Health professionals prefer to communicate risk-related numerical information using “1-in-X” ratios.

Miroslav Sirota[1], PhD; Marie Juanchich[1], PhD; Dafina Petrova[2], PhD; Rocio Garcia-Retamero[2], PhD; Lukasz Walasek[3], PhD; Sudeep Bhatia[4], PhD

[1]Department of Psychology, University of Essex

[2]Department of Experimental Psychology, Mind, Brain, and Behavior Research Center, University of Granada

[3]Department of Psychology, University of Warwick

[4]Department of Psychology, University of Pennsylvania

Author Note

Correspondence concerning this paper should be addressed to Miroslav Sirota, Department of Psychology, University of Essex, Wivenhoe Park, Colchester, CO4 3SQ, United Kingdom. Email: msirota@essex.ac.uk, Phone: (+44) 1206 874 229.

Financial support for Study 1 was provided by a grant awarded to Miroslav Sirota by the Department of Psychology, University of Essex. Financial support for Study 3 was partially provided by Ministerio de Economía y Competitividad (Spain) (PSI2014-51842-R to Rocio Garcia-Retamero and Dafina Petrova). The funding agreement ensured the authors' independence in designing the study, interpreting the data and writing and publishing the report.

Supplementary Materials and all data sets are available at:
https://osf.io/9eb66/?view_only=beb69bc3c4e54a31aba5484ebe8fb1e7

Note: This is an accepted manuscript version for an article to be published in the journal *Medical Decision Making* but it might differ slightly from the final version.

ABSTRACT

Previous research showed that format effects such as the "1-in-X" effect – whereby "1-in-X" ratios lead to a higher perceived probability than "N-in-N*X" ratios – alter perceptions of medical probabilities. We do not know, however, how prevalent this effect is in practice – whether health professionals often use "1-in-X" ratios. We assembled four different sources of evidence, involving experimental work and corpus studies, to examine the use of "1-in-X" and other numerical formats quantifying probability. Our results revealed that the use of "1-in-X" ratios is prevalent and that health professionals prefer this format compared with other numerical formats (i.e., the "N-in-N*X", %, and decimal formats). In Study 1, UK family physicians preferred to communicate prenatal risk using a "1-in-X" ratio (80.4%, n = 131) across different risk levels and regardless of patients' numeracy levels. In Study 2, a sample from the UK adult population (n = 203), reported that most GPs (60.6%) preferred to use "1-in-X" ratios compared with other formats. In Study 3, "1-in-X" ratios were the most commonly used format in a set of randomly sampled drug leaflets describing the risk of side effects (100%, n = 94). In Study 4, the "1-in-X" format was the most commonly used numerical expression of medical probabilities or frequencies on the UK's NHS website (45.7%, n = 2,469 sentences). The prevalent use of "1-in-X" ratios magnifies the chances of increased subjective probability. Further research should establish clinical significance of the "1-in-X" effect.

Keywords: format preference, "1-in-X" effect, "1-in-X" ratio, subjective probability, risk communication

INTRODUCTION

Health professionals regularly communicate health information, including details relating to probability and frequency, in order to promote informed and shared decision-making [1, 2]. Patients, however, do not always adequately represent, process and understand such numerical information [3, 4]. Sometimes, as is the case with the “1-in-X” effect, mathematically equivalent information might mean different things to different people. According to the “1-in-X” effect, people often perceive the same objective probability presented in “1-in-X” formats (e.g., “1 in 10”) as being subjectively higher and more worrying than in “N-in-N*X” formats (e.g., “10 in 100”), or percentages (e.g., “10%”) [5, 6]. The effect has been replicated across different samples, cultures and scenarios [5-7], including pregnant women waiting for prenatal screenings [8]. People with different levels of numeracy, education and cognitive reflection ability are similarly prone to this effect [6]. The “1-in-X” effect also extends to positive outcomes (e.g., the chance of winning the lottery) [7]. Ratio artefacts cannot account for the “1-in-X” effect [5]. Finally, the effect does not seem to be an instance of “group diffusion” [9]. Evidence suggests the higher probability in multiple ratio comparisons is a function of the “1” in the numerator rather than a function of the increase in the size of the denominator [6].

Some important questions, however, remain unanswered. Theoretically speaking, we do not understand the mechanisms and boundary conditions of the effect [5-7]. From a practical viewpoint, we do not know (i) whether and to what extent the “1-in-X” format is distorting the subjective probability and (ii) how prevalent such ratios are in practice. It is important to establish this to provide robust evidence supporting (or opposing) the calls for banning the format from risk communication [10, 11]. The first issue, not addressed in this manuscript, concerns *accuracy*. The “1-in-X” effect causes higher subjective probability estimates but this does not necessarily mean the format reduces the accuracy of risk estimation (e.g., “N-in-N*X” formats might lead to underestimation of probabilities). Some findings, however, suggest reduced accuracy. First, the “1-in-X” format increases subjective probability estimates not only when compared with the “N-in-N*X” format, but when compared with other numerical formats such as percentages and visual representations [5, 6]. It seems unlikely that all the compared formats lead to underestimation. Second, a meta-analysis of four experiments showed that participants receiving the “1-in-X” ratios overestimated probabilities – using probability and arbitrary frequency scales – presented alongside or immediately after the risk information was displayed; the overestimation was significantly higher than that generated by the “N-in-N*X” ratios [12].

The second issue, addressed in this manuscript, concerns the *prevalence* of the “1-in-X” ratios. If the “1-in-X” ratio distorts subjective probability estimates, then the prevalence of this format in risk communication will indicate how common this distortion is. Around seven instances of the “1-in-X” ratio communication are needed for one probability estimate to be higher than that based on an “N-in-N*X” ratio (Sirota et al. estimated the meta-analytical effect of Hedges’ g = 0.42 [6], which is an equivalent Number Needed to Treat of 7.3). The prevalence of the ratio will determine its cumulative impact on accuracy at a population level. For example, assuming that only a single risk is described per consultation, and there are 340 million consultations with GPs in England annually, [13] then only a high prevalence would scale up the overestimation. If the format used is uncommon, such as in 0.01% of instances, the perceived probability will only be elevated in around 1 in 73,000 consultations (i.e., in every seventh consultation of the 4,657 consultations where the “1-in-X” ratio was used out of a total of 340,000,000). If the use of this format is highly prevalent, such as in 99.9% of instances, around 1 in 7 consultations (or in every seventh consultation of the 339,660,000 consultations where the “1-in-X” ratio was used out of a total of 340,000,000) will yield a higher subjective probability.

