project (Cynthia Baur, Sara Cjaza, Angela Fagerlin, Carolyn Petersen, Rima Rudd, Michael Wolf, and Steven Woloshin). We are grateful to Marianne Sharko, MD, MS, Andrew Z. Liu, MPH, and Lisa Grossman Liu, MD, PhD, for contributions to article screening and risk of bias assessment. We also thank Jordan Brutus for assisting with data management.

ORCID iDs

Jessica S. Ancker https://orcid.org/0000-0002-3859-9130
Brian J. Zikmund-Fisher https://orcid.org/0000-0002-1637-4176

Availability of Research Resources

All research resources are available at the Making Numbers Meaningful Project at OSF (https://osf.io/rvxf2/). This project includes a Methodology Files folder (containing the search strategy, the data extraction instrument, and the study risk of bias [S-ROB] rubric), the list of each included article mapped to the Making Numbers Meaningful review article that covers it, and a Probability Findings folder displaying the extracted findings for each of the Making Numbers Meaningful review articles.

References

1. Ancker JS, Benda NC, Sharma MM, et al. Scope, methods, and overview findings for the Making Numbers Meaningful evidence review of communicating probabilities in health: a systematic review. *MDM Policy Pract*. 2025;10(1): 23814683241255334. DOI: 10.1177/238146832 41255334
2. Ancker JS, Benda NC, Sharma MM, Johnson SB, Weiner S, Zikmund-Fisher BJ. Taxonomies for synthesizing the evidence on communicating numbers in health: goals, format, and structure. *Risk Anal*. 2022;42(12):2656–70. DOI: 10.1111/risa.13875
3. Ancker JS, Benda NC, Sharma MM, et al. How point (single-probability) tasks are affected by probability format, part 1: a Making Numbers Meaningful systematic review. *MDM Policy Pract*. 2025;10(1):23814683241255333. DOI: 10.1177/23814683241255333
4. Gigerenzer G. What are natural frequencies? *BMJ*. 2011;343:d6386. DOI: 10.1136/bmj.d6386
5. Hoffrage U, Gigerenzer G. Using natural frequencies to improve diagnostic inferences. *Acad Med*. 1998;73(5): 538–40.
6. Fair AKI, Murray PG, Thomas A, Cobain MR. Using hypothetical data to assess the effect of numerical format and context on the perception of coronary heart disease risk. *Am J Health Promot*. 2008;22(4):291–6. DOI: 10.4278/061030140R2.1
7. Chapman AR, Litton E, Chamberlain J, Ho KM. The effect of prognostic data presentation format on perceived risk among surrogate decision makers of critically ill patients: a randomized comparative trial. *J Crit Care*. 2015;30(2):231–5. DOI: 10.1016/j.jcrc.2014.11.005
8. Sirota M, Juanchich M, Kostopoulou O, Hanak R. Decisive evidence on a smaller-than-you-think phenomenon revisiting the "1-in-X" effect on subjective medical probabilities. *Med Decis Making*. 2014;34(4):419–29. DOI: 10.1177/0272989X13514776
9. Pighin S, Savadori L, Barilli E, Cremonesi L, Ferrari M, Bonnefon J-F. The 1-in-X effect on the subjective assessment of medical probabilities. *Med Decis Making*. 2011;31(5):721–9. DOI: 10.1177/0272989X11403490
10. Sirota M, Juanchich M, Bonnefon JF. "1-in-X" bias: "1-in-X" format causes overestimation of health-related risks. *J Exp Psychol Appl*. 2018;24(4):431–9. DOI: 10.1037/xap0000190
11. Pighin S, Savadori L, Barilli E, et al. Communicating Down syndrome risk according to maternal age: "1-in-X" effect on perceived risk. *Prenat Diagn*. 2015;35(8):777–82. DOI: 10.1002/pd.4606
12. Knapp P, Raynor DK, Woolf E, Gardner PH, Carrigan N, McMillan B. Communicating the risk of side effects to patients: an evaluation of UK regulatory recommendations. *Drug Saf*. 2009;32(10):837–49. DOI: 10.2165/11316570-000000000-00000
13. Wu L, Zeng S, Wu Y. Affect heuristic and format effect in risk perception. *Soc Behav Pers*. 2018;46(8):1331–44. DOI: 10.2224/SBP.6957
14. Gurmankin AD, Baron J, Armstrong K. Intended message versus message received in hypothetical physician risk communications: exploring the gap. *Risk Anal*. 2004;24(5):1337–47.
15. Knapp P, Gardner P, McMillan B, Raynor DK, Woolf E. Evaluating a combined (frequency and percentage) risk

expression to communicate information on medicine side effects to patients. *Int J Pharm Pract*. 2013;21(4):226–32. DOI: 10.1111/j.2042-7174.2012.00254.x
16. Hill WT, Brase GL. When and for whom do frequencies facilitate performance? On the role of numerical literacy. *Q J Exp Psychol*. 2012;65(12):2343–68. DOI: 10.1080/17470218.2012.687004
17. Peters E, Hart PS, Fraenkel L. Informing patients: the influence of numeracy, framing, and format of side effect information on risk perceptions. *Med Decis Making*. 2011;31(3):432–6. DOI: 10.1177/0272989X10391672
18. Halpern DF, Blackman S, Salzman B. Using statistical risk information to assess oral contraceptive safety. *Appl Cogn Psychol*. 1989;3(3):251–60. DOI: 10.1002/acp.2350030305
19. Peters E, Västfjäll D, Slovic P, Mertz CK, Mazzocco K, Dickert S. Numeracy and decision making. *Psychol Sci*. 2006;17(5):407–13. DOI: 10.1111/j.1467-9280.2006.01720.x
20. Chen Y, Yang ZJ. Message formats, numeracy, risk perceptions of alcohol-attributable cancer, and intentions for binge drinking among college students. *J Drug Educ*. 2015;45(1):37–55.
