



# HHS Public Access

Author manuscript

*Med Decis Making*. Author manuscript; available in PMC 2016 November 01.

Published in final edited form as:

*Med Decis Making*. 2015 November ; 35(8): 937–947. doi:10.1177/0272989X15584922.

## Presenting numeric information with percentages and descriptive risk labels: A randomized trial

Aleksandr Sinayev, M.A., Ellen Peters, Ph.D., Martin Tusler, M.S., and Liana Fraenkel, M.D., M.P.H

The Ohio State University (M.A. Sinayev, Ph.D. Peters and M.S. Tusler), Yale School of Medicine (M.D., M.P.H Fraenkel)

### Abstract

**Background**—Previous research demonstrated that providing (vs. not providing) numeric information about medications' adverse effects (AEs) increased comprehension and willingness to use medication, but left open the question about which numeric format is best.

**Objective**—To determine which of four tested formats (percentage, frequency, percentage+risk label, frequency+risk label) maximizes comprehension and willingness to use medication across age and numeracy levels.

**Design**—In a cross-sectional internet survey ( $N=368$ ; American Life Panel, 5/15/08–6/18/08), respondents were presented with a hypothetical prescription medication for high cholesterol. AE likelihoods were described using one of four tested formats. Main outcome measures were risk comprehension (ability to identify AE likelihood from a table) and willingness to use the medication (7-point scale; not likely=0, very likely=6).

**Results**—The percentage+risk label format resulted in the highest comprehension and willingness to use the medication compared to the other three formats (mean comprehension in percentage + risk label format=95% vs mean across the other three formats = 81%; mean willingness= 3.3 vs 2.95, respectively). Comprehension differences between percentage and frequency formats were smaller among the less numerate. Willingness to use medication depended less on age and numeracy when labels were used.

**Limitations**—Generalizability is limited by use of a sample that was older, more educated, and better off financially than national averages.

**Conclusions**—Providing numeric AE-likelihood information in a percentage format with risk labels is likely to increase risk comprehension and willingness to use a medication compared to other numeric formats.

---

Correspondence: Ellen Peters, Ph.D., Department of Psychology, The Ohio State University, Columbus, OH 43210 (peters.498@osu.edu, 614-688-3477, FAX 614-688-3984).

Disclosures: The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health or the National Science Foundation. Data collection was supported by the National Institute on Aging, Grant Numbers R01AG20717 and P30 AG024962. The funding agreements ensured the authors' independence in designing the study, interpreting the data, writing, and publishing the report.

## Keywords

risk comprehension; risk communication; pharmaceutical decision making; adherence; numeracy; aging; informed decision making; statins

---

## INTRODUCTION

Patients often do not understand and are fearful of adverse events (AEs) from prescription drugs (1). Quantitative information about the likelihoods of AEs is rarely provided in the United States (2) although prescription drugs are probabilistic in terms of their AEs and benefits. One reason used to justify not providing numbers is low numeracy (the ability to understand and use mathematical and probabilistic concepts) in the general population (3). In a recent study, however, we demonstrated that providing numeric information (compared to verbal descriptions) decreased overestimation of AE likelihoods and increased reported willingness to use a hypothetical drug across numeracy levels (4). More and less numerate individuals overestimated risks substantially less and reported being more willing to use the medication when provided numeric information compared to when they were not provided numbers (although the effects were smaller among the less numerate). Similar results have been demonstrated in non-medical domains (5,6). A further concern about providing numbers is the lack of certainty about the AE likelihood estimates. A possible solution is to provide information about the uncertainty (e.g., provide a range estimate instead of or in addition to a point estimate). However, people appear to discount such range estimates as inexact and take little away from them (7).

Based on our prior results, we recommended providing numeric risk information, when it is available, to decrease overestimations of risk and increase willingness to use prescribed drugs for most patients. We did not, however, provide recommendations of which numeric format(s), if any, might be best. Numeric information about AE likelihood can be presented in a number of different numeric formats, including frequencies (e.g., out of 100 people, 14 will get dry mouth) and percentages (e.g., 14% of people will get dry mouth). It can also be presented with vs. without verbal risk labels (called “risk labels” from here on). For example, the European Commission Guidelines recommended prescription-drug risk labels ranging from “very common” (for AE likelihoods >10%) to “very rare” (<0.001%). In the present paper, we aim to extend previous results by testing the effects of different numeric formats and the presence vs. absence of risk labels on comprehension of AE likelihood and reported willingness to use a prescribed medication.

Recent reviews have summarized evidence on the effects of numeric formats and risk labels on comprehension and behavioral intentions (8,9). The reviews concluded that while no ideal numeric format exists, some are better than others. In particular, in most studies, the use of percentage formats increased comprehension (10,11,12,13,14) and decreased perceived risk (15,16,17) compared to frequency formats (but see 11,18 for different results).

Risk labels may be beneficial because they enhance the evaluative context (i.e., they make information easier to evaluate as good or bad; 19) and facilitate comparisons (20). Both

reviews recommended the use of risk labels because they enhanced comprehension (19,20); however, this recommendation may be misguided in some cases (20,21). Knapp et al. (22), for example, found that adding risk labels to frequency information about medication AEs significantly reduced risk perceptions and increased comprehension for one AE, but had no significant effects for the others. Their study included only 72 participants who saw exact numeric information, however, so that the non-significant findings may be due to the small sample size. Motivated by Knapp et al.'s and other findings (19,20) and the general idea that labels may enhance the evaluative context (19), we expected labels to result in higher comprehension. Overall, we expected percentage formats augmented with risk labels to facilitate comprehension and willingness to take medications compared to other formats, although previous evidence is somewhat mixed.