The present research

We aimed to accumulate different sources of evidence on the format use and preference in diverse healthcare samples and cultures. We examined oral (Studies 1 and 2) and written (Studies 3 and 4) risk communication preferences in different health contexts. We examined risk communication about hypothetical medical conditions, elicited in controlled conditions (Studies 1 and 2) plus “real life” risk communication currently available to patients (Studies 3 and 4). In Study 1, we evaluated GPs’ format preferences for communicating the risk of prenatal screening. In the vignettes, we also manipulated the patients’ numeracy levels (i.e., by providing information about the numeracy test results of the patients) and risk levels to test the robustness of the GPs’ preferences. In Study 2, we assessed the patients’ lay intuitions regarding the GPs’ format preferences in a sample of UK adults with diverse education levels. This study was designed to conceptually replicate the findings in Study 1. In Study 3, we examined the format preference for communicating side effects in medication leaflets for the most frequently sold drugs in Spain (in 2014). In Study 4, we assessed the most commonly used format on the UK’s NHS website.

We hypothesized that the “1-in-X” format (e.g., 1 in 10) would be preferable to other numerical formats, such as “N-in-N*X” (e.g., 10 in 100), percentages (e.g., 10%), or decimals between 0 and 1 (e.g., 0.10) due to its reputation for being simple to understand [14, 15] in relation to anecdotal evidence of its extensive use [8]. Given this reputation, we also hypothesized that perceived low numeracy (Study 1) and education levels (Study 2) in patients would amplify the preference for the “1-in-X” format.

STUDY 1

Here, we studied GPs’ preferences for formats of risk communication in three prenatal screening scenarios. We used different probability magnitudes to assess the robustness of the preference and formulate different ratios to investigate whether the format, rather than the specific numerical quantity, influences their preferences. We also manipulated the level of numeracy in a hypothetical patient to whom the GPs communicated the information. We hypothesized that the GPs would prefer the “1-in-X” format across all the risk magnitudes and numeracy levels, more so when the patients were less numerate, as this format might be considered simpler to understand [14, 15].

Method

The task was completed by 131 GPs (61.1% of whom were male; 18.3% practiced in an inner city, 57.3% in an urban area and 24.4% in a rural area). Their years of experience ranged from 1 to 35 (M = 14.1, SD = 9.4 years). Participants received £7 for completing a short online questionnaire (comprising also other scenarios reported elsewhere [16]).

A professional research software company (Qualtrics) recruited the sample via an online panel. Only certified physicians currently practicing in the UK were eligible to participate (the response rate was 28.2%). We excluded responses from seven participants based on a priori exclusion criteria indicating careless responding [17]: failure to pass the instructional manipulation check and/or rushed completion (i.e., shorter than 240 seconds, which was one third of the median response time on the questionnaire for the first five valid responses).

We used a mixed 3(risk: low, medium, high) × 3(numeracy: low, baseline, high numeracy) experimental design, with risk as a within-subjects factor and numeracy as a between-subjects factor. The scenarios were presented randomly.

After giving informed consent and answering sociodemographic questions (gender, experience and practice location), the GPs diagnosed and managed some patient vignettes unrelated to this research [16] and answered three questions focused on their risk format preference for informing three pregnant patients seeking advice regarding antenatal Down syndrome screening. The patients were described as having low, high or undetermined numeracy levels [18]. For instance, the low numeracy level was manipulated as follows: “As part of a new initiative by the NHS, Alice took a short online test before making an appointment. This test showed that she has a very low understanding of numbers and probabilities” (see full materials in the Appendix). We manipulated the risk of having a baby affected by Down syndrome by increasing the age of the patient (1%, 2%, and 12.5% for a 40-, 43- and 49-year-old woman, respectively). These risk magnitudes were based on actual patient data [8]. The GPs selected one out of four possible ways to communicate the risk to the patient: “1-in-X”, “N-in-N*X”, probability expressed as a percentage (i.e., a scale of 0 to 100%), and probability expressed on a scale ranging from 0 to 1. For the low risk scenario the options were: “1 in 100”, “10 in 1,000”, “1%”, or “0.01”. The options were presented randomly.

Results

The GPss largely preferred the “1-in-X” format (80.4% of choices) to communicate the risk of having a child affected by Down syndrome, followed by the risk expressed as a percentage (14.5% of choices), the “N-in-N*X” format (3.8% of choices), and the 0-1 probability scale (1.3% of choices). The difference between formats was statistically significant, χ^2 (3) = 658.96, $p < .001$, $V = 0.75$. The GPs’ preferences were consistent across the scenarios: most GPs (80.2%) selected the same format for all three of them. The patients’ numeracy levels did not affect the format preference, with the “1-in-X” format being the most common preference across the conditions (Figure 1). The results of a generalized estimating equations multinomial logistic regression with numeracy as a factor, scenarios with different risk levels as a factor and repeated measures and preference for one of the four choices as the dependent variable confirmed this results pattern. We found no significant effect of risk magnitudes, Wald $\chi^2(2) = 3.98$, $p = .136$, of numeracy levels, Wald $\chi^2(2) = .72$, $p = .696$, nor a significant interaction between numeracy levels and risk magnitude, Wald $\chi^2(4) = .2.37$, $p = .667$. Thus, our second hypothesis about a more pronounced preference for the “1-in-X” ratio in the low numeracy condition was not supported by our data.

Insert Figure 1 around here

STUDY 2

In Study 1, GPs preferred to use “1-in-X” ratios to communicate risk, whether the risk was 1%, 2%, or 12.5% and regardless of patients’ numeracy levels. A critical reader, however, could object that the pattern of preference does not reflect what the physicians are doing, but what they think they should be doing. In Study 2, we addressed this issue by eliciting intuitions about GPs’ preferences for risk communication in a sample of the UK adult population. If the GPs’ pattern of preference reported in Study 1 describes actual GP behavior, then we should expect a similar pattern in the risk communication format among patients. These intuitions may vary with patients’ education levels, as suggested by the “simplicity reputation” of the 1-in-X ratio [14].

Method

A sample of 203 participants recruited via an online panel (Prolific Academic) completed an online questionnaire with the following inclusion criteria: i) participants resided in the UK, ii) participants successfully completed at least 90% of past online panel tasks; and iii) participants passed the instructional manipulation check. The participants were reimbursed (£1 for a 10-minute questionnaire). They were mostly women (56.7%), mostly with a high school education (34.0%) or an undergraduate degree (43.3%), mostly consisted of working professionals (30.5%) and unemployed, including students and homemakers (26.6%). Their ages ranged from 18 to 73 (*M* = 34.4, *SD* = 12.1 years).

After giving informed consent and solving unrelated word problems, the participants read a scenario adopted from Study 1, involving a 40-year-old woman seeking advice regarding antenatal screening from a GP. The participants selected the option most likely to be used by the GP to communicate the risk of having a child with Down syndrome (i.e., “1 in 100”, “10 in 1,000”, “1%”, or “0.01”). The participants were instructed to infer the physician’s preference based on their prior experience and/or beliefs. (See instructions and scenario in the Appendix.) Finally, the participants answered some sociodemographic questions and were debriefed.

