Presentation of Quantitative Benefit Information

in

Direct-to-Consumer (DTC) Television and Print Advertisements for Prescription Drugs:

A Randomized Study

Final Technical Report

Office of Prescription Drug Promotion Research Team
Office of Medical Policy
Center for Drug Evaluation and Research
Food and Drug Administration

[2012]

Table of Contents

Index of Tables and Figures	3
Executive Summary	4
Introduction	9
Hypotheses	13
Preface	13
Statistical Format Effects in Print and Television Ads	14
Visual Format Effects in Print and Television Ads	16
Drug Efficacy Level Effects in Print and Television Ads	17
Method	18
Design	18
Measures	23
Procedure	25
Data Collection	27
Analyses	29
Results and Discussion	32
Participants	32
Statistical Format Results	34
Visual Format Results	46
Drug Efficacy Level Results	53
General Discussion	71
References	81
Appendix A: Example of Mock Print Ad	86
Appendix B: Questionnaire	87
Appendix C: Supplementary Tables	98

Index of Tables and Figures

Table	Title	Page
1	Study Conditions	18
2	Accuracy Measures	23
3	Statistical Formats and Their Wording	34
4	List of Statistical Format Hypotheses and Findings	34
5	Visual Format Conditions	46
6	List of Visual Format Hypotheses and Findings	46
7	List of Drug Efficacy Hypotheses and Findings	53
Figure		
1	Pie Chart	20
2	Bar Chart	21
3	Pictograph	21
4	Table	22
5	Study Sample Flowchart	33
6	Description of Statistical Format Effects on Benefit Accuracy	40
7	Description of Visual Format Effects on Benefit Accuracy	50

Executive Summary

FDA is committed to fostering the safe and effective use of prescription drugs and believes that improvement in peoples' understanding of risk and benefit information is essential to this commitment. This study evaluated the effect of including quantitative benefit information in various statistical and visual formats (e.g., relative or absolute frequency, bar graphs, tables) in direct-to-consumer (DTC) print and television advertisements (ads). FDA was interested in evaluating, for example, to what extent viewers of quantitative benefit information understood and could accurately recall such information and whether including such information changed their attitude toward the drug, their perception of how well the drug works, or how risky the drug is. FDA was also interested in whether including quantitative benefit information affected viewers' intentions to get more information about the drug or to take the drug. Finally, it was important to determine if including quantitative benefit information had a detrimental effect on the recall of risk information.

The study was guided by the following research questions:

(1) Does presenting quantitative benefit information in a statistical format in DTC ads help people recall quantitative benefit information in DTC ads? If so, which statistical formats are most helpful?

(2) Do visual aids help people recall quantitative benefit information in DTC ads? If so, which types of visuals are most helpful?

To answer these questions, FDA designed and implemented a randomized, controlled study exposing participants to a DTC prescription drug ad for a mock drug containing quantitative benefit information. Participants saw either a print DTC ad or a television DTC ad; note that the

television and print ad conditions were not designed for comparison with one another and some differences existed in the administration of these two conditions. The ad contained information about either a high-efficacy or a low-efficacy cholesterol drug. This benefit information about the drug was presented either in a statistical format or a visual format. The statistical formats tested were absolute frequency (for example, 65 out of 100), percent (for example, 65%), relative frequency (for example, 33 times more likely), a combination of absolute frequency and percent, and a combination of relative frequency and percent. The visual formats tested were pie charts, bar charts, tables, and pictographs. Participants in a control condition saw an ad without quantitative benefit information. Approximately 4,800 participants who had been diagnosed with high cholesterol responded to the study via the Internet.

Findings

Our results can be grouped into three categories: the effects of statistical format, the effects of visual format, and the effects of drug efficacy level.