Visschers and colleagues proposed a dual-process explanation of format effects. In this framework, information is processed either heuristically, relying on incidental information such as mood, numeric format or label, or it is processed analytically, ignoring such incidental information and concentrating on the numeric estimates of the risk. Which process is used, according to this review, depends on relevant situational variables including cognitive capacities, motivation and available time. As evidence for this approach, they noted that several of the format effects reviewed bear some resemblance to known heuristics such as anchoring and availability. They did not, however, provide any evidence that format effects are moderated by their key situational variables.

In the present paper, we consider numeracy as one possible individual difference that increases deliberation with numbers and, thus, may decrease the influence of format effects. Consistent with this possibility, studies outside the scope of the Visschers et al. review have found that numeric formats and risk labels influence risk perceptions less for the more numerate than they do for the less numerate (16,23). Thus, we further test Visschers et al.'s dual-process conceptualization of format effects and extend previous results to the clinically relevant dependent variable of willingness to take a medication. Since older adults experience declines in deliberative efficiency (24), they, like the less numerate, also should experience larger effects of numeric format and risk label.

**Hypothesis 1** The use of percentage formats instead of frequency formats will increase comprehension of AEs and willingness to take a prescribed medication. Numeric format effects will be strongest among older and less numerate individuals.

**Hypothesis 2** The use of risk labels will increase comprehension of drug side effects. The effects of labels will be strongest among older and less numerate individuals.

In addition to testing Visschers et al.'s theory of format effects, we attempted to create a fuller picture of numeric format and label effects by following methodological suggestions. Visschers et al. recommended that more studies be done on non-student samples. West et al suggested that future studies should avoid problematic order effects and explore behavioral intentions rather than relying only on comprehension and risk perception. First, we examined a diverse sample, including older adults. Because prescription drug use increases

with age (25), studying this population is important. Second, we asked for behavioral intentions in the form of willingness to use a medication. Third, we used a between-subject design to avoid order effects. Fourth, though this was not explicitly recommended, we assessed the effects of numeric format in the presence of multiple pieces of numeric information rather than a single AE or benefit; this has rarely been done in previous studies on format effects (10,11,12). This is important because if numeric information is to be presented to patients, it will likely include a list of multiple AEs. Previous research shows that presenting risks simultaneously may reduce comprehension, compared to presenting them one at a time (26). Hence, results obtained in experiments in which one AE is presented at a time may not be as relevant to practicing physicians.

## METHODS

From 5/15/08–6/18/08, we conducted a randomized experiment over the Internet in the American Life Panel (ALP; [www.rand.org/labor/roybalfd/american\\_life.html](http://www.rand.org/labor/roybalfd/american_life.html)). ALP respondents are paid \$20 for each half-hour interview.

Respondents reported whether they regularly took a prescription medication to lower cholesterol. Because individuals making a new prescription-drug decision were more clinically relevant to our scenario, our analyses concentrated on the 61% of respondents who reported not regularly taking prescription medication to lower cholesterol ( $N=370$ )<sup>5</sup>. Respondents were told “Imagine you have been diagnosed with high cholesterol, a major cause of heart disease and stroke. Your doctor has prescribed you a new medication to lower your cholesterol, but it has possible adverse effects. Below is a list of possible adverse effects that may occur while taking this medicine.” They then were given risk information about medication AEs in one of six formats, randomly assigned. AE information was hypothetical but modified from statin information available from a pharmacy and the Physicians’ Desk Reference. The scenario was designed to include one serious rare AE and other less serious common AEs to provide respondents with a relatively realistic scenario. Four formats included numeric information and were retained for analysis in the present paper. The other two formats were non-numeric. Results comparing numeric and non-numeric conditions are available in our previous paper (4). Descriptions of AEs were placed in a table in descending order of frequency with accompanying information about AE likelihood.

Respondents were randomly assigned to condition in a 2 (numeric format: frequencies vs. percentages)  $\times$  2 (risk labels: present vs. absent) between-subjects design. Thus, there were four formats: percentage + risk label consistent with the European Commission guidelines (e.g., 14% of people=very common) (27), percentage without risk label, frequency + risk label (e.g., 14,000 out of 100,000 people=very common) and frequency without risk label. Consistent with the guidelines, frequency bands (e.g., 1 – 10% of patients) are often used with these labels instead of precise numbers, though we do not test these in the present

---

<sup>5</sup>Results among respondents who already use a similar medication are available from the first author. Given the wording of the scenario, it is not entirely clear how these respondents read the scenario. For example, they may have understood the scenario to mean that they would be taking an additional drug for their high cholesterol or that they would be taking a replacement drug for their current medication.

experiment. The percentage + risk label condition is shown in Figure 1, and the other formats are described in Table 1.