Results

The pattern of preference among the participants was similar to that of the physicians. Most of them selected the “1-in-X” ratio (60.6%), followed by percentages (28.1%), the “N-in-N*X” ratio (5.9%) and probabilities expressed on a 0-1 scale (5.4%). The difference between the information formats was significant, $\chi^2(3) = 164.35$, $p < .001$, $V = 0.52$. These intuitions were consistent with the findings from Study 1 that GPs mostly use “1-in-X” ratios when communicating health risks. We observed only a slight variation in predicted preferences according to education level (with vs. without higher education) as depicted in Figure 2. This difference was statistically insignificant, $\chi^2(3) = 2.90$, $p = .407$, $V = 0.12$. We did not support our second expectation that these intuitions would vary with patients’ education levels.

Insert Figure 2 around here

STUDY 3

In the first two studies, we assessed GPs’ preferences and participants’ intuitions about

GPs' preferences for different risk communication formats. We focused on a *hypothetical* scenario whereby physicians *communicated* risks to patients. In Study 3, we investigated the "1-in-X" format in real life written health communication i.e., drug leaflets. To enhance the generalizability, we investigated this issue in Spain. We expected the "1-in-X" format would be the most common. For side effects, we expected the preference for "1 in X" ratios to be stronger given that this format is endorsed and used as an example in EU guidelines [19].

Method

To ascertain the most commonly bought drugs, we determined the most often sold active ingredients as reported by the Spanish National Health Service (n = 26 unique active ingredients) [20]. We downloaded a list of the commercialized medications containing these active ingredients (N = 1,777). We then randomly selected five medications per ingredient while maximizing the variability of risk-reporting practices (i.e., avoiding multiple drugs from the same producer). If an ingredient was found in fewer than five medications, all of them were selected. This resulted in 112 medications from 62 pharmaceutical companies. For seven of these, we were unable to obtain a drug leaflet and 13 did not use numerical expressions to communicate the frequency of side effects. The final sample contained 92 drug leaflets. The details are described in the Appendix.

Our analysis focused on the leaflets' "undesirable side effects". We coded the format in which the frequency of adverse events was communicated as "1-in-X", "N-in-N*X", percentage, 0-1 probability scale or "other". We also coded whether a different format was used for interval expressions (e.g., an interval consisting of two point estimates might have used consistent or inconsistent formats such as "between 1 in 10 and 1 in 100 patients" and "between 1 and 10 in

100 patients”). The formats used to present the side effects were homogeneous: they were usually communicated using the same format, in tables or grouped paragraphs. Both the number and type of side effects varied so we therefore used leaflets as a unit of analysis.

The leaflets were coded by one of the authors (DP) and another (GR) coded 20 randomly-selected leaflets to assess the reliability of the coding. The inter-rater reliability by computing the percentage of agreement was very good (> 95%) on all variables except for the use of intervals (75%). An inspection showed that disagreements on intervals were due to different definitions used by the coders. The definition was refined to the one specified above and coding was repeated for this variable, resulting in complete agreement.

Results

All 92 leaflets communicated numerical information using some version of the “1-in-X” format. Twenty-two (24%) leaflets used an additional numerical format: two (2%) used percentages, each on a single occasion, and 20 (22%) used the “N-in-N*X” format (the latter only in intervals with “1-in-X”, see below). Of the 92 (100%) leaflets using the “1-in-X” format, 60 (65%) communicated the frequency of events using point-estimates (e.g., about 1 in 10 patients; less than 1 in 100 patients). The remaining 32 (35%) used intervals. Of these, 12 (38%) communicated the intervals using only the “1-in-X” format. This means the interval was composed of two “1-in-X” fractions with different denominators (e.g., >1/100 to <1/10; in at least 1 in 1.000 and fewer than 1 in 100 patients; at least 1 in 100 but fewer than 1 in 10 patients). The remaining 20 (63%) leaflets communicated the intervals using a combination of the “1-in-X” and “N-in-N*X” formats. These leaflets contained the following risk intervals: between 1 and 10 out of 100 (*n* = 19; 95%), between 1 and 10 out of 1,000 (*n* = 17; 85%), between 1 and 100 out of 10,000 (*n* = 1; 5%), and between 1 and 10 in 10,000 (*n* = 16; 80%).

STUDY 4

In Study 4, we extended our investigation by performing a corpus analysis of the NHS website, the main source of health information for the UK public. We expected the “1-in-X” format to be the most prevalent one used to communicate the frequency/probability information relevant to risk. We also expected the “1-in-X” format to be used more often for smaller probabilities. This expectation relies on the ability of “1-in-X” ratios to express different levels of probability, which increases exponentially with the increasing size of the denominator, thus implicating lower probabilities. For instance, “1 in 2” expresses 50% probability and the closest lower ratio of “1 in 3” expresses 33.3% (i.e., a 16.6% decrease), whereas “1 in 78” expresses 1.28% and the closest lower ratio “1 in 77” expresses 1.30% (a 0.02% decrease). Furthermore, “1-in-X” cannot convey probabilities over 50% while using integers. Such restrictions do not apply to the other formats. We also explored the possibility of association of the format and severity of the outcome.

Method

We downloaded content from the NHS website and extracted phrases containing information about numerical probabilities/frequencies. We extracted 2,649 suitable sentences and manually coded whether these quantified a risk-related frequency and probability or something else, which reduced the database to 2,469 sentences. These details are described in the Appendix. The unit of analysis was one piece of numerical information (i.e., if an interval was provided, that interval was broken down into two one-point pieces of information).

We coded the mode of presentation of the format (i.e., numerals such as “1 in 3” or words such as “one in three”) and the probability magnitude conveyed. We also randomly selected 100 instances of occurrence for each of the formats (“1-in-X”, “N-in-N*X”, and percentages). We coded the outcomes of those 300 cases in the following categories: i/ risk of the consequences of a medical condition; ii/ positive consequence of a treatment (e.g., remission); iii/ negative consequence of a treatment (e.g., negative side effect), and iv/ risk of contracting an illness. These details are described in the Appendix.

Results

Probabilities were most commonly conveyed in the “1-in-X” format (45.7%), followed by percentages (40.6%), and the “N-in-N*X” format (13.7%). The difference between the proportions was significant, $\chi^2(2) = 438.52$, $p < .001$, $V = 0.30$; the proportion of the “1-in-X” format was significantly higher than that of the percentages, $\chi^2(1) = 7.57$, $p = .006$, $V = 0.06$. Thus, the “1-in-X” format was the most common format qualifying risk on the NHS website. Risk was expressed mostly using numerals (73.4%): percentages were exclusively expressed numerically, whereas the “1-in-X” ratios and “N-in-N*X” ratios were conveyed using either words or numerals (54.4% and 57.7% of numerals respectively).

For very small to medium probabilities, the “1-in-X” format was preferred over the other formats, whereas probabilities over 40% were mostly conveyed using percentages (Figure 3). The preference for a specific format to communicate different probability magnitudes was examined separately for probabilities below and above 50%. Very low to low probabilities were mainly conveyed using the “1-in-X” format and an increase in probability was associated with an increased use of percentages (Table 1). For probabilities over 50%, the analysis showed that an increased probability magnitude was related to an increased use of percentages. The variations in probabilities below 50% or over 50% were related to format preference, $\chi^2(10, N = 1919) = 288.50$, $p < .001$, $V = .27$ and $\chi^2(10, N = 550) = 62.42$, $p < .001$, $V = .34$, respectively.