21. Zikmund-Fisher BJ, Fagerlin A, Roberts TR, Derry HA, Ubel PA. Alternate methods of framing information about medication side effects: incremental risk versus total risk occurence. *J Health Commun*. 2008;13(2):107–24.
22. Shepperd JA, Lipkus IM, Sanderson SC, McBride CM, O'Neill SC, Docherty S. Testing different communication formats on responses to imagined risk of having versus missing the GSTM1 gene. *J Health Commun*. 2013;18(1):124–37. DOI: 10.1080/10810730.2012.688245
23. Zikmund-Fisher BJ, Ubel PA, Smith DM, et al. Communicating side effect risks in a tamoxifen prophylaxis decision aid: the debiasing influence of pictographs. *Patient Educ Couns*. 2008;73(2):209–14.
24. Berry DC, Knapp P, Raynor T. Expressing medicine side effects: assessing the effectiveness of absolute risk, relative risk, and number needed to harm, and the provision of baseline risk information. *Patient Educ Couns*. 2006;63(1–2):89–96. DOI: 10.1016/j.pec.2005.09.003
25. Sullivan HW, O'Donoghue AC, Aikin KJ. Communicating benefit and risk information in direct-to-consumer print advertisements: a randomized study. *Ther Innov Regul Sci*. 2015;49(4):493–502. DOI: 10.1177/2168479015572370
26. Tait AR, Voepel-Lewis T, Zikmund-Fisher BJ, Fagerlin A. The effect of format on parents' understanding of the risks and benefits of clinical research: a comparison between text, tables, and graphics. *J Health Commun*. 2010;15(5):487–501. DOI: 10.1080/10810730.2010.492560
27. Tait AR, Voepel-Lewis T, Zikmund-Fisher BJ, Fagerlin A. Presenting research risks and benefits to parents: does format matter? *Anesth Analg*. 2010;111(3):718–23. DOI: 10.1213/ANE.0b013e3181e8570a
28. Schwartz LM, Woloshin S, Welch HG. Using a drug facts box to communicate drug benefits and harms: two randomized trials. *Ann Intern Med*. 2009;150(8):516–27. DOI: 10.1059/0003-4819-150-8-200904210-00106
29. Witteman HO, Zikmund-Fisher BJ, Waters EA, Gavaruzzi T, Fagerlin A. Risk estimates from an online risk calculator are more believable and recalled better when expressed as integers. *J Med Internet Res*. 2011;13(3):e54. DOI: 10.2196/jmir.1656
30. Bonner C, Jansen J, Newell BR, et al. Is the "heart age" concept helpful or harmful compared to absolute cardiovascular disease risk? An experimental study. *Med Decis Making*. 2015;35(8):967–78. DOI: 10.1177/0272989X15597224
31. Soureti A, Hurling R, Murray P, Van Mechelen W, Cobain M. Evaluation of a cardiovascular disease risk assessment tool for the promotion of healthier lifestyles. *Eur J Prev Cardiol*. 2010;17(5):519–23. DOI: 10.1097/HJR.0b013e328337ccd3
32. Ubel PA, Smith DM, Zikmund-Fisher BJ, et al. Testing whether decision aids introduce cognitive biases: results of a randomized trial. *Patient Educ Couns*. 2010;80(2):158–63. DOI: 10.1016/j.pec.2009.10.021
33. Miron-Shatz T, Hanoch Y, Graef D, Sagi M. Presentation format affects comprehension and risk assessment: the case of prenatal screening. *J Health Commun*. 2009;14(5):439–50. DOI: 10.1080/10810730903032986
34. Weinstein ND, Kolb K, Goldstein BD. Using time intervals between expected events to communicate risk magnitudes. *Risk Anal*. 1996;16(3):305–8. DOI: 10.1111/j.1539-6924.1995.tb00792.x
35. Knapp P, Gardner P, McMillan B, Raynor DK, Woolfb L. The effects of numeracy on the perceived risk of medicine side effects. *Int J Pharm Pract*. 2010;2:23–24. DOI: 10.1111/j.2042-7174.2010.tb00509.x
36. Strathie A, Netto G, Walker GH, Pender G. How presentation format affects the interpretation of probabilistic flood risk information. *J Flood Risk Manag*. 2017;10(1):87–96. DOI: 10.1111/jfr3.12152
37. Stone ER, Bruine de Bruin W, Wilkins AM, Boker EM, MacDonald Gibson J. Designing graphs to communicate risks: understanding how the choice of graphical format influences decision making. *Risk Anal*. 2017;37(4):612–28. DOI: 10.1111/risa.12660
38. Okan Y, Stone ER, Bruine de Bruin W. Designing graphs that promote both risk understanding and behavior change. *Risk Anal*. 2018;38(5):929–46. DOI: 10.1111/risa.12895
39. Okan Y, Stone ER, Parillo J, Bruine de Bruin W, Parker AM. Probability size matters: the effect of foreground-only versus foreground + background graphs on risk aversion diminishes with larger probabilities. *Risk Anal*. 2020;40(4):771–88. DOI: 10.1111/risa.13431
40. Stone ER, Reeder EC, Parillo J, Long C, Walb L. Salience versus proportional reasoning: rethinking the mechanism behind graphical display effects. *J Behav Decis Mak*. 2018;31(4):473–86. DOI: 10.1002/bdm.2051
41. Weinstein ND, Sandman PM, Hallman WK. Testing a visual display to explain small probabilities. *Risk Anal*. 1994;14(6):895–6. DOI: 10.1111/j.1539-6924.1994.tb00053.x

42. Schapira MM, Nattinger AB, McAuliffe TL. The influence of graphic format on breast cancer risk communication. *J Health Commun*. 2006;11(6):569–82.