To assess willingness to use the drug, respondents answered the question “How likely is it that you will take this drug?” on a 7-point scale (0=not likely; 6=very likely). To assess risk comprehension, respondents were asked “About how many people do you think will get an upset stomach out of 100,000 people who take this medication?” in the frequency condition and “What percent (%) of people who take this medication will get an upset stomach?” in the percent condition (correct response= 400 or 0.4%, depending on condition). Percentage responses were transformed into frequencies out of 100,000 for analysis purposes. Finally, numeracy was measured using a new 8-item scale reduced from 18 items obtained from existing numeracy scales. The new scale was developed across two large, independent samples that varied widely in age and educational level. It had excellent psychometric properties based on a Rasch analysis and good predictive abilities relative to existing scales, supporting its predictive validity (28). The primary usefulness of the new scale is that it can be used in a wide range of populations, allowing for a clearer understanding of how numeracy influences the decision process across the lifespan.

## Analyses

We used R v3.0.2 (29), software for statistical computing, to conduct the analysis. Numeracy was mean-centered; we used a mean split to illustrate interactions, but the continuous measure was used in analyses. Age was treated as semi-continuous, with the young adults (age 18–39) coded as  $-1$ , middle-aged adults (40–59) coded as  $0$ , and older adults (60–89) coded as  $+1$  to match the results of our previous manuscript (4). The risk label variable was coded as  $-1/2$  when absent and  $+1/2$  when present. Numeric format was coded as  $-1/2$  for percent and  $+1/2$  for frequency. The coding of variables affects the size and significance of coefficients for main effects and lower order interactions when a higher order interaction is present (30). This coding was chosen so that our coefficients reflect effects on average across the experimental conditions for a middle-aged person of mean numeracy.

All analyses began with a logistic model (for comprehension) or a linear model (for willingness to use the drug) with the four-way interaction between numeracy, age, numeric format, risk-label presence, all lower order interactions with these terms, and all main effects as predictors. We examined the highest order non-significant interaction(s) currently in the model and removed the one that provided the least independent explanatory power (the term with the highest p-value), and re-estimated the model without that term. The goal was to progressively reduce the complexity of the model until only statistically significant terms remained. If all interactions of the highest order were significant ( $p < .05$ , two-tailed), this was the final model. We did not keep non-significant interactions, except for those that were of a lower order than a significant interaction. Hence, the absence of an interaction between variables in the final model implies it was non-significant and was removed from the model.

The study was well powered with a 97% probability of detecting a small effect size of  $d=0.2$  given the sample size.

The funding agreements ensured the authors' independence in designing the study, interpreting the data, writing, and publishing the report. Data collection was conducted and approved by RAND Corporation.

## RESULTS

The overall response rate for the survey was 78.9%. Non-respondents did not differ from respondents in education or gender but were slightly younger (49.6 years vs. 51.3 years,  $p=.04$ ). The final sample used in the present paper comprised 370 respondents who answered all questions ( $n=98, 99, 81, \text{ and } 92$ , respectively, in percentage, frequency, percentage+risk label, and frequency+risk label conditions) and who reported not currently using a statin medication. Two respondents gave birth dates outside of the feasible range so their data were not used in the analyses (as many of our analyses concerned age), leaving 368 respondents for analysis. Of the respondents in the conditions analyzed in the present paper, 69% were female (mean/median age=51/52 years; median income=\$60,000 – \$74,999; median education=Bachelor's degree; 88.9% Caucasian, 7.3% Black, 2.4% Asian, 0.5% Native American, and 0.8% reported Other). Compared to the 2010 U.S. Census, the sample is older, better educated, and higher in income (see Table 2). Age was not correlated with numeracy ( $r=0.04, p>0.05$ ) (28).

To begin, we conducted regressions predicting each of the demographic variables with condition. Numeracy and education differed by condition (see Table 3). Specifically, respondents who were assigned to either of the risk-label conditions were better educated and more numerate than the respondents assigned to a no-label condition. Numeracy was already included in all of our models, so any confounding effects were accounted for. Participants in either percent condition were more likely to be white. We reran all reported regression models controlling for education and ethnicity. These models are not reported for simplicity, but they returned substantially similar results which are available from the first author. Cronbach's alpha for the numeracy scale was 0.71.

### Risk Comprehension

Among our respondents, 83.7% answered correctly, 6.8% underestimated the risk, and 9.5% overestimated it, indicating that the question was quite easy. The more numerate comprehended more than the less numerate (92% and 78%, respectively, answered correctly,  $b=0.39, p<0.001$ ; see Table 4 for final model). This result was consistent with previous research (31,32,33). As expected, percentages resulted in higher comprehension than frequencies ( $b=-1.18, p=0.003$ ). Numeric format interacted with numeracy ( $b = -0.55, p=0.005$ ), but not in the direction predicted in Hypothesis 1. Specifically, the more numerate answered correctly more often with percentages (99%) than with frequencies (85%), whereas the difference was smaller among the less numerate (81% and 76% answered correctly with percentages and frequencies, respectively; see Figure 2). Consistent with previous research, the presence of risk labels (compared to their absence) increased comprehension as a main effect, with 78% of respondents answering correctly without risk labels and 90% answering correctly with risk labels ( $b=1.00, p=0.002$ ). No other interactions, including the hypothesized interactions with age, were significant.