Insert Figure 3 around here

Insert Table 1 around here

We also explored the association between the format and the nature of the outcome (using a random sample of 300 sentences, 100 of each format). Most of the outcomes denoted the prevalence of an illness, followed by the negative consequence associated with a medical condition, the positive consequence of a treatment and the negative consequence of a treatment (see Table 2). The three formats were used to quantify the three negative outcomes approximately equally (around 25–40%), but the “1-in-X” format was used less often for the positive outcome describing the positive consequences of treatments, $\chi^2(6, N = 300) = 18.72$, $p = .005$, $V = .18$. Given that the outcome categories differed in the probability magnitude, $F(3, 299) = 19.65$, $p < .001$, $\eta^2_p = .17$ (Table 2, last column) and that the “1-in-X” format is unable to express probabilities higher than 50%, we only investigated probabilities below 50% (Table 3). In a series of multinomial regressions, we found that the four outcome categories did not predict the preference for the “1-in-X” format, $\chi^2(6) = 8.65$, $p = .194$; however when the outcome categories were reduced to negative vs. positive, negative outcomes were more often described using “1-in-X” ratios than “N-in-N*X” ratios or percentages, $\chi^2(2) = 7.17$, $p = .028$. The “N-in-N*X” format was roughly as likely to describe negative outcomes as the “1-in-X” format, $OR = 0.96$, 95% CI[0.1, 10.9], whereas the percentage format was preferred to the “1-in-X” format for describing positive outcomes, $OR = 5.9$, 95% CI[1.2, 29.4]. This relationship was reduced but not completely eliminated when adjusting for probability magnitude, $\chi^2(2) = 5.13$, $p = .077$.

Insert Table 2 around here

Insert Table 3 around here

GENERAL DISCUSSION

In the four studies we found converging and robust evidence that “1-in-X” ratios are the most prevalent expressions of numerical risk used in healthcare communication. We employed traditional experimental methods (with GPs and the general adult population) as well as analyses of written materials and websites. The “1-in-X” ratio preference was not limited to a specific communication context, since we found it in both written and oral communication; among different healthcare communicators, including GPs, pharmaceutical companies and the UK’s NHS; in different domains such as communication of screening risks, risks associated with side effects of medications or diagnosis descriptions, and in different cultures (the UK and Spain).

The current findings extend the research on the “1-in-X” effect, which has demonstrated that “1-in-X” ratios lead to subjectively higher [5-7] and less accurate probability estimates [12]. It seems plausible to assume that the common use of “1-in-X” ratios has some systematic roots embedded in communication guidelines [19] or the education of GPs. Future research should identify the roots of this practice and the conditions under which the preference is reinforced or weakened. The common use of the “1-in-X” ratio implies that the “1-in-X” effect is highly prevalent too. This raises the question of clinical meaningfulness of the “1-in-X” effect: Future studies should establish to what extent different formats alter related decision-making. We believe this is possible. First, health-behavior models assume that the probability perception affects health behavior (e.g., [21]). It is unclear why this would not be the case here. Second, limited evidence has demonstrated that the effect can transfer into related decision-making: participants presented with a risk of contracting a disease in a “1-in-X” ratio were more likely to cancel a trip to a country where the disease could be contracted than in an “N-in-N*X” ratio [12].

We also found the “1-in-X” ratio dominates risk communication, regardless of patients’ numeracy levels (Study 1), the level of the communicated risk within the possible ranges of the “1-in-X” ratios (i.e., Study 1) or education (Study 2). None of these variables contributed to substantial variation in preference for the “1-in-X” ratios. This could mean patients with low as well as high numeracy levels are equally likely to be exposed to the format. It is an open question whether these two groups are equally affected by it – in terms of perception accuracy and related decisions.

Although we did not observe shifts in the preferences of the GPs as a function of the probability magnitude levels (Study 1), the corpus analysis of the NHS website demonstrated that “1-in-X” ratios were increasingly associated with low probability outcomes, especially with those below 1%, and mostly with negative outcomes (Study 4). This prompts several questions regarding the mechanisms underlying the effect. We will outline two non-exclusive possibilities. First, we found that “1-in-X” ratios were strongly associated with low probabilities. Given that people tend to overweight very low probabilities [22, 23], we can speculate that the overweighting of the very small probabilities commonly expressed with “1-in-X” ratios is activated by the “1-in-X” format itself. Such activation would lead to a spill-over effect and all “1-in-X” ratios – even those conveying higher probabilities – would be overweighted. Second, we found that “1-in-X” formats are used more often for negative medical outcomes than for positive ones. Given that people tend to overestimate the probability of severe outcomes [24, 25], we can also hypothesize that due to strong co-occurrence of severe outcomes and the “1-in-X” format, the “1-in-X” format activates the overestimation mechanism. This would again lead to spill-over effects. Future research should test these hypotheses.

Despite the pattern’s robust nature, we should consider some limitations of the present studies. Our analysis of GPs’ communication (Studies 1 and 2) focused only on one type of situation (the risk of undertaking a screening test). It is possible that format preference is, to some extent, domain-specific. Thus, it would be wrong to generalize this to all communication situations without further research. Furthermore, it could be argued that people’s intuitions about GPs’ preferences for the “1-in-X” ratio do not reflect the GPs’ practice but rather people’s own preference for the format. This is plausible; however we argue that our conclusion should remain unaffected. GPs use this format frequently and their preference is either an expression of people’s preference for the format or some domain-specific reasons. In Study 3, we focused on drug leaflets. However, future research should consider other written materials used for communicating with the public e.g., vaccination leaflets. Different leaflets aim to achieve different health-related goals. Communicators may cleverly use different formats of risk communication to best achieve their goals whilst telling the truth. In Study 4, we focused on one provider, but different providers may favor different formats. Finally, we identified a high prevalence of the “1-in-X” format among numerical formats; verbal representations of risk might be more common in practice.

To conclude, we found that the “1-in-X” format is the most prevalent numerical format among commonly used numerical quantifications of probability and frequency of risk. This preference spans across different domains, types of communication and health communicators. Prior work suggests that the “1-in-X” ratio leads to a higher subjective probability and the overestimation of actual probability. Future studies should assess to what extent this effect is clinically meaningful but, in light of the present results, caution is warranted.

REFERENCES

[1] Epstein RM, Alper BS, Quill TE. Communicating evidence for participatory decision making. JAMA. 2004; 291(19):2359-66.

[2] Paling J. Strategies to help patients understand risks. BMJ: British Medical Journal. 2003; 327(7417):745-8.

[3] Fagerlin A, Zikmund-Fisher BJ, Ubel PA. Helping patients decide: Ten steps to better risk communication. JNCI: Journal of the National Cancer Institute. 2011; 103(19):1436-43.

[4] Gigerenzer G, Gaissmaier W, Kurz-Milcke E, Schwartz LM, Woloshin S. Helping doctors and patients make sense of health statistics. Psychological Science in the Public Interest. 2007; 8(2):53-96.