43. Kreuzmair C, Siegrist M, Keller C. Does iconicity in pictographs matter? The influence of iconicity and numeracy on information processing, decision making, and liking in an eye-tracking study. *Risk Anal*. 2017;37(3):546–56. DOI: 10.1111/risa.12623
44. Zikmund-Fisher BJ, Witteman HO, Dickson M, et al. Blocks, ovals, or people? Icon type affects risk perceptions and recall of pictographs. *Med Decis Making*. 2014;34(4): 443–53. DOI: 10.1177/0272989X13511706
45. McDowell M, Gigerenzer G, Wegwarth O, Rebitschek FG. Effect of tabular and icon fact box formats on comprehension of benefits and harms of prostate cancer screening: a randomized trial. *Med Decis Making*. 2019;39(1):41–56. DOI: 10.1177/0272989X18818166
46. Navar AM, Wang TY, Mi X, et al. Influence of cardiovascular risk communication tools and presentation formats on patient perceptions and preferences. *JAMA Cardiol*. 2018;3(12):1192–9. DOI: 10.1001/jamacardio.2018.3680
47. Adarkwah CC, Jegan N, Heinzel-Gutenbrunner M, et al. Time-to-event versus ten-year-absolute-risk in cardiovascular risk prevention—does it make a difference? Results from the Optimizing-Risk-Communication (OptRisk) randomized-controlled trial. *BMC Med Inform Decis Mak*. 2016;16(1):152. DOI: 10.1186/s12911-016-0393-1
48. Adarkwah CC, Jegan N, Heinzel-Gutenbrunner M, et al. The Optimizing-Rrisk-Ccommunication (OptRrisk) randomized trial—impact of decision-aid-based consultation on adherence and perception of cardiovascular risk. *Patient Prefer Adherence*. 2019;13:441–52. DOI: 10.2147/PPA.S197545
49. Ancker JS, Weber EU, Kukafka R. Effect of arrangement of stick figures on estimates of proportion in risk graphics. *Med Decis Making*. 2011;31(1):143–50.
50. Wright AJ, Whitwell SCL, Takeichi C, Hankins M, Marteau TM. The impact of numeracy on reactions to different graphic risk presentation formats: an experimental analogue study. *Br J Health Psychol*. 2009;14(pt 1):107–25. DOI: 10.1348/135910708X304432
51. Wright AJ, Takeichi C, Whitwell SCL, Hankins M, Marteau TM. The impact of genetic testing for Crohn's disease, risk magnitude and graphical format on motivation to stop smoking: an experimental analogue study. *Clin Genet*. 2008;73(4):306–14. DOI: 10.1111/j.1399-0004.2008.00964.x
52. Han PKJ, Klein WMP, Killam B, Lehman T, Massett H, Freedman AN. Representing randomness in the communication of individualized cancer risk estimates: effects on cancer risk perceptions, worry, and subjective uncertainty about risk. *Patient Educ Couns*. 2012;86(1):106–13. DOI: 10.1016/j.pec.2011.01.033
53. Janssen E, Ruiter RAC, Waters EA. Combining risk communication strategies to simultaneously convey the risks of four diseases associated with physical inactivity to sociodemographically diverse populations. *J Behav Med*. 2018;41(3):318–32. DOI: 10.1007/s10865-017-9894-3
54. Siegrist M, Orlow P, Keller C. The effect of graphical and numerical presentation of hypothetical prenatal diagnosis results on risk perception. *Med Decis Making*. 2008;28(4): 567–74. DOI: 10.1177/0272989X08315237
55. Gibson JM, Rowe A, Stone ER, Bruine De Bruin W. Communicating quantitative information about unexploded ordnance risks to the public. *Environ Sci Technol*. 2013;47(9):4004–13. DOI: 10.1021/es305254j
56. Emmons KM, Wong M, Puleo E, Weinstein N, Fletcher R, Colditz G. Tailored computer-based cancer risk communication: correcting colorectal cancer risk perception. *J Health Commun*. 2004;9(2):127–41. DOI: 10.1080/10810730490425295
57. Hamilton JG, Genoff Garzon M, Shah IH, et al. Illustrating cancer risk: patient risk communication preferences and interest regarding a novel BRCA1/2 genetic risk modifier test. *Public Health Genomics*. 2020;23(1–2):6–19. DOI: 10.1159/000505854
58. Damman OC, Vonk SI, van den Haak MJ, van Hooijdonk CMJ, Timmermans DRM. The effects of infographics and several quantitative versus qualitative formats for cardiovascular disease risk, including heart age, on people's risk understanding. *Patient Educ Couns*. 2018;101(8):1410–8. DOI: 10.1016/j.pec.2018.03.015
59. Keller C, Siegrist M. Effect of risk communication formats on risk perception depending on numeracy. *Med Decis Making*. 2009;29(4):483–90. DOI: 10.1177/0272989X09333122
60. Galesic M, Garcia-Retamero R, Gigerenzer G. Using icon arrays to communicate medical risks: overcoming low numeracy. *Health Psychol*. 2009;28(2):210–6. DOI: 10.1037/a0014474
61. Colome A, Rodriguez-Ferreiro J, Tubau E. Too worried to judge: on the role of perceived severity in medical decision-making. *Front Psychol*. 2018;9:1906. DOI: 10.3389/fpsyg.2018.01906
62. Fraenkel L, Nowell WB, Stake CE, et al. The impact of information presentation format on preference for total knee replacement surgery. *Arthritis Care Res (Hoboken)*. 2019;71(3):379–84. DOI: 10.1002/acr.23605