Because percentages (compared to frequencies) and risk-label presence (compared to absence) improved comprehension, the percentage+risk label format maximized comprehension (95% answered correctly when shown percentages+labels, compared to 86% for frequencies+labels, 82% for percentages only, and 73% for frequencies only).

### Willingness to use the drug

Multiple regression analyses of willingness to use the drug were conducted following the same statistical approach described above (see Table 5 for the final model). More numerate respondents reported being more willing to use the drug than the less numerate ( $M=3.64$  and  $2.79$  for the more and less numerate, respectively;  $b=0.15$ ,  $p=0.007$ ). This result was consistent with previous research indicating that the more numerate perceive less risk in a variety of domains (34,35) and are more likely to adhere to medication plans (36). This main effect was qualified by an interaction with risk labels. Although adding risk labels had no main effect ( $b=0.20$ ;  $p=0.32$ ), their presence did increase willingness among the less numerate somewhat, but at (what might be considered) the cost of decreasing willingness among the more numerate (willingness means for the highly numerate were  $3.64$  and  $2.96$  without and with the risk labels, respectively; the same means for the less numerate were  $2.79$  and  $2.93$ ; interaction  $b=-0.22$ ;  $p=0.045$ ).

Effects of age also emerged. Older respondents reported being less likely to use the medication than middle-aged or young respondents ( $M=2.51$ ,  $3.00$ , and  $3.81$ , respectively;  $b=-0.62$ ,  $p<0.001$ ). Age also interacted with risk label ( $b=0.58$ ,  $p=0.047$ ). With labels, differences between older, middle-aged and younger respondents were smaller ( $M=2.73$ ,  $2.89$ , and  $3.52$ , respectively) than without labels ( $2.29$ ,  $3.10$ , and  $4.00$ , respectively). The presence of labels increased willingness to use the drug among older respondents but reduced it among younger respondents.

Numeric format had no main effect ( $M=3.19$  and  $2.89$  for percentages and frequencies;  $b=-0.27$ ,  $p=0.18$ ). However, an interaction with risk labels emerged ( $b=0.83$ ,  $p=0.04$ ). When labels were absent, there was indeed no effect of numeric format ( $M=3.09$  and  $3.12$ , respectively, for percentages and frequencies). When labels were present, however, respondents provided percentages were more willing to use the medication than those with frequencies ( $M=3.32$  and  $2.62$ , respectively; interaction  $t=2.07$ ,  $p=0.04$ ; See Figure 3). Numeric format did not interact with age or numeracy as hypothesized.

As a result, the percentages+risk labels format ( $M=3.32$ ) maximized willingness to use the medication compared to frequencies+labels, percentages, and frequencies ( $M=2.62$ ,  $3.09$ , and  $3.12$ , respectively).

## DISCUSSION

The aim of the present study was to extend previous results on numeric format and risk label effects and, if possible, identify the best format for presenting numeric likelihood information about AEs. We tested the effects of numeric formats (percentage vs. frequency), risk labels (presence vs. absence) and individual differences in numeracy and age on comprehension and willingness to use medication. Although the use of a medical treatment

is ultimately a choice that each respondent should make based on his or her preferences, participants were told that their physician prescribed the medication for them; hence, increasing willingness to use such a drug in similar cases may be advantageous at the population level. We recommend presenting a medication's AE likelihoods using percentages with risk labels over the other three formats due to increased comprehension and willingness to use the medication.

Percentages increased comprehension over frequencies, supporting Hypothesis 1 in part. This main effect of numeric format is consistent with people having more experience with percentages (13). Percentages also may be easier to process, since they have an implicit and constant denominator (of 100). Processing frequencies requires understanding the denominator, which varies across settings and is sometimes ignored (37).

Consistent with the main effect proposed in Hypothesis 2, the presence of risk labels also increased comprehension of numeric information. In the present study, what the labels described was consistent with the comprehension question asked; our results would not extend necessarily to other types of comprehension questions. When risk labels describe a facet of the quantitative information inconsistent with the comprehension question, they can actually decrease comprehension (20).

We found little evidence, however, for Visschers et al.'s dual-process proposal of format effects (8). Hypothesis 1's interaction was not supported, since neither younger nor more numerate participants demonstrated smaller numeric format effects for comprehension or willingness to use the medication. Instead, percentages, compared to frequencies, improved comprehension regardless of age, and the effect was larger among the more numerate. Percentages increased willingness to use the medication regardless of age or numeracy (though only when labels were present). The interaction in Hypothesis 2 also was not supported; labels improved comprehension regardless of age or numeracy.

The lack of numeric format differences in the less numerate may have been due to the very small risk that required us to represent it as a decimal percent (0.4%) rather than the integer percents that have been tested in prior research (e.g., 7%). Because integer percents are easier to recall and approximate than decimal percents (38) and the less numerate are more susceptible to a variety of format effects (16), our use of a decimal percent could explain the lack of a percentage vs. frequency numeric format difference in the less numerate. In particular, any advantages of seeing the information as percentages may have been equal to the disadvantages of seeing it as a decimal percent for these respondents. If this is the case, the less numerate may still benefit from percentages when AE likelihoods are larger and can be expressed in integer form.