[5] Pighin S, Savadori L, Barilli E, Cremonesi L, Ferrari M, Bonnefon JF. The 1-in-X Effect on the subjective assessment of medical probabilities. Medical Decision Making. 2011; 31(5):721-9.

[6] 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. Medical Decision Making. 2014; 34(4):419-29.

[7] Oudhoff JP, Timmermans DR. The effect of different graphical and numerical likelihood formats on perception of likelihood and choice. Medical Decision Making. 2015; 35(4):487-500.

[8] Pighin S, Savadori L, Barilli E, Galbiati S, Smid M, Ferrari M, et al. Communicating Down syndrome risk according to maternal age:“1-in-X” effect on perceived risk. Prenatal Diagnosis. 2015; 35(8):777-82.

[9] Yamaguchi S. Biased risk perceptions among Japanese: Illusion of interdependence among risk companions. Asian Journal of Social Psychology. 1998; 1(2):117-31.

[10] Zikmund-Fisher BJ. Time to retire the 1-in-X risk format. Medical Decision Making. 2011; 31(5):703-4.

[11] Zikmund-Fisher BJ. Continued use of 1-in-X risk communications is a systemic problem. Medical Decision Making. 2013:0272989X13516198.

[12] Sirota M, Juanchich M, Bonnefon JF. "1-in-X" ratios lead to medical probability overestimation. *16th Biennial European Meeting of the Society for Medical Decision Making*. London, UK 2016.

[13] NHS England. Improving general practice - A call to action. Evidence Pack. Aug, 2013-14.

[14] Schapira MM, Nattinger AB, McHorney CA. Frequency or probability? A qualitative study of risk communication formats used in health care. Medical Decision Making. 2001; 21(6):459-67.

[15] Brase GL. Which statistical formats facilitate what decisions? The perception and influence of different statistical information formats. Journal of Behavioral Decision Making. 2002; 15(5):381-401.

[16] Sirota M, Round T, Samaranayaka S, Kostopoulou O. Expectations for antibiotics increase their prescribing: Causal evidence about localized impact. Health Psychol. 2017; 36(4):402-9.

[17] Meade AW, Craig SB. Identifying careless responses in survey data. Psychol Methods. 2012; 17(3):437-55.

[18] Petrova D, Kostopoulou O, Delaney B, Cokely E, Garcia-Retamero R. Strengths and gaps in physicians' risk communication: A scenario study of the influence of numeracy on cancer screening communication. Medical Decision Making. 2017.

[19] European Commision. A guideline on summary of product characteristics. 2009 01/02/2017]Available from: http://ec.europa.eu/health//sites/health/files/files/eudralex/vol-2/c/smpc_guideline_rev2_en.pdf

[20] Spanish Ministry of Health Social Services and Equality. Informe Anual del Sistema Nacional de Salud 2015 [Annual Report of the National Health System 2015]. Ministerio de Sanidad, Servicios Sociales e Igualdad 2015.

[21] Waters EA, McQueen A, Cameron LD. Perceived risk and health risk communication. The Routledge Handbook of Language and Health Communication. 2014:47.

[22] Burns Z, Chiu A, Wu G, Cochran JJ, Cox LA, Keskinocak P, et al. Overweighting of small probabilities. *Wiley Encyclopedia of Operations Research and Management Science*: John Wiley & Sons, Inc. 2010.

[23] Kahneman D, Tversky A. Prospect theory: An analysis of decision under risk. Econometrica: Journal of The Econometric Society. 1979:263-91.

[24] Harris AJL, Corner A. Communicating environmental risks: Clarifying the severity effect in interpretations of verbal probability expressions. Journal of Experimental Psychology-Learning Memory and Cognition. 2011; 37(6):1571-8.

[25] Mazur DJ, Merz JF. How age, outcome severity, and scale influence general medicine clinic patients' interpretations of verbal probability terms. Journal of General Internal Medicine. 1994; 9(5):268-71.

Table 1. Format preference as a function of different probability magnitudes.

			Format		
Probability magnitude	**1-in-X (*n* = 1129)**	**N-in-X*N (*n* = 338)**	**Percentage (*n* = 1002)**	**Total (N = 2469)**	**Total in % (N = 2469)**
<1%	**80.9%**	18.0%	1.1%	356	14.4%
1-10%	**56.3%**	5.4%	38.4%	871	35.3%
10-20%	**53.7%**	3.5%	42.8%	255	10.3%
20-30%	**55.9%**	6.2%	37.9%	145	5.9%
30-40%	**59.7%**	14.8%	25.6%	176	7.1%
40-50%	24.1%	9.5%	**66.4%**	116	4.7%
50-60%	–	34.8%	**65.2%**	46	1.9%
60-70%	–	33.3%	**66.7%**	93	3.8%
70-80%	–	**52.0%**	48.0%	148	6.0%
80-90%	–	25.9%	**74.1%**	162	6.6%
90-100%	–	5.9%	**94.1%**	101	4.1%

Note: For the probability magnitude categories, values matching the lower bound were excluded and values matching the upper bound were included.

Table 2. Format preference as a function of the outcome described.

Outcome and Format	**%**	**N** ***(n = 300)***	***Probability M (SD)***
Negative consequence of a treatment			
1-in-X	37.5%	6	
N-in-X*N	25.0%	4	
Percentage	37.5%	6	
Total	5.3%	16	16.6 (18.8)
Positive consequence of a treatment			
1-in-X	5.9%	2	
N-in-X*N	38.2%	13	
Percentage	55.9%	19	
Total	11.3%	34	68.2 (29.7)
Prevalence of a medical condition			
1-in-X	40.7%	61	
N-in-X*N	29.3%	44	
Percentage	30.0%	45	
Total	50.0%	150	25.1 (31.4)
Negative cons. of a medical condition			
1-in-X	31.0%	31	
N-in-X*N	39.0%	39	
Percentage	30.0%	30	
Total	33.3%	100	34.5 (31.6)

Note: For the probability magnitude categories, values matching the lower bound were excluded and values matching the upper bound were included.

Table 3. Format preference as a function of type of outcome and probability of occurrence (for outcomes up to 50% probability).