63. Cameron LD, Marteau TM, Brown PM, Klein WM, Sherman KA. Communication strategies for enhancing understanding of the behavioral implications of genetic and biomarker tests for disease risk: the role of coherence. *J Behav Med*. 2012;35(3):286–98. DOI: 10.1007/s10865-011-9361-5
64. Cozmuta R, Wilhelms E, Cornell D, Nolte J, Reyna V, Fraenkel L. Influence of explanatory images on risk perceptions and treatment preference. *Arthritis Care Res*. 2018;70(11):1707–11. DOI: 10.1002/acr.23517
65. Stone ER, Gabard AR, Groves AE, Lipkus IM. Effects of numerical versus foreground-only icon displays on understanding of risk magnitudes. *J Health Commun*. 2015; 20(10):1230–41. DOI: 10.1080/10810730.2015.1018594
66. Leonhardt JM, Robin Keller L. Do pictographs affect probability comprehension and risk perception of

multiple-risk communications? *J Consum Aff*. 2018;52(3): 756–69. DOI: 10.1111/joca.12185
67. Han PKJ, Klein WMP, Lehman T, Killam B, Massett H, Freedman AN. Communication of uncertainty regarding individualized cancer risk estimates: effects and influential factors. *Med Decis Making*. 2011;31(2):354–66. DOI: 10.1177/0272989X10371830
68. Lipkus IM, Klein WM, Rimer BK. Communicating breast cancer risks to women using different formats. *Cancer Epidemiol Biomarkers Prev*. 2001;10(8):895–8.
69. Lee DH, Mehta MD. Evaluation of a visual risk communication tool: effects on knowledge and perception of blood transfusion risk. *Transfusion*. 2003;43(6):779–87.
70. Timmermans DR, Oudhoff J. Indicating risks in the Dutch Cancer Society Cancer Risk Test: indicating population risks improves risk perception. *Ned Tijdschr Geneeskd*. 2012;156(21):A4961.
71. Henneman L, Oosterwijk JC, Van Asperen CJ, et al. The effectiveness of a graphical presentation in addition to a frequency format in the context of familial breast cancer risk communication: a multicenter controlled trial. *BMC Med Inform Decis Mak*. 2013;13(1):55. DOI: 10.1186/1472-6947-13-55
72. European Commission. Guideline on the readability of the labelling and package leaflet of medicinal products for human use. 2009. Available from: https://ec.europa.eu/health/sites/default/files/files/eudralex/vol-2/c/2009_01_12_readability_guideline_final_en.pdf
73. Berry DC, Knapp P, Raynor DK. Provision of information about drug side-effects to patients. *Lancet*. 2002;359(9309): 853–4.
74. Berry DC, Hochhauser M. Verbal labels can triple perceived risk in clinical trials. *Ther Innov Regul Sci*. 2006;40(3):249–58. DOI: 10.1177/009286150604000302
75. Peters E, Hart PS, Tusler M, Fraenkel L. Numbers matter to informed patient choices: a randomized design across age and numeracy levels. *Med Decis Making*. 2014;34(4): 430–42.
76. Young S, Oppenheimer DM. Effect of communication strategy on personal risk perception and treatment adherence intentions. *Psychol Health Med*. 2009;14(4):430–42. DOI: 10.1080/13548500902890103
77. Knapp P, Raynor DK, Berry DC. Comparison of two methods of presenting risk information to patients about the side effects of medicines. *Qual Saf Health Care*. 2004;13:176–80.
78. Berry D, Raynor T, Knapp P, Bersellini E. Over the counter medicines and the need for immediate action: a further evaluation of European Commission recommended wordings for communicating risk. *Patient Educ Couns*. 2004;53(2):129–34. DOI: 10.1016/S0738-3991(03)00111-3
79. Schwartz PH, Imperiale TF, Perkins SM, Schmidt KK, Althouse S, Rawl SM. Impact of including quantitative information in a decision aid for colorectal cancer screening: a randomized controlled trial. *Patient Educ Couns*. 2019;102(4):726–34. DOI: 10.1016/j.pec.2018.11.010
80. Sawant R, Sansgiry S. Communicating risk of medication side-effects: role of communication format on risk perception. *Pharm Pract*. 2018;16(2):1174. DOI: 10.18549/PharmPract.2018.02.1174
81. Dahlstrom MF, Dudo A, Brossard D. Precision of information, sensational information, and self-efficacy information as message-level variables affecting risk perceptions. *Risk Anal*. 2012;32(1):155–66. DOI: 10.1111/j.1539-6924.2011.01641.x
82. Sheridan SL, Sutkowi-Hemstreet A, Barclay C, et al. A comparative effectiveness trial of alternate formats for presenting benefits and harms information for low-value screening services: a randomized clinical trial. *JAMA Intern Med*. 2016;176(1):31–41. DOI: 10.1001/jamainternmed.2015.7339
83. Char SJL, Evans LR, Malvar GL, White DB. A randomized trial of two methods to disclose prognosis to surrogate decision makers in intensive care units. *Am J Respir Crit Care Med*. 2010;182(7):905–9. DOI: 10.1164/rccm.201002-0262OC
84. Knapp P, Gardner PH, Woolf E. Combined verbal and numerical expressions increase perceived risk of medicine side-effects: a randomized controlled trial of EMA recommendations. *Health Expect*. 2016;19(2):264–74. DOI: 10.1111/hex.12344
85. Marteau TM, Saidi G, Goodburn S, Lawton J, Michie S, Bobrow M. Numbers or words? A randomized controlled trial of presenting screen negative results to pregnant women. *Prenat Diagn*. 2000;20(9):714–8.