Statistical format:

- Participants who did not see any quantitative benefit information about the drug were the least likely to accurately report how well the drug worked.
- Descriptively, presenting information using absolute frequency and percent formats appears to be best at helping participants accurately recall how well a drug works.
- There was a match between the kind of quantitative information participants viewed and the kind of quantitative information participants were able to accurately report. For instance, participants who viewed the benefit information as an absolute frequency (for example, 65 out of 100), compared with those who did not see any quantitative benefit

information, were better able to report how well the drug worked as an absolute frequency and a percent but not as a relative frequency (for example, 33 times better).

• In general, participants who saw the benefit information presented in two formats (for example, 65 out of 100 and 65%) were the most likely to accurately report how well the drug worked.

• The statistical format that participants saw did not affect their ability to recall the drug's risks, their attitude toward the drug, their perceptions of how well the drug works and how risky it is, or their intentions to get more information about the drug or to take the drug.

Visual format:

• When viewing print ads, participants who saw a bar chart or table, compared with those who saw no visual display, were more likely to accurately recall how well the drug worked. The bar chart was also better than the pictograph, and the table was better than the pie chart at helping participants accurately recall how well the drug worked.

• When viewing television ads, participants who saw any visual display, compared with those who saw no visual display, were more likely to accurately recall how well the drug worked. The bar chart was also better at helping participants accurately recall how well the drug worked than the pictograph and the table.

• The type of visual display that participants saw did not affect their ability to recall the drug's risks, their attitude toward the drug, their perceptions of how well the drug works and how risky it is, or their intentions to get more information about the drug or to take the drug.

Drug Efficacy level:

• Participants who saw quantitative information describing the high-efficacy drug had a more positive attitude toward the drug, thought the drug worked better, and reported more intentions to do things like get more information about the drug compared with participants who saw quantitative information describing the low-efficacy drug.

• Participants generally thought that the high-efficacy drug was less risky than the low-efficacy drug, despite identical risk profiles.

• The efficacy of the drug (high or low) did not affect participants' ability to recall the drug's risks.

Overall, our results showed that benefit recall was low, regardless of the particular presentation of information. This is likely an effect of our procedure, in which participants were not able to refer back to the print ad or television ad as they were answering the questions.

Conclusions

Our findings demonstrate that participants can accurately recall quantitative benefit information from DTC prescription drug print and television ads for a mock prescription drug, and that providing this information does not adversely influence their recall or perceptions of the product's risk. Overall, presenting information using absolute frequency and percent formats may be best at helping participants accurately recall how well a drug works. Presenting a visual aid also appears to help participants accurately recall how well a drug works, with bar charts and tables demonstrating advantages over other visual formats. In general, providing information to participants enables them to see the information and answer questions about it correctly, although it does not necessarily change: (1) their attitude toward the drug, (2) their perception of how well the drug works and how risky it is, or (3) their intentions to get more information about the drug or to take the drug. At the same time, including quantitative benefit information did not

have a detrimental effect on the recall of risk information. Thus, the inclusion of quantitative benefit information in DTC print and television ads has the potential to help people make informed decisions about speaking with their health care professional about prescription drugs.

A major contribution of this research is that, to our knowledge, it is the first study to systematically examine the addition of quantitative information in television DTC ads. In fact, to our knowledge, the risk communication literature has focused only on print (or online text) modalities, making this the first study to examine the addition of quantitative information in a dynamic, television modality.

FULL TECHNICAL REPORT

Presentation of Quantitative Benefit Information in Direct-to-Consumer (DTC) Television and Print Advertisements for Prescription Drugs

A Randomized Study

The Federal Food, Drug, and Cosmetic Act (FD&C Act) requires that manufacturers, packers, and distributors (sponsors) who advertise prescription human and animal drugs, including biological products for humans, disclose in advertisements (ads) certain information about the advertised product's uses and risks.[1] The presentation of this information is likely to evoke active trade-offs, in which people compare the perceived risks of taking or not taking a treatment and the perceived benefits of taking or not taking a treatment (Schwartz, Woloshin, et al. 1997). Because FDA has an interest in fostering safe and proper use of prescription drugs, it seeks ways to improve peoples' understanding of this risk and benefit information.