Percentages also did not, in general, increase willingness to take the medication inconsistent with part of Hypothesis 1. The lack of numeric format effects on willingness to use the medication could be explained by previous research on evaluability. A single number is more difficult to evaluate by itself than it is in the context of other numbers (39). Thus, in previous studies, participants may have found it difficult to appraise the riskiness of a single AE and may have simplified information processing by using the easier-to-evaluate numeric



format and its ensuing affect as a substitute to judge risk instead of using the magnitude of the AE likelihood itself. When several AEs are presented together, however, the AE likelihood may become easier to evaluate in the comparative context (39,40,41). Thus, presenting multiple AEs (along with their likelihoods) may have eased the evaluation of the magnitudes of AE likelihoods, and this evaluability effect may have mitigated most numeric differences when risk labels were absent. But with risk labels, respondents given the percentage format were more willing to use the medication than those given the frequency format. Perhaps the addition of different labels to different numeric values made the numeric values difficult to compare and therefore reduced evaluability, causing respondents to rely on the peripheral cue of numeric format. If this is the case, recommendations to use a specific numeric format may be overly simplistic and should only be made in the context of the presence or absence of a risk label. Future research should examine these questions.

Although not hypothesized, the presence of risk labels (compared to their absence) increased willingness to use the medication in older and less numerate respondents (though at the cost of decreasing willingness somewhat among the younger and more numerate respondents). As a result, differences in willingness to use the medication between numeracy and age levels were effectively erased. Similarly, in our previous study, verbal descriptions of risk elicited similar willingness among respondents of different ages and numeracy levels, whereas numeric descriptions of risk resulted in higher willingness among younger and more numerate respondents (4). In a recent study, however, when risk labels were added to frequency bands, greater risk was perceived from the medication's side effects, compared to when labels were absent (42; they did not test comprehension). This possible difference in results between our study and theirs may have been due to their study participants being younger and more educated on average (in which case, our results are similar), or it could have been due to the more frequent and caustic side effects of the medication in their study. It may be that risk labels increased comprehension of these dire side effects and led to greater perceived risk in their study. In our study, on the other hand, increased comprehension of the less likely and less serious side effects may have led to no change in perceived risk on average. In other words, the addition of risk labels may have caused appropriate reactions in both studies. Alternatively, it may be that when risk labels were present, participants ignored the more difficult to evaluate frequency ranges (7) in place of the easier to evaluate labels. If this alternative possibility is true, then participants would have effectively been comparing numbers (in the frequency band condition) to verbal labels only, and we would expect lower risk perceptions in the numeric condition (4).

One limitation of our study is that it was conducted using an online panel; our sample was more educated, and likely more numerate, than the general population. Any numeracy effects may have suffered from restriction of range as a result. The clinical relevance of the results is also unclear given the present design. In addition, we chose to study only those respondents not currently taking a similar medication, thus limiting the scope of our study. These people may differ from those taking a similar medication, for example, in their beliefs about the effectiveness of medication in general, feelings about medications, desires to avoid identifying themselves as a "sick person" (43), or willingness to protect themselves against future ill health. The latter may be an important concern in willingness to treat high

cholesterol, because it is initially asymptomatic. The applicability of the present scenario was unclear, however, for those already taking a similar medication.

We examined willingness to use a medication and comprehension at levels of probability that were realistic for the medication studied, but format effects at other levels of probability also need to be tested. Finally, the risk labels used were consistent with those recommended by the European Commission guidelines. The numbers these verbal labels represent, however, do not reflect how people interpret them (44). For example, people interpret “almost impossible” to mean a probability as high as 8%, whereas the verbal descriptor that the European Commission uses for a probability between 1–10% is “common.” Guidelines should be set to use more psychologically feasible labels. Perfect labels are impossible, however, given that individuals vary considerably both in how they interpret verbal labels numerically and what verbal labels they would use for various numeric values (45,46). More important would be to choose a fixed set of verbal descriptors and numeric values to be used consistently so that they can become familiar to patients and providers.

This study adds to the growing literature on the effects of numeric formats and labels. Hypotheses based on a dual-process approach were not supported (e.g., percentages were understood better than frequencies, but mostly among the more numerate), and some new results emerged (e.g., the effect of numeric format on behavioral intentions depended on the presence versus absence of risk labels). Overall, we recommend percentages with risk labels for presenting likelihood information about medication AEs to patients who do not already use a similar medication.

## Acknowledgments

Financial support for this study was provided by National Science Foundation SES-1047757 and 1155924 (Peters) and by the National Institute of Arthritis and Musculoskeletal and Skin Diseases, part of the National Institutes of Health, under Award Number AR060231-01 (Fraenkel).

## References

1. World Health Organization. [Accessed August 22, 2012] Adherence to long-term therapies: evidence for action. [http://www.who.int/chp/knowledge/publications/adherence\\_report/en/](http://www.who.int/chp/knowledge/publications/adherence_report/en/). Published 2003
2. Ostrove, N. Background and overview of the CMI, PPI, Medication Guide Programs; Presented at: Food and Drug Administration Risk Communication Advisory Committee Meeting; February 26, 2009; Washington, DC.
3. Schwartz PH. Decision aids, prevention, and the ethics of disclosure. *Hastings Cent Rep.* 2011; 41(2):30–9. [PubMed: 21495515]
4. 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):430–442. [PubMed: 24246563]
5. Budescu DV, Por HH, Broomell SB, Smithson M. The interpretation of IPCC probabilistic statements around the world. *Nat Clim Change.* 2014; 4(6):508–512.
6. Budescu DV, Broomell SB, Por HH. Improving communication of uncertainty in the reports of the Intergovernmental Panel on Climate Change. *Psychol Sci.* 2009; 20(3):299–308. [PubMed: 19207697]