86. de Wit JB, Das E, Vet R. What works best: objective statistics or a personal testimonial? An assessment of the persuasive effects of different types of message evidence on risk perception. *Health Psychol*. 2008;27(1):110–5. DOI: 10.1037/0278-6133.27.1.110
87. Gutierrez KM, Cohn LD. Perceived risk of emerging recreational drugs: impact of anecdotal and statistical evidence. *J Drug Issues*. 2018;48(3):435–51. DOI: 10.1177/0022042618770632
88. Betsch C, Ulshöfer C, Renkewitz F, Betsch T. The influence of narrative v. statistical information on perceiving vaccination risks. *Med Decis Making*. 2011;31(5):742–53. DOI: 10.1177/0272989X11400419
89. Kuhn KM. Message format and audience values: interactive effects of uncertainty information and environmental attitudes on perceived risk. *J Environ Psychol*. 2000;20(1): 41–51. DOI: 10.1006/jevp.1999.0145
90. Zikmund-Fisher BJ, Fagerlin A, Keeton K, Ubel PA. Does labeling prenatal screening test results as negative or positive affect a woman's responses? *Am J Obstet Gynecol*. 2007;197(5):528.e1-6.
91. Bonner C, Newell BR. How to make a risk seem riskier: the ratio bias versus construal level theory. *Judgm Decis Mak*. 2008;3(5):411–6.
92. Weinstein ND, Atwood K, Puleo E, Fletcher R, Colditz G, Emmons K. Colon cancer: risk perceptions and risk communication. *J Health Commun*. 2004;9(1):53–65.

93. Lipkus IM, Biradavolu M, Fenn K, Keller P, Rimer BK. Informing women about their breast cancer risks: truth and consequences. *Health Commun*. 2001;13(2):205–26. DOI: 10.1207/S15327027HC1302_5
94. Hess R, Visschers VHM, Siegrist M. Risk communication with pictographs: the role of numeracy and graph processing. *Judgm Decis Mak*. 2011;6(3):263–74.
95. Harris P, Smith V. When the risks are low: the impact of absolute and comparative information on disturbance and understanding in US and UK samples. *Psychol Health*. 2005;20(3):319–30. DOI: 10.1080/08870440512331317689
96. Raghubir P. Is 1/10 > 10/100? The effect of denominator salience on perceptions of base rates of health risk. *Int J Res Mark*. 2008;25(4):327–34. DOI: 10.1016/j.ijresmar.2008.07.006
97. Schwartz PH, Perkins SM, Schmidt KK, Muriello PF, Althouse S, Rawl SM. Providing quantitative information and a nudge to undergo stool testing in a colorectal cancer screening decision aid: a randomized clinical trial. *Med Decis Making*. 2017;37(6):688–702. DOI: 10.1177/02729 89X17698678
98. Dillard AJ, McCaul KD, Kelso PD, Klein WMP. Resisting good news: reactions to breast cancer risk communication. *Health Commun*. 2006;19(2):115–23. DOI: 10.1207/s15327027hc1902_3
99. Garcia-Retamero R, Galesic M. How to reduce the effect of framing on messages about health. *J Gen Intern Med*. 2010;25(12):1323–9. DOI: 10.1007/s11606-010-1484-9
100. Williams T, Clarke V, Borland R. Effects of message framing on breast-cancer-related beliefs and behaviors: the role of mediating factors. *J Appl Soc Psychol*. 2001;31(5):925–50. DOI: 10.1111/j.1559-1816.2001.tb02656.x
101. Farrell K, Ferguson E, James V, Lowe KC. Confidence in the safety of blood for transfusion: the effect of message framing. *Transfusion*. 2001;41(11):1335–40. DOI: 10.1046/j.1537-2995.2001.41111335.x
102. Kalluru R, Petrie KJ, Grey A, et al. Randomised trial assessing the impact of framing of fracture risk and osteoporosis treatment benefits in patients undergoing bone densitometry. *BMJ Open*. 2017;7(2):e013703. DOI: 10.1136/bmjopen-2016-013703
103. Webster RK, Weinman J, Rubin GJ. Positively framed risk information in patient information leafets reduces side effect reporting: a double-blind randomized controlled trial. *Ann Behav Med*. 2018;52(11):920–9. DOI: 10.1093/abm/kax064
104. Peng J, Li H, Miao D, Feng X, Xiao W. Five different types of framing effects in medical situation: a preliminary exploration. *Iran Red Crescent Med J*. 2013;15(2):161–5.