Under the FD&C Act, FDA engages in a variety of communication activities to ensure that patients and health care professionals have the information they need to make informed decisions about treatment options, including the use of prescription drugs. FDA regulations (21 CFR 201.57) describe the content of required product labeling, and FDA reviewers ensure that labeling contains accurate and complete information about the known risks and benefits of each drug.

FDA regulations require that prescription drug ads that make promotional claims about a product also include risk information in a "balanced" manner (21 CFR 202.1(e)(5)(ii)), both in terms of the content and presentation of the information. This balance applies to including information to the front, display page of an ad, as well as including information "in brief

[1] For prescription drugs and biologics, the FD&C Act requires advertisements to contain "information in brief summary relating to side effects, contraindications, and effectiveness" (21 U.S.C. 352(n)).

summary" about the advertised product's "side effects, contraindications, and effectiveness,"[2] usually, but not always, on a separate page.

FDA recently provided a draft guidance to sponsors about ways to present risk information in prescription drug ads.[3] Although the guidance cited much research on the presentation of information, research addressing specifically how to present benefit information in prescription drug ads is limited. For example, benefit claims appearing in DTC ads are often presented in general language that does not inform patients of the likelihood of that benefit and are often simply variants of the indication, as listed in the FDA-approved product labeling. One content analysis of DTC advertising by Woloshin, Schwartz, et al. (2001) found that information about product benefits and risks is often presented in an unbalanced fashion. The researchers classified the "promotional techniques" used in the ads. Emotional appeals were observed in 67 percent of the ads while vague and qualitative benefit terminology was found in 87 percent of the ads. Only 9 percent contained data. However, for risk information, half the ads used data to describe side-effects, typically with lists of side-effects that generally occurred infrequently. Similarly, a content analysis (Frosch, Krueger, et al. 2007) found that only a small proportion of product-claim ads gave specific information about the population prevalence of the medical condition being advertised. The authors criticize DTC for presenting "best-case scenarios that can distort and inflate consumers' expectations about what prescription drugs can accomplish" (Frosch et al. 2007, p. 12) without disclosing how many people are likely to experience that benefit.

---

[2] See section 502(n) of the FD&C Act.

[3] Draft guidance for industry: Presenting Risk Information in Prescription Drug and Medical Device Advertising. Available at http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM155480.pdf, last accessed September 26, 2012.

There is a substantial literature on the communication of *risks* in quantitative form (for reviews, see Fagerlin, Ubel, et al. 2007; Lipkus 2007). In general, this research has found that some statistical formats (such as N out of 100) are easier for people to understand than others (such as relative frequency; Fagerlin et al. 2007) and that some visuals (such as pictographs; Fagerlin, Wang, et al. 2005) can help people understand risk information. However, previous research shows that individuals have great difficulty processing numerical concepts (e.g., Beyth-Marom 1982; Bowman 2002; Cohen, Ferrell, et al. 2002). This literature also shows that quantitative information is not helpful for everyone. Numeracy is the degree to which a person can work with numbers, including basic probability and mathematical concepts (Lipkus, Samsa, et al. 2001)—people with low numeracy are less able to understand quantitative risk information.

Given this broader literature on risk communication and numeracy, it was not clear whether adding quantitative benefit information to DTC prescription drug ads would help people and if so, how best to do it (as little research exists on the topic). To our knowledge, only one series of studies has addressed whether providing quantitative information about benefits enables people to make better choices about potential therapy. This series of studies tested one possible format (termed the "drug facts box") for presenting quantitative benefit and risk information to participants (Schwartz, Woloshin, et al. 2007, 2009; Woloshin and Schwartz, 2011; Woloshin, Schwartz, et al. 2004). In these studies, the drug facts box format contained information about the benefits and risks in terms of percent and frequency formats (how many participants in the clinical trial experienced a benefit or side effect compared to placebo). As expected, these studies showed that people who were provided drug efficacy information used it to make decisions. Participants given a drug facts box were more likely to correctly choose the product

with the higher drug efficacy rate than participants who saw an ad that did not contain quantitative benefit information.