7. Schapira MM, Nattinger AB, McHorney CA. Frequency or probability? A qualitative study of risk communication formats used in health care. *Med Decis Making*. 2001; 21(6):459–467. [PubMed: 11760103]
8. Visschers VH, Meertens RM, Passchier WW, De Vries NN. Probability information in risk communication: a review of the research literature. *Risk Anal*. 2009; 29(2):267–287. [PubMed: 19000070]
9. West LW, Squiers LB, McCormack L, Southwell BG, Brouwer ES, Ashok M. Communicating quantitative risks and benefits in promotional prescription drug labeling or print advertising. *Pharmacoepidemiol Drug Saf*. 2013; 22(5):447–458. [PubMed: 23440924]
10. Waters EA, Weinstein ND, Colditz GA, Emmons K, et al. Formats for improving risk communication in medical tradeoff decisions. *J Health Commun*. 2006; 11(2):167–182. [PubMed: 16537286]
11. Cuite CL, Weinstein ND, Emmons K, Colditz G. A test of numeric formats for communicating risk probabilities. *Med Decis Making*. 2008; 28(3):377–384. [PubMed: 18480036]
12. Woloshin S, Schwartz LM. Communicating data about the benefits and harms of treatment: a randomized trial. *Ann Intern Med*. 2011; 155(2):87–96. [PubMed: 21768582]
13. Callison C, Gibson R, Zillmann D. How to report quantitative information in news stories. *Newsp Res J*. 2009; 30(2):43–55.
14. Zillmann D, Callison C, Gibson R. Quantitative media literacy: individual differences in dealing with numbers in the news. *Media Psychol*. 2009; 12(4):394–416.
15. Slovic P, Monahan J, MacGregor DG. Violence risk assessment and risk communication: the effects of using actual cases, providing instruction, and employing probability versus frequency formats. *Law and Hum Behave*. 2000; 24(3):271–296.
16. Peters E, Västfjäll D, Slovic P, Mertz CK, Mazzocco K, Dickert S. Numeracy and decision making. *Psychol Sci*. 2006; 17(5):407–413. [PubMed: 16683928]
17. 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):184–187. [PubMed: 15827666]
18. Knapp P, Gardner PH, Carriga N, Raynor DK, Woolf E. Perceived risk of medicine side effects in users of a patient information website: a study of the use of verbal descriptors, percentages and natural frequencies. *Br J Health Psychol*. 2009; 14(3):579–594. [PubMed: 18992183]
19. Peters E, Dieckmann NF, Västfjäll D, Mertz CK, Slovic P, Hibbard JH. Bringing meaning to numbers: the impact of evaluative categories on decisions. *J Exp Psychol Appl*. 2009; 15(3):213–227. [PubMed: 19751072]
20. Dieckmann NF, Peters E, Gregory R, Tusler M. Making sense of uncertainty: advantages and disadvantages of providing an evaluative structure. *J Risk Res*. 2012; 15(7):717–735.
21. 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):709–717. [PubMed: 14551000]
22. 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):267–271. [PubMed: 19896321]
23. 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):432–436. [PubMed: 21191122]
24. Peters E, Hess TM, Västfjäll D, Auman C. Adult age differences in dual information processes: implications for the role of affective and deliberative processes in older adults' decision making. *Perspect Psychol Sci*. 2007; 2(1):1–23. [PubMed: 26151915]
25. Roe CM, McNamara AM, Motheral BR. Gender and age-related prescription drug use patterns. *Ann Pharmacother*. 2002; 36(1):30–39. [PubMed: 11816254]
26. Zikmund-Fisher BJ, Angott AM, Ubel PA. The benefits of discussing adjuvant therapies one at a time instead of all at once. *Breast Cancer Res Treat*. 2011; 129(1):79–87. [PubMed: 20945090]