105. Peng J, Jiang Y, Miao D, Li R, Xiao W. Framing effects in medical situations: distinctions of attribute, goal and risky choice frames. *J Int Med Res*. 2013;41(3):771–6. DOI: 10.1177/0300060513476593
106. Kreiner H, Gamliel E. Are highly numerate individuals invulnerable to attribute framing bias? Comparing numerically and graphically represented attribute framing. *Eur J Soc Psychol*. 2017;47(6):775–82. DOI: 10.1002/ejsp.2272
107. Bigman CA, Cappella JN, Hornik RC. Effective or ineffective: attribute framing and the human papillomavirus (HPV) vaccine. *Patient Educ Couns*. 2010;81(suppl 1):S70–6. DOI: 10.1016/j.pec.2010.08.014
108. Zamarian L, Benke T, Buchler M, Wenter J, Delazer M. Information about medications may cause misunderstanding in older adults with cognitive impairment. *J Neurol Sci*. 2010;298(1–2):46–51. DOI: 10.1016/j.jns.2010.08.061
109. Gamliel E, Kreiner H, Garcia-Retamero R. The moderating role of objective and subjective numeracy in attribute framing. *Int J Psychol*. 2016;51(2):109–16. DOI: 10.1002/ijop.12138
110. Biswas D, Pechmann C. What do these clinical trial results mean? How product efficacy judgments are affected by data partitioning, framing, and quantification. *Organ Behav Hum Decis Process*. 2012;117(2):341–50. DOI: 10.1016/j.obhdp.2011.11.007
111. Levin IP, Schnittjer SK, Thee SL. Information framing effects in social and personal decisions. *J Exp Soc Psychol*. 1988;24(6):520–9. DOI: 10.1016/0022-1031(88)90050-9
112. Johnson BB, Slovic P. Presenting uncertainty in health risk assessment: initial studies of its effects on risk perception and trust. *Risk Anal*. 1995;15(4):485–94. DOI: 10.1111/j.1539-6924.1995.tb00341.x
113. Sladakovic J, Jansen J, Hersch J, Turner R, McCaffery K. The differential effects of presenting uncertainty around benefits and harms on treatment decision making. *Patient Educ Couns*. 2016;99(6):974–80. DOI: 10.1016/j.pec.2016.01.009
114. Longman T, Turner RM, King M, McCaffery KJ. The effects of communicating uncertainty in quantitative health risk estimates. *Patient Educ Couns*. 2012;89(2):252–9. DOI: 10.1016/j.pec.2012.07.010
115. Yamagishi K.When a 86% mortality is more dangerous than 24.14%: implications for risk communication. *Appl Cogn Psychol*. 1997;11(6):495–506. DOI: 10.1002/(SICI)1099-0720(199712)11:6<495::AID-ACP481>3.0.CO;2-J
116. Ancker JS, Weber EU, Kukafka R. Effects of game-like interactive graphics on risk perceptions and decisions. *Med Decis Making*. 2011;31(1):130–42.
117. Witteman HO, Fuhrel-Forbis A, Wijeysundera HC, et al. Animated randomness, avatars, movement, and personalization in risk graphics. *J Med Internet Res*. 2014;16(3):e80. DOI: 10.2196/jmir.2895
118. Harle CA, Downs JS, Padman R. Effectiveness of personalized and interactive health risk calculators: a randomized trial. *Med Decis Making*. 2012;32(4):594–605. DOI: 10.1177/0272989X11431736
119. Natter HM, Berry DC. Effects of active information processing on the understanding of risk information. *Appl Cogn Psychol*. 2005;19:123–35. DOI: 10.1002/acp.1068

120. Lopez-Gonzalez AA, Aguilo A, Frontera M, et al. Effectiveness of the Heart Age tool for improving modifiable cardiovascular risk factors in a Southern European population: a randomized trial. *Eur J Prev Cardiol*. 2015;22(3): 389–96. DOI: 10.1177/2047487313518479
121. Wadhwa M, Zhang K. When numbers make you feel: impact of round versus precise numbers on preventive health behaviors. *Organ Behav Hum Decis Process*. 2019;150:101–11.
122. Garcia-Retamero R, Cokely ET. The influence of skills, message frame, and visual aids on prevention of sexually transmitted diseases. *J Behav Decis Mak*. 2014;27(2): 179–89. DOI: 10.1002/bdm.1797
123. Sinayev A, Peters E, Tusler M, Fraenkel L. Presenting numeric information with percentages and descriptive risk labels: a randomized trial. *Med Decis Making*. 2015;35(8): 937–47. DOI: 10.1177/0272989X15584922
124. Tan SB, Goh C, Thumboo J, Che W, Chowbay B, Cheung YB. Risk perception is affected by modes of risk presentation among Singaporeans. *Ann Acad Med Singapore*. 2005;34(2):184–7.