Although these results suggest that providing quantitative benefit information to people is better than not providing any quantitative benefit information at all, they do not demonstrate whether some statistical formats would be more useful for people than others, or whether visuals would help people understand this information. In addition, this research focused on the addition of quantitative information to the second page or brief summary of prescription drug print ads; it did not address the presence or absence of quantitative information on the first page. We are unaware of research looking into the integration of quantitative benefit information into the body of the ad itself.

It also remains unknown if ways of communicating product benefit work equally well across print and television DTC media. To our knowledge, research on presenting quantitative information in risk communication has been conducted exclusively with static modalities. The ideal format for presenting quantitative information may vary as a function of presentation. The amount of mental processing capacity each individual can devote to understanding a message varies depending on: (1) how long individuals have to look at the material and (2) whether the material is self-paced or presented at an uncontrollable speed. As a result, some forms of quantitative information may lend themselves to print, rather than broadcast. This particular understanding is crucial to the risk-benefit tradeoff that patients must make in consultation with a health care professional to achieve the best health outcomes.

The goal of the study was to evaluate the effect of including quantitative benefit information, in various statistical and visual formats, in both DTC print and television ads. The study was guided by the following research questions:

(1) Does presenting quantitative benefit information in a statistical format in DTC ads help people recall quantitative benefit information in DTC ads? If so, which statistical formats are most helpful?

(2) Do visual aids help people recall quantitative benefit information in DTC ads? If so, which types of visuals are most helpful?

Hypotheses

Preface

The proposed research had two main objectives: to examine different statistical formats and to examine different visual formats. First, we tested several statistical formats to determine whether the presentation of benefit information in different formats affects perception and understanding of that benefit. The risk communication literature suggests that presenting numerical risk information as an absolute frequency (e.g., N out of 100) may be the most easily understood format (Fagerlin et al. 2007). Percentages and combinations of absolute frequency and percentages, represent increasingly complex statistical formats; however, they may not differ from the baseline of absolute frequency for average people. In contrast, the risk communication literature also suggests that presenting numerical risk information as a relative frequency (e.g., 33 times more likely) is a markedly more complex statistical format that biases perceptions (Fagerlin et al. 2007). Thus, presenting benefit information as a relative frequency, compared to absolute frequency, may affect perceptions of that benefit. Presenting the combination of absolute frequency and relative frequency may mitigate this effect.

Second, we tested several visual formats to determine whether the presentation of a visual format, in conjunction with the presentation of absolute frequency information, affects perceptions of benefit. The risk communication literature suggests that the addition of visual

formats, such as bar charts, tables, and pictographs, increases the understanding of numerical information (Ancker et al. 2006; Lipkus and Hollands, 1999). However, not all visual formats are equally helpful; for instance, pie charts may only help when people are comparing proportions (Lipkus 2007). Thus, presenting benefit information with a *bar chart*, *table*, and *pictograph*—but not necessarily with a *pie chart*—may affect people's understanding of benefit information, in comparison to when there is no visual format.

Testing two levels of drug efficacy enabled us to ensure that participants cognitively processed the information about the drug benefits. It also enabled us to determine whether our statistical and visual format findings differed depending on the magnitude of the numbers involved.

In addition, measuring numeracy enabled us to assess the magnitude of these effects across participants. Similarly, the separate television and print portions of the study enabled an assessment of whether these effects were consistent across these modalities.

Statistical Format Effects in Print and Television Ads

(S1) We will test two explanations for behavioral intentions, attitude toward drug and perceived benefit.

(S1a) **Overestimation hypothesis**: The first hypothesis rests on the assumption that, in the absence of any quantitative information, participants will overestimate the effectiveness of drugs. Accordingly, we would predict that behavioral intentions, attitude toward drug, and perceived benefit will be **higher** for participants in the *no* statistical format condition, compared to *all other* statistical format conditions. Support for this interpretation will be found if estimates of the benefits are higher in the *no* statistical format condition than in *all other* statistical format conditions.