Outcome	**1-in-X**	**N-in-X*N**	**%**	***n***
Negative consequence of a treatment				
<1%	66.7%	33.3%	0.0%	3
1-10%	28.6%	14.3%	57.1%	7
10-20%	0.0%	0.0%	100.0%	1
20-30%	0.0%	0.0%	0.0%	0
30-40%	50.0%	50.0%	0.0%	4
40-50%	0.0%	0.0%	0.0%	0
Total	**40.0%**	**26.7%**	**33.3%**	**15**
Positive consequence of a treatment				
<1%	0.0%	0.0%	0.0%	0
1-10%	50.0%	0.0%	50.0%	4
10-20%	0.0%	0.0%	100.0%	1
20-30%	0.0%	0.0%	0.0%	0
30-40%	0.0%	33.3%	66.7%	3
40-50%	0.0%	0.0%	100.0%	2
Total	**20.0%**	**10.0%**	**70.0%**	**10**
Prevalence of a medical condition				
<1%	58.3%	41.7%	0.0%	36
1-10%	52.0%	14.0%	34.0%	50
10-20%	58.3%	16.7%	25.0%	12
20-30%	20.0%	20.0%	60.0%	5
30-40%	50.0%	30.0%	20.0%	10
40-50%	14.3%	0.0%	85.7%	7
Total	**50.8%**	**23.3%**	**25.8%**	**120**
Negative cons. of a medical condition				
<1%	71.4%	28.6%	0.0%	7
1-10%	46.9%	15.6%	37.5%	32
10-20%	33.3%	22.2%	44.4%	9
20-30%	50.0%	16.7%	33.3%	6
30-40%	44.4%	22.2%	33.3%	9
40-50%	11.1%	77.8%	11.1%	9
Total	**43.1%**	**26.4%**	**30.6%**	**72**
Total	**46.1%**	**24.0%**	**30.0%**	**217**

Note: For the probability magnitude categories, values matching the lower bound were excluded and values matching the upper bound were included.

FIGURE LEGENDS

Figure 1. Physicians' format preference ("1-in-X", "N-in-N*X", "0-100% scale", "0-1 scale") as a function of the numeracy of a patient (low, baseline, high) and of the level of risk (low, medium, high).

Note. The bars without numerical information (in %) indicate 0%.

Figure 2. Physicians' format preference as perceived by the general public ("1-in-X", "N-in-N*X", "0-100% scale", "0-1 scale") as a function of education.

Note. The bars without numerical information in % indicate 0%.

Figure 3. Format occurrence in the NHS website ("1-in-X", "N-in-N*X", "0-100% scale") as a function of probability categories.

Note. The bars without numerical information in % indicate 0%.

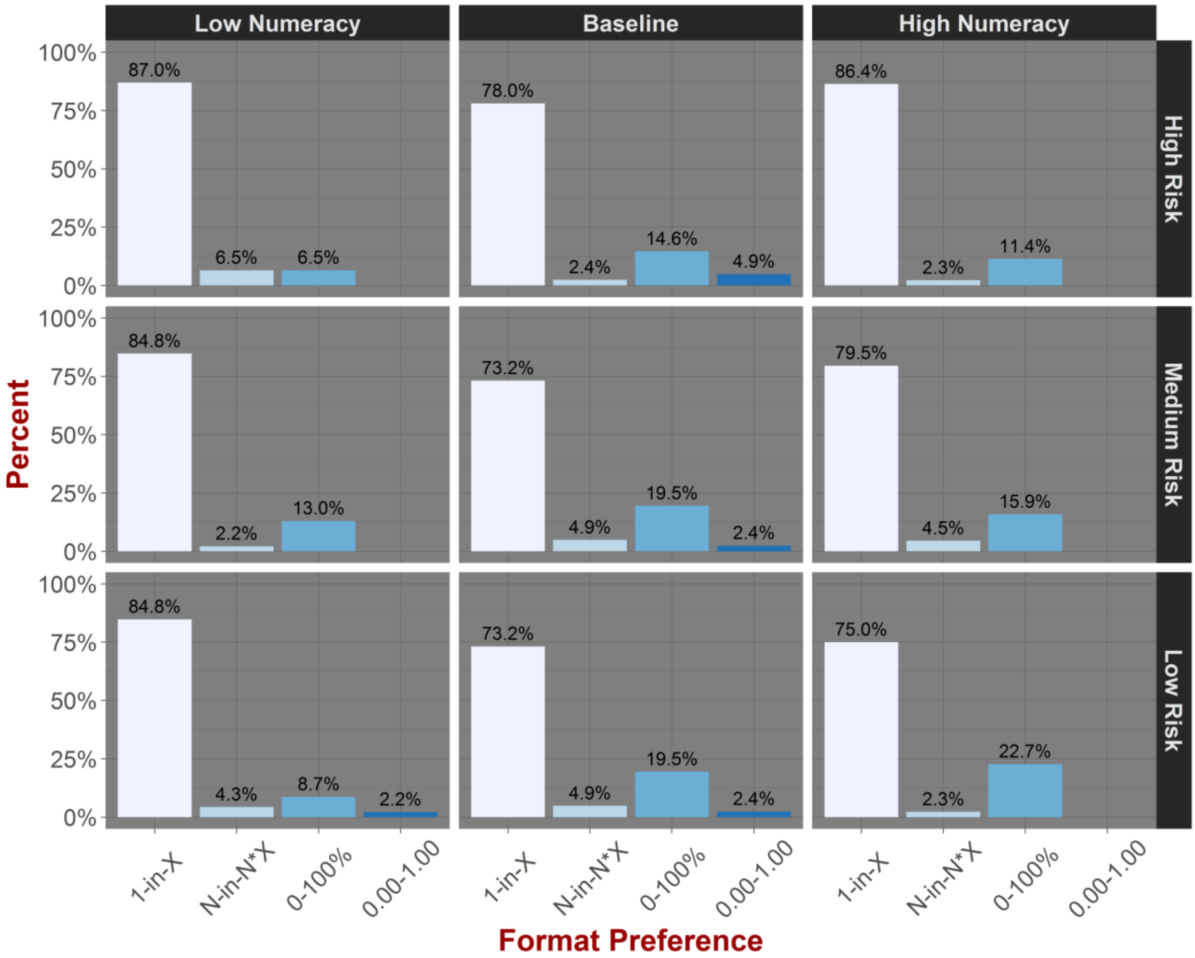
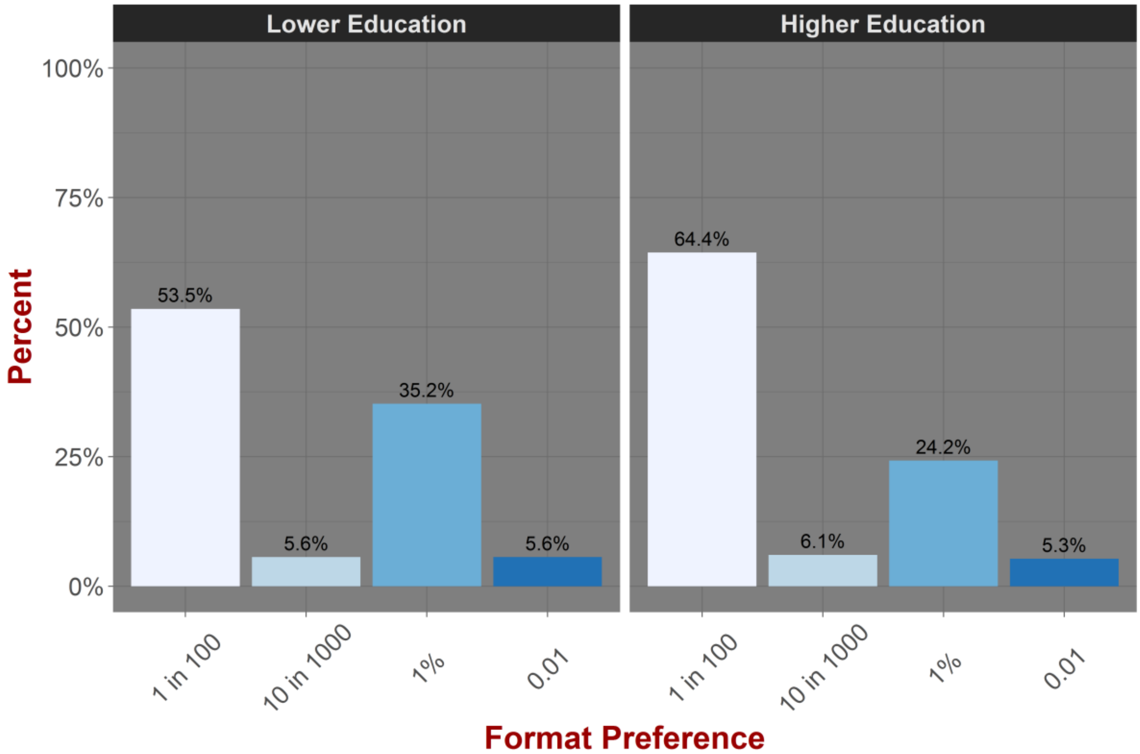
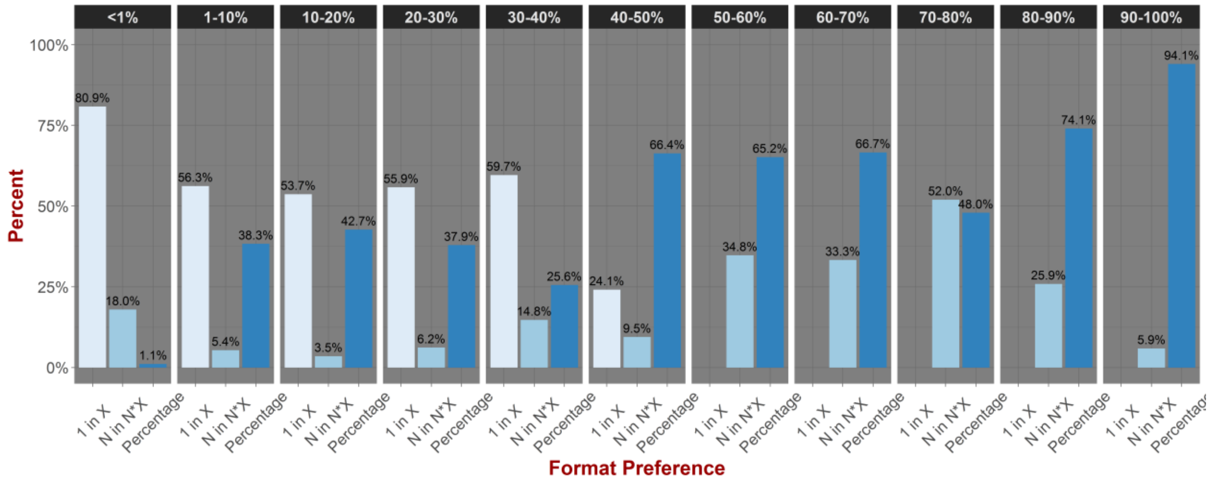