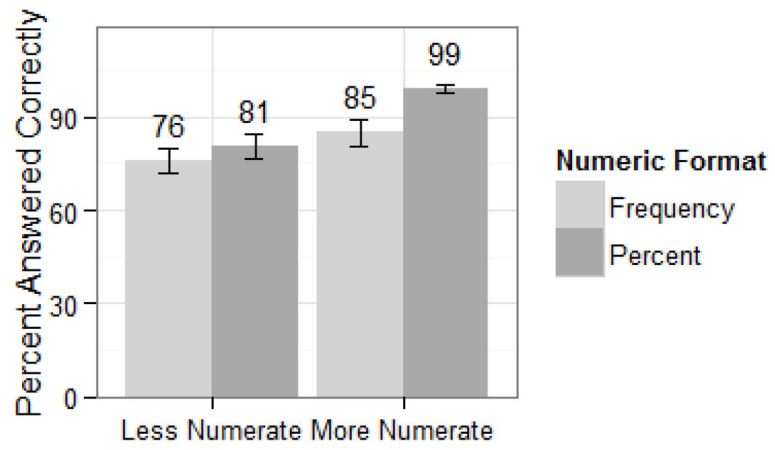
27. Berry D, Raynor T, Knapp P, Bersellini E. Over the counter medicines and the need for immediate action: a further evaluation of European Commission recommended wordings for communicating risk. *Patient Educ Couns*. 2004; 53(2):129–134. [PubMed: 15140451]
28. Weller J, Dieckmann NF, Tusler M, Mertz CK, Burns W, Peters E. Development and testing of an abbreviated numeracy scale: a Rasch Analysis approach. *J Behav Decis Mak*. Advance online publication. In press. 10.1002/bdm.1751
29. R [computer program]. Version 3.0.2. Auckland (New Zealand): R Core Team; 2013.
30. Irwin JR, McClelland GH. Misleading Heuristics and Moderated Multiple Regression Models. *J Marketing Res*. 2001; 38(1):100–109.
31. Schwartz LM, Woloshin S, Black WC, Welch HG. The role of numeracy in understanding the benefit of screening mammography. *Ann Intern Med*. 1997; 127(11):966–972. [PubMed: 9412301]
32. Peters E, Hibbard J, Slovic P, Dieckmann N. Numeracy skill and the communication, comprehension, and use of risk-benefit information. *Health Aff*. 2007; 26(3):741–748.
33. Woloshin S, Schwartz LM, Moncur M, Gabriel S, Tosteson AN. Assessing values for health: numeracy matters. *Med Decis Making*. 2001; 21(5):382–390. [PubMed: 11575488]
34. Woloshin S, Schwartz LM, Black WC, Welch HG. Women’s perceptions of breast cancer risk how you ask matters. *Med Decis Making*. 1999; 19(3):221–229. [PubMed: 10424829]
35. Weinstein ND, Atwood K, Puleo E, Fletcher R, Colditz G, Emmons KM. Colon cancer: risk perceptions and risk communication. *J Health Commun*. 2004; 9(1):53–65. [PubMed: 14761833]
36. Kalichman SC, Ramachandran B, Catz S. Adherence to combination antiretroviral therapies in HIV patients of low health literacy. *J Gen Intern Med*. 1999; 14(5):267–273. [PubMed: 10337035]
37. Denes-Raj V, Epstein S. Conflict between intuitive and rational processing: when people behave against their better judgment. *J Pers Soc Psychol*. 1994; 66(5):819–829. [PubMed: 8014830]
38. Peters E, Dieckmann NF, Dixon A, Hibbard JH, Mertz CK. Less is more in presenting quality information to consumers. *Med Care Res Rev*. 2007; 64(2):169–190. [PubMed: 17406019]
39. Hsee CK. The evaluability hypothesis: an explanation for preference reversals between joint and separate evaluations of alternatives. *Organ Behav Hum Decis Process*. 1996; 67(3):247–257.
40. Bazerman MH, Moore DA, Tenbrunsel AE, Wade-Benzoni KA, Blount S. Explaining how preferences change across joint versus separate evaluation. *J Econ Behav Organ*. 1999; 39(1):41–58.
41. Hsee CK, Loewenstein GF, Blount S, Bazerman MH. Preference reversals between joint and separate evaluations of options: a review and theoretical analysis. *Psychol Bull*. 1999; 125(5):576–590.
42. Knapp, P.; Gardner, PH.; Woolf, E. Health Expect [Internet]. 2015. Combined verbal and numerical expressions increase perceived risk of medicine side-effects: a randomized controlled trial of EMA recommendations. [cited 2015 Feb 2]. Epub 2015 Jan 26
43. Rosenbaum L. Beyond belief – how people feel about taking medications for heart disease. *N Engl J Med*. 2015; 372(2):183–187. [PubMed: 25564902]
44. Wallsten TS, Budescu DV. A review of human linguistic probability processing: general principles and empirical evidence. *Knowl Eng Rev*. 1995; 10(1):43–62.
45. Olson M, Budescu DV. Patterns of preference for numerical and verbal probabilities. *J Behav Decis Mak*. 1997; 10(2):117–132.
46. Karelitz TM, Budescu DV. You say probable and I say likely: improving interpersonal communication with verbal probability phrases. *J Exp Psychol Appl*. 2004; 10(1):25–41. [PubMed: 15053700]

You have been diagnosed with high cholesterol, a major cause of heart disease and stroke. Your doctor has prescribed you a new medication to lower your cholesterol, but it has possible adverse events. Below is a list of possible adverse events that may occur while taking this medicine, grouped by how common or rare each one is. The number of people out of 100,000 people who take this medication and have a given adverse event are listed in the first column of the table below. The second column is a verbal indication of how common or rare the adverse event is.