125. Cheung YB, Wee HL, Thumboo J, et al. Risk communication in clinical trials: a cognitive experiment and a survey. *BMC Med Inform Decis Mak*. 2010;10:55. DOI: 10.1186/1472-6947-10-55
126. Eyler RF, Cordes S, Szymanski BR, Fraenkel L. Utilization of continuous "spinners" to communicate risk. *Med Decis Making*. 2017;37(6):725–9. DOI: 10.1177/0272989X17707198
127. Berry DC, Raynor DK, Knapp P. Communicating risk of medication side effects: an empirical evaluation of EU recommended terminology. *Psychol Health Med*. 2003;8(3): 251–63. DOI: 10.1080/1354850031000135704
128. Young SD, Oppenheimer DM. Different methods of presenting risk information and their influence on medication compliance intentions: results of three studies. *Clin Ther*. 2006;28(1):129–39. DOI: 10.1016/j.clinthera.2006.01.013
129. Lipkus IM, Crawford Y, Fenn K, et al. Testing different formats for communicating colorectal cancer risk. *J Health Commun*. 1999;4(4):311–24. DOI: 10.1080/108107399126841
130. Gurm HS, Litaker DG. Framing procedural risks to patients: is 99% safe the same as a risk of 1 in 100? *Acad Med*. 2000;75(8):840–2. DOI: 10.1097/00001888-200008000-00018
131. Donovan RJ, Jalleh G. Positive versus negative framing of a hypothetical infant immunization: the influence of involvement. *Health Educ Behav*. 2000;27(1):82–95. DOI: 10.1177/109019810002700108
132. Taylor J, Seyed-Hosseini M, Quest D. Framing the risk of an OTC medication side effect. *Can Pharm J*. 2011;144(1): 34–39. DOI: 10.3821/1913-701X-144.1.34
133. Wilson DK, Kaplan RM, Schneiderman LJ. Framing of decisions and selections of alternatives in health care. *Soc Behav*. 1987;2(1):51–59.
134. Kuhn KM. Communicating uncertainty: framing effects on responses to vague probabilities. *Organ Behav Hum Decis Process*. 1997;71(1):55–83. DOI: 10.1006/obhd.1997.2715
135. Brewer NT, Richman AR, Defrank JT, Reyna VF, Carey LA. Improving communication of breast cancer recurrence risk. *Breast Cancer Res Treat*. 2012;133(2):553–61. DOI: 10.1007/s10549-011-1791-9
136. Graham PH, Martin RM, Browne LH. Communicating breast cancer treatment complication risks: when words are likely to fail. *Asia Pac J Clin Oncol*. 2009;5(3):193–9. DOI: 10.1111/j.1743-7563.2009.01232.x
137. Henneman L, van Asperen CJ, Oosterwijk JC, Menko FH, Claassen L, Timmermans DR. Do preferred risk formats lead to better understanding? A multicenter controlled trial on communicating familial breast cancer risks using different risk formats. *Patient Prefer Adherence*. 2020;14:333–42. DOI: 10.2147/PPA.S232941
138. Nagle C, Hodges R, Wolfe R, Wallace E. Reporting Down syndrome screening results: women's understanding of risk. *Prenat Diagn*. 2009;29(3):234–9. DOI: 10.1002/pd.2210
139. Gurmankin AD, Baron J, Armstrong K. The effect of numerical statements of risk on trust and comfort with hypothetical physician risk communication. *Med Decis Making*. 2004;24(3):265–71.
140. Brewer NT, Tzeng JP, Lillie SE, Edwards AS, Peppercorn JM, Rimer BK. Health literacy and cancer risk perception: implications for genomic risk communication. *Med Decis Making*. 2009;29(2):157–66. DOI: 10.1177/0272989X08327111
141. Steiner MJ, Dalebout S, Condon S, Dominik R, Trussell J. Understanding risk: a randomized controlled trial of communicating contraceptive effectiveness. *Obstet Gynecol*. 2003;102(4):709–17.
142. Lobb EA, Butow PN, Meiser B, et al. Women's preferences and consultants' communication of risk in consultations about familial breast cancer: impact on patient outcomes. *J Med Genet*. 2003;40(5):e56.
143. Hovick SR, Tan N, Morr L, et al. Understanding BRCA mutation carriers' preferences for communication of genetic modifiers of breast cancer risk. *J Health Commun*. 2019;24(4):377–84. DOI: 10.1080/10810730.2019.1604912
144. Knapp P, Gardner PH, Raynor DK, Woolf E, McMillan B. Perceived risk of tamoxifen side effects: a study of the use of absolute frequencies or frequency bands, with or without verbal descriptors. *Patient Educ Couns*. 2010;79(2): 267–71. DOI: 10.1016/j.pec.2009.10.002
145. Ghosh K, Crawford BJ, Pruthi S, et al. Frequency format diagram and probability chart for breast cancer risk communication: a prospective, randomized trial. *BMC Womens Health*. 2008;8:18. DOI: 10.1186/1472-6874-8-18
146. Edmonds SW, Cram P, Lu X, et al. Improving bone mineral density reporting to patients with an illustration of personal fracture risk. *BMC Med Inform Decis Mak*. 2014;14(1):101. DOI: 10.1186/s12911-014-0101-y

147. Etnel JRG, de Groot JM, El Jabri M, et al. Do risk visualizations improve the understanding of numerical risks? A randomized, investigator-blinded general population survey. *Int J Med Inform*. 2020;135:104005. DOI: 10.1016/j.ijmedinf.2019.104005
148. Connelly NA, Knuth BA. Evaluating risk communication: examining target audience perceptions about four presentation formats for fish consumption health advisory information. *Risk Anal*. 1998;18(5):649–59. DOI: 10.1111/j.1539-6924.1998.tb00377.x
149. Clarke MG, Kennedy KP, MacDonagh RP. Discussing life expectancy with surgical patients: do patients want to know and how should this information be delivered? *BMC Med Inform Decis Mak*. 2008;8:24. DOI: 10.1186/1472-6947-8-24
150. Hoesseini A, Dronkers EAC, Sewnaik A, Hardillo JAU, Baatenburg de Jong RJ, Offerman MPJ. Head and neck cancer patients' preferences for individualized prognostic information: a focus group study. *BMC Cancer*. 2020;20(1):399. DOI: 10.1186/s12885-020-6554-8
151. Gett RM, Cooray AR, Gold D, Danta M. Evaluating informed consent for colonoscopy. *Surg Laparosc Endosc Percutan Tech*. 2014;24(4):345–52. DOI: 10.1097/SLE.0000000000000053
152. Kiely BE, McCaughan G, Christodoulou S, et al. Using scenarios to explain life expectancy in advanced cancer: attitudes of people with a cancer experience. *Support Care Cancer*. 2013;21(2):369–76. DOI: 10.1007/s00520-012-1526-4
153. Hagerty RG, Butow PN, Ellis PA, et al. Cancer patient preferences for communication of prognosis in the metastatic setting. *J Clin Oncol*. 2004;22(9):1721–30.