Low Numeracy
Baseline
High Numeracy
High Risk
Medium Risk
Low Risk
Percent
100%
75%
50%
25%
0%
87.0%
6.5%
6.5%
78.0%
2.4%
14.6%
4.9%
86.4%
2.3%
11.4%
84.8%
2.2%
13.0%
73.2%
4.9%
19.5%
2.4%
79.5%
4.5%
15.9%
84.8%
4.3%
8.7%
2.2%
73.2%
4.9%
19.5%
2.4%
75.0%
2.3%
22.7%
1-in-X
N-in-N*X
0-100%
0.00-1.00
Format Preference

Figure 1.



Lower Education
Higher Education
Percent
100%
75%
50%
25%
0%
53.5%
5.6%
35.2%
5.6%
64.4%
6.1%
24.2%
5.3%
1 in 100
10 in 1000
1%
0.01
1 in 100
10 in 1000
1%
0.01
Format Preference

Figure 2.



<1%
1-10%
10-20%
20-30%
30-40%
40-50%
50-60%
60-70%
70-80%
80-90%
90-100%
100%
75%
50%
25%
0%
Percent
80.9%
18.0%
1.1%
56.3%
5.4%
38.3%
53.7%
3.5%
42.7%
55.9%
6.2%
37.9%
59.7%
14.8%
25.6%
24.1%
9.5%
66.4%
34.8%
65.2%
33.3%
66.7%
52.0%
48.0%
25.9%
74.1%
5.9%
94.1%
1 in X
N in N*X
Percentage
Format Preference

Figure 3.

APPENDIX

Study 1: Instruction and Scenarios

Instruction

Now, you are going to see three different pregnant patients who are seeking advice regarding antenatal screening and will ask about their risk of having a child affected by Down's syndrome. We are interested in knowing how you would communicate the risk. We will provide you with accurate numerical information presented in different risk formats (e.g., ratio, percentage, probability and so on). There is no right or wrong answer, since all the formats convey the same numerical information; we are only interested in the format you prefer.

Low risk scenarios

Alice Roberts is a 40-year-old woman. She is in her 13th week of pregnancy. She has no significant illnesses; the pregnancy is without complications. *As part of a new initiative by the NHS, Alice took a short online test before making an appointment. This test showed that she has a very low understanding of numbers and probabilities. {This test showed that she has a very good understanding of numbers and probabilities.}*

Today, she came in for her first antenatal appointment. During the consultation, she asked about her risk of having a child affected by Down's syndrome. She wants to know what her chance is of having a child affected by Down's syndrome at her age.

"The chance of having a child affected by Down's syndrome in your age group is ….."

[select the numerical expression you would prefer to use]

1 in 100; 10 in 1000; 1%; 0.01

**Note*: Italics (added) indicate numeracy manipulations (low numeracy in the text; high numeracy in the curly brackets); the sentences describing the numeracy test were missing in the baseline condition.

Medium risk scenarios

Becky Williams is a 43-year-old woman. She is in her 13th week of pregnancy. She has no significant illnesses; the pregnancy is without complications. *As part of a new initiative by the NHS, Becky took a short online test before making an appointment. This test showed that she has a very low understanding of numbers and probabilities. {This test showed that she has a very good understanding of numbers and probabilities.}*

Today, she came in for her first antenatal appointment. During the consultation, she asked about her risk of having a child affected by Down's syndrome. She wants to know what her chance is of having a child affected by Down's syndrome at her age.

"The chance of having a child affected by Down's syndrome in your age group is ….."

[select the numerical expression you would prefer to use]

1 in 50; 20 in 1000; 2%; 0.02

**Note*: Italics (added) indicate numeracy manipulations (low numeracy in the text; high numeracy in the curly brackets); the sentences describing the numeracy test were missing in the baseline condition.