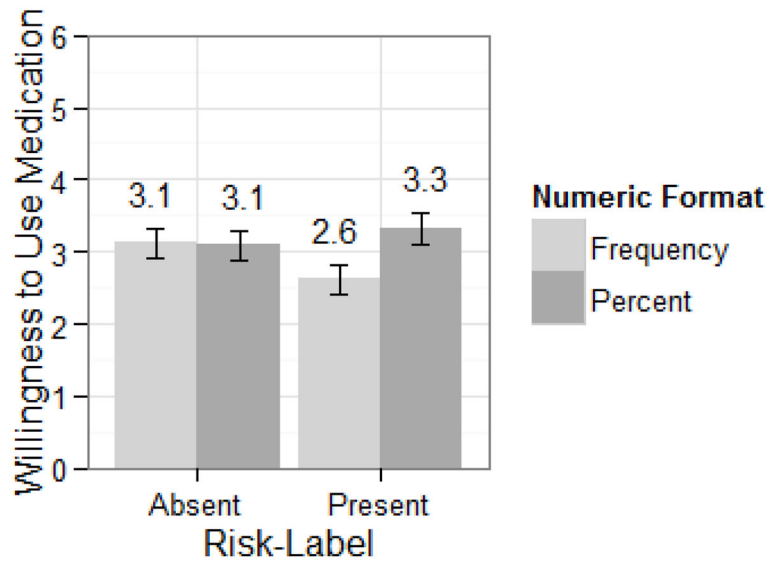
Possible Adverse events:  
out of 100,000 people

14%	Very common	dry mouth
7%	Common	constipation, headache
0.4%	Uncommon	stomach upset
0.05%	Rare	Dizziness
0.001%	Very rare	muscle damage leading to a very serious condition called rhabdomyolysis

**Figure 1.**  
Percent-plus-Risk-label example



**Figure 2.** Comprehension by numeric format and a mean split of numeracy. Error bars represent  $\pm 1$  standard error.



**Figure 3.** Mean willingness to use the medication by risk-label presence or absence and numeric format. Error bars represent +/- 1 standard error.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

**Table 1**

Description of four numeric likelihood formats

<b>Numeric Format</b>	<b>Format description</b>
Percent only	The table included percentages in place of frequencies (e.g., 14%, 0.05%). No risk labels were provided.
Frequency only	The table included frequencies (e.g., 14,000 or 50 out of 100,000 people). No risk labels were provided.
Percent+Risk label	This is Figure 1. Percentages and risk labels were provided.
Frequency+Risk label	The table included risk labels and frequencies instead of percentages.

Note. In each case, a table similar to that in Figure 1 was provided. Differences from Figure 1 are noted under the format description



**Table 2**

Demographics of our sample compared to the 2010 Census

Characteristic	Sample	2010 Census
Age		
18–34	20.1%	30.7%
35–64	51.6%	52.7%
65–89	28.2%	16.6%
Education		
Less than high school	1.9%	14.6%
High school diploma	12.5%	28.5%
Some college/vocational school	35.6%	31.1%
College graduate or more	50.1%	25.8%
Household Income		
Less than \$20,000	8.8%	18.9%
\$20,000–\$39,999	19.2%	21.5%
\$40,000–\$59,999	20.5%	17.4%
\$60,000 or more	51.5%	42.1%

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

**Table 3**

## Demographics by condition

	Percent Only	Frequency Only	Percent + risk label	Frequency + risk label
N	97	98	81	92
Mean age in years (s)	50 (14)	50 (16)	51 (13)	54 (13)
Education, % associate degree or less * †	48%	60%	54%	37%
Mean household income (s) ‡	11.5 (2.8)	11.3 (3.2)	11.3 (3.3)	11.8 (3.0)
Ethnicity, % white*	94%	89%	93%	80%
Mean numeracy (s) *	4.1 (1.9)	3.7 (1.8)	4.2 (1.9)	4.5 (1.7)

Note. An asterisk denotes that the variable differed significantly between label and no label conditions. The letter "s" denotes a standard deviation.

† Similar results are obtained with a more detailed education scale that approximately mirrors years of education (from 1=less than first grade to 16 = doctorate degree)

‡ Income categories range from 1 (less than \$5,000) to 14 (more than \$75,000).

**Table 4**

## Logistic Regression Results for Final Comprehension Model

Variable	Unstandardized b (SE)
Intercept	<b>2.10 (0.21)</b>
Numeracy (score)	<b>0.39 (0.10)</b>
Age (-1 for 18–39, 0 for 40–59, 1 for 59–89)	-0.26 (0.22)
Num. Format (0 for percent, 1 for frequency)	<b>-1.18 (0.39)</b>
Risk Label (0 when absent, 1 when present)	<b>1.00 (0.32)</b>
Numeracy × Num. Format	<b>-0.55 (0.20)</b>
Nagelkerke R <sup>2</sup>	0.17
X <sup>2</sup> (df=8)	39.4 (p<0.001)

Note: Significant predictors are in bold for emphasis. As is standard in logistic regression, the b-values indicated the change in log odds of getting the comprehension item correct for each one-unit increase in the predictor.

**Table 5**

Multiple Regression Results for Final Willingness to Use Medication Model

Variable	b (SE)
Intercept	<b>3.11 (0.10)</b>
Age (-1 for 18–39, 0 for 40–59, 1 for 59–89)	<b>-0.62 (0.15)</b>
Num. Format (-0.5 for percent, 0.5 for frequency)	-0.27 (0.20)
Risk Label (-0.5 when absent, 0.5 when present)	-0.20 (0.20)
Numeracy (mean centered score)	<b>0.15 (0.05)</b>
Age × Risk Label	<b>0.58 (0.29)</b>
Num. Format × Risk Label	<b>-0.83 (0.40)</b>
Risk Label × Numeracy	<b>-0.22 (0.11)</b>
Adjusted R <sup>2</sup>	0.083
F(7, 360)	5.8 (p<0.001)

Note: Significant predictors are in bold for emphasis.