(S1b). **Peripheral cue hypothesis**: This hypothesis rests on the assumption that any statistical information will be used as a peripheral cue; that is, participants will not process the quantitative information provided in the various statistical formats, but will view it as scientific proof of the drug's benefit. Accordingly, we would predict that behavioral intentions, attitude toward drug, and perceived benefit will be **lower** for participants in the *no* statistical format condition, compared to *all other* statistical format conditions. Support for this interpretation will be found if, in addition to perceived benefit effects, estimates on attitude toward the ad peripheral cue measures—ratings of how believable, persuasive, informative, etc, the ad is—are lower in the *no* statistical format condition than in *all other* statistical format conditions.

(S2) Based on the risk communication literature, we predict that the absolute frequency, percentage, and absolute frequency and percentage conditions may not differ on behavioral intentions, attitude toward drug, or perceived benefit. However, we predict that behavioral intentions, attitude toward drug, and perceived benefit will be higher in the *relative frequency* condition than in the *absolute frequency*, *percent*, *absolute frequency + percent*, and *absolute frequency + relative frequency* conditions.

(S3) The effects tested in hypotheses 1-2 will be modified by numeracy. For instance, we expect that the difference between the *relative frequency* and the *absolute frequency + relative frequency* conditions will be greater for high-numeracy participants than for low-numeracy participants (because high-numeracy participants will be more likely to use the additional information provided by the absolute frequency).

(S4) Benefit accuracy will be lowest in the *no* statistical format condition and highest in the absolute frequency condition (Slovic, Monahan, et al. 2000). Tests of other relations between

statistical formats will be exploratory. For instance, we might see information overload with some formats (e.g., absolute frequency and relative frequency), which impedes benefit accuracy.

(S5) The effects tested in (S4) will be modified by numeracy, such that low-numeracy participants will show greater differences in benefit accuracy across statistical formats than will high-numeracy participants (Peters et al. 2006).

(S6) We expect that risk recall will *not* differ by statistical format, but we will conduct exploratory analyses to determine whether information overload impedes risk recall.

(S7) We expect that perceived risk will be lowest in the relative frequency condition if perceived benefit is indeed highest in this condition (Slovic and Peters, 2006).

Visual Format Effects in Print and Television Ads

(V1) We will examine two explanations for benefit accuracy, behavioral intentions, attitude toward drug, and perceived benefit.

(V1a) **Visual information facilitation hypothesis**: The first hypothesis rests on the assumption that participants will, to the extent possible, process and use the information in the visual formats. The risk communication literature suggests that visual representations of risk can increase understanding, and that participants will have a more difficult time processing this kind of information in pie charts, as compared to other visual formats. Therefore, our first hypothesis is that benefit accuracy will be higher in the *bar chart, table,* and *pictograph* conditions—but not necessarily the *pie chart* condition—than in the *no* visual format condition. Tests of other relations between visual formats will be exploratory.

(V1b) **Information overload hypothesis**: Alternatively, there may be no differences across visual formats on behavioral intentions, attitude toward drug, perceived benefit, or

benefit accuracy if the visual serves as a distraction or is too much information to process.

(V1c) **Peripheral cue hypothesis**: Behavioral intentions, attitude toward drug, and perceived benefit—but not benefit accuracy—may be higher in *all visual* conditions than in the *no* visual condition if the visual information serves as a peripheral cue.

(V2) The effects tested in (V1) will be modified by numeracy. For instance, we expect that high-numeracy participants will be more likely to process the information in the visual formats and, thus, more likely to show the pattern of effects outlined in 1a, compared to low-numeracy participants.

(V3) We expect that perceived risk and risk recall will *not* differ by visual format, but we will conduct exploratory analyses to determine whether information overload impedes risk recall.

Drug Efficacy Effects in Print and Television Ads

(E1) Behavioral intentions, attitude toward drug, and perceived benefit will be higher in *high*-efficacy conditions than in *low*-efficacy conditions.

(E2) We will explore whether there are differences between the *no* efficacy condition (control) and the *low*- and *high*-efficacy condition on behavioral intentions, attitude toward drug, and perceived benefit.

(E3) Benefit accuracy will be higher in the *low*- and *high*-efficacy conditions than in the *no* efficacy condition. There will be no difference between the *low*- and *high*-efficacy conditions.

(E4) The effects tested in (E1) and (E3) will be modified by numeracy, such that high-numeracy participants will be more likely to show these effects than will low-numeracy participants.

(E5) Risk recall will not differ by drug efficacy level (no, low, high).

(E6) Perceived risk will be lower in the high-efficacy condition compared with the low-efficacy condition because, according to the Affect Heuristic (Slovic and Peters, 2006), people perceive things that are more beneficial as less risky.

Method

Design

To answer these questions, we designed and implemented a randomized, controlled study, exposing participants to DTC prescription drug print and television ads containing quantitative benefit information. The study design was reviewed by the Office of Biostatistics within the Center for Drug Evaluation and Research at FDA. The ads were tailored into control and study arms, with each study arm delivering the quantitative information in a different statistical or visual format (Table 1). As shown in Table 1, for each ad type (print or television) there were 10 experimental arms plus a control condition in the statistical format design and 10 experimental arms plus a control condition in the visual format design.

Table 1. Study Conditions

Statistical Format	Drug Efficacy Level	Visual Format				
		No Visual	Pie Chart	Bar Chart	Table	Pictograph
Absolute Frequency	High	Arm 1	Arm 11	Arm 13	Arm 15	Arm 17
	Low	Arm 2	Arm 12	Arm 14	Arm 16	Arm 18
Percent	High	Arm 3				
	Low	Arm 4				
Absolute Frequency + Percent	High	Arm 5				
	Low	Arm 6				
Relative Frequency	High	Arm 7				
	Low	Arm 8				

Relative Frequency + Percent	High	Arm 9
	Low	Arm 10
Control	—	Arm 19

Four characteristics of the study's prescription drug ads were varied (manipulated): ad type, statistical format, visual format, and drug efficacy level.

Ad Type. Participants were assigned to view either a print or television ad for the mock prescription drug for high cholesterol, named Votrea.[4] The two-page print ad contained both a display page with the study manipulations and a brief summary page listing important information (brief summary) about the drug. The 90-second television ad was filmed with two professional actors portraying a married couple, similar to many DTC television ads. The actors and visual setting did not differ by ad type.

Statistical Format. Participants were randomly assigned to study arm, with approximately half of the study sample in the Statistical Format design. These participants viewed an ad delivering the quantitative benefit information in one of five statistical formats. These formats were applied to both the high-efficacy and low-efficacy conditions.

- **Absolute Frequency**—With Votrea, 65 out of 100 people lowered their bad cholesterol to normal levels versus 2 out of 100 people with no treatment.
- **Percent**—With Votrea, 65 percent of people lowered their bad cholesterol to normal levels versus 2 percent of people with no treatment.
- **Relative Frequency**—With Votrea, people were 33 times more likely to lower their bad cholesterol to normal levels compared to no treatment.

[4] Participants who accessed the survey through dial-up were assigned to the print conditions due to bandwidth limitations. See the **Study Arm Assignment** section for the description of study arm and condition assignment.

- **Absolute Frequency + Percent**—With Votrea, 65 out of 100 people (65 percent) lowered their bad cholesterol to normal levels versus 2 out of 100 people (2 percent) with no treatment.
- **Relative Frequency + Percent** —With Votrea, 65 percent of people lowered their bad cholesterol to normal levels versus 2 percent of people with no treatment—that's 33 times more effective.

Visual Format. Participants were randomly assigned to study arm, with approximately half of the participants in the Visual Format design. These participants viewed an ad that contained the absolute frequency benefit information plus one of five visual aid formats: pie chart, bar chart, pictograph, table, or no visual (Figures 1-4). These visual aids were applied to both the high-efficacy and low-efficacy conditions.

Figure 1. Pie Chart.