High risk scenarios

Emilia Wright is a 49-year-old woman. She is in her 13th week of pregnancy. She has no significant illnesses; the pregnancy is without complications. *As part of a new initiative by the NHS, Emilia took a short online test before making an appointment. This test showed that she has a very low understanding of numbers and probabilities. {This test showed that she has a very good understanding of numbers and probabilities.}*

Today, she came in for her first antenatal appointment. During the consultation, she asked about her risk of having a child affected by Down’s syndrome. She wants to know what her chance is of having a child affected by Down’s syndrome at her age.

“The chance of having a child affected by Down’s syndrome in your age group is …..”

[select the numerical expression you would prefer to use]

1 in 8; 125 in 1000; 12.5%; 0.125

**Note*: Italics (added) indicate numeracy manipulations (low numeracy in the text; high numeracy in the curly brackets); the sentences describing the numeracy test were missing in the baseline condition.

Study 2: Instruction and Scenario

Instruction

Now, you are going to read about a pregnant woman who is seeking advice regarding antenatal screening (she will ask a GP about her risk of having a child affected by Down's syndrome). We are interested in knowing how a GP (general practitioner) would communicate the risk to the patient based on your experience and/or beliefs. We will provide you with accurate numerical information presented in different risk formats (e.g., ratio, percentage, probability and so on). There is no right or wrong answer, since all the formats convey the same numerical information; we are only interested in the format you think a GP would prefer to use.

Scenario [adopted from Study 1]

Alice Roberts is a 40-year-old woman. She is in her 13th week of pregnancy. She has no significant illnesses; the pregnancy is without complications.

Today, she came in for her first antenatal appointment. During the consultation, she asked about her risk of having a child affected by Down's syndrome. She wants to know what her chance is of having a child affected by Down's syndrome at her age.

Her GP would probably say: "The chance of having a child affected by Down's syndrome in your age group is ….."

[select the numerical expression you believe the GP would probably prefer to use]

1 in 100; 10 in 1000; 1%; 0.01

Study 3: Method (details)

To obtain a list of the most commonly bought drugs, we first determined the most often sold active ingredients as reported by the National Health Service in Spain (*n* = 26 unique active ingredients) [19]. The report provided data (for the year 2014) about the most often consumed 15 active ingredients according to (a) number of packages sold (responsible for 34% of the total drug consumption) and (b) cost (responsible for 25% of the total drug consumption). We extracted 26 unique active ingredients across the two lists. The highest ranking ingredients included medications used to treat acid reflux, pain, fever, high lipids, asthma, epilepsy, anxiety and abnormal blood sugar levels, among others.

We downloaded a list of all the authorized and commercialized medications containing the 26 active ingredients from the website of the Spanish Agency for Medications and Health Products (www.aemps.gob.es). This resulted in a list of 1,777 medications sold by 170 pharmaceutical companies. We randomly selected five medications per active ingredient. In order to capture the variety of possible risk communication practices, if two of the selected medications were sold by the same pharmaceutical company, we replaced the second one with a randomly selected one from a different pharmaceutical company. Whenever there were fewer than five medications containing a certain active ingredient, the same rule was followed and only one medication per pharmaceutical company was selected. These steps left us with a total of 112 medications, 7 of which did not have a drug leaflet that could be accessed. Thus the final sample for review contained 105 drug leaflets from 62 pharmaceutical companies. In addition, 13 leaflets did not use numerical expressions to communicate the frequency of side effects. Thus, the final sample for review contained 92 drug leaflets.

Study 4: Method (details)

Data collection and extraction

We downloaded content from the UK National Health Service (hereafter, NHS) website. Specifically, we first downloaded the names of all the conditions listed on the NHS website (http://www.nhs.uk/Conditions/Pages/hub.aspx). For each individual condition we extracted all the text from its main page (i.e., the page that the condition's link directs to) as well as from all the connected subpages. These subpages predominantly include information about clinical trials, treatment, recovery, symptoms, causes, prevention, diagnosis and community. However, not every subpage was present for every condition. We extracted and saved all the html scripts from all the available pages. For each html file, we first extracted the paragraphs (using the <p> html tag), which we then divided into individual sentences. We processed the text by removing any non-English symbols, links to other websites and white spaces. Finally, we searched through each document, identifying sentences in which one of the capturing phrases occurred.

We then selected the capturing phrases that featured numerical information conveying probability/frequency information relevant to risk. First, we generated a list of phrases that conveyed information about proportions. We used numbers ranging from zero to 100 and then all the prominent whole numbers between 1,000 and 9,000, 10,000 and 90,000 and 100,000 and 900,000, also including 1,000,000. Both verbal and nonverbal expressions of these numbers were paired with the following words: "woman", "man", "child", "person", "family", "patient" and the plural version of these words when appropriate. These expressions were either followed by "out of" or "in", creating two sets of possible proportions (e.g., "1 in", "two patients out of", "one family in"). We only used exact and complete matches when scanning the NHS text data.

For percentages, we used the “%” symbol, as well as “percent” and “percentage” to identify relevant mentions of probability. We also extracted sentences with numerals between 0 and 1 that could match the 0-1 probability scale. We found 70 cases, all of which represented quantities and not risk (e.g., “cutting your salt intake to less than 6g (0.2oz) a day”), so these were not analysed further.

Data preparation and coding

Sentences that had two ratios or a range of percentages were counted twice, one for each value. For example, ratios that provided a range (e.g., 1 in 20-30 people) were broken down into two cases: one for each bound of the ratios (i.e., 1 in 20 and 1 in 30). Similarly, a probability of 20%-30% was broken down as two probabilities: 20% and 30%.

We manually reviewed all of the cases and filtered out any sentences that were not suitable for our analyses (180 of the 2,649 sentences). First, we excluded sentences that did not feature a quantity (e.g., “a small percentage of stillbirths are caused by problems with the mother’s health”). We then excluded sentences that featured a quantity that was not a frequency or a probability. For example: “they work by stopping the small blood vessels […], reducing blood loss by about 50%” was coded as a quantity (of blood) whereas “around 90% of breast lumps are benign” was coded as a frequency – 90% of the time, the breast lumps are benign. We also excluded vague values (e.g., “estimates of the rates of serious complications range widely, from one in tens of thousands to one in millions of treatments”). Overall, our analyses focused on 2,469 sentences that featured a percentage or a ratio phrase describing a frequency or a probability.

Our coding provided details about the mode of presentation of the format (i.e., numerals or words), the probability conveyed and the target outcome (coded on a subsample of 300 sentences). We coded as words all of the values that exclusively featured words (e.g., one in two) or those that featured both words and numbers (e.g., 1 in two or one in 2). For each value, we derived the probability conveyed by identifying the numerator and denominator. Some probabilities were above 100% because they reflected relative risk. We randomly selected 100 cases for each of the three formats ("1-in-X", "N-in-X*N" and percentages). We coded the outcomes of those 300 cases in these categories: i/ risk of the consequences of a medical condition; ii/ positive consequence of a treatment (e.g., remission); iii/ negative consequence of a treatment (e.g., negative side effect); iv/ risk of contracting an illness.