154. Hill S, Spink J, Cadilhac D, et al. Absolute risk representation in cardiovascular disease prevention: comprehension and preferences of health care consumers and general practitioners involved in a focus group study. *BMC Public Health*. 2010;10:108. DOI: 10.1186/1471-2458-10-108
155. Carey M, Herrmann A, Hall A, Mansfield E, Fakes K. Exploring health literacy and preferences for risk communication among medical oncology patients. *PLoS One*. 2018;13(9):e0203988. DOI: 10.1371/journal.pone.0203988
156. Freeman TR, Bass MJ. Risk language preferred by mothers in considering a hypothetical new vaccine for their children. *CMAJ*. 1992;147(7):1013–7.
157. Hallowell N, Statham H, Murton F, Green J, Richards M. "Talking about chance": the presentation of risk information during genetic counseling for breast and ovarian cancer. *J Genet Couns*. 1997;6(3):269–86. DOI: 10.1023/A:1025624221369
158. Bloch EV, DiSalvo M, Hall BD, Epstein CJ. Alternative ways of presenting empiric risks. *Birth Defects Orig Artic Ser*. 1979;15(5C):233–44.
159. Kaplowitz SA, Campo S, Chiu WT. Cancer patients' desires for communication of prognosis information. *Health Commun*. 2002;14(2):221–41. DOI: 10.1207/S15327027HC1402_4
160. Garcia-Retamero R, Cokely ET. Simple but powerful health messages for increasing condom use in young adults. *J Sex Res*. 2015;52(1):30–42. DOI: 10.1080/00224499.2013.806647
161. Dorval M, Bouchard K, Chiquette J, et al. A focus group study on breast cancer risk presentation: one format does not fit all. *Eur J Hum Genet*. 2013;21(7):719–24. DOI: 10.1038/ejhg.2012.248
162. Corso PS, Hammitt JK, Graham JD. Valuing mortality-risk reduction: using visual aids to improve the validity of contingent valuation. *J Risk Uncertain*. 2001;23(2):165–84. DOI: 10.1023/A:1011184119153
163. Schonlau M, Peters E. Comprehension of graphs and tables depend on the task: empirical evidence from two Web-based studies. *Stat Politics Policy*. 2012;3:1–35.
164. Betsch C, Renkewitz F, Haase N. Effect of narrative reports about vaccine adverse events and bias-awareness disclaimers on vaccine decisions: a simulation of an online patient social Network. *Med Decis Making*. 2013;33(1):14–25. DOI: 10.1177/0272989X12452342
165. Pighin S, Bonnefon J-F, Savadori L. Overcoming number numbness in prenatal risk communication. *Prenat Diagn*. 2011;31(8):809–13. DOI: 10.1002/pd.2771
166. Johnson BB. Communicating air quality information: experimental evaluation of alternative formats. *Risk Anal*. 2003;23(1):91–103. DOI: 10.1111/1539-6924.00292
167. Keller C, Siegrist M, Visschers V. Effect of risk ladder format on risk perception in high- and low-numerate individuals. *Risk Anal*. 2009;29(9):1255–64. DOI: 10.1111/j.1539-6924.2009.01261.x
168. Fischhoff B, Wong-Parodi G, Garfin DR, Holman EA, Silver RC. Public understanding of ebola risks: mastering an unfamiliar threat. *Risk Anal*. 2018;38:71–83. DOI: 10.1111/risa.12794
169. Figuerias M, Ghorayeb J, Coutinho M, Maroco J, Thomas J. Levels of trust in information sources as a predictor of protective health behaviors during COVID-19 pandemic: a UAE cross-sectional study. *Front Psychol*. 2021;12:633550. DOI: 10.3389/fpsyg.2021.633550
170. Hocevar K, Metzger M, Flanagin A.Source credibility, expertise, and trust in health and risk messaging. In: Nussbaum J, ed. *Oxford Research Encyclopedias: Communication*. Oxford (UK): Oxford University Press; 2017.
171. Becker MH. The health belief model and personal health behavior. *Health Educ Monogr*. 1974;(2):324–508.
172. Witte K. Putting the fear back into fear appeals: the extended parallel process model. *Commun Monogr*. 1992;59:329–49.
173. Witte K, Allen M. A meta-analysis of fear appeals: implications for effective public health campaigns. *Health Educ Behav*. 2000;27(5):591–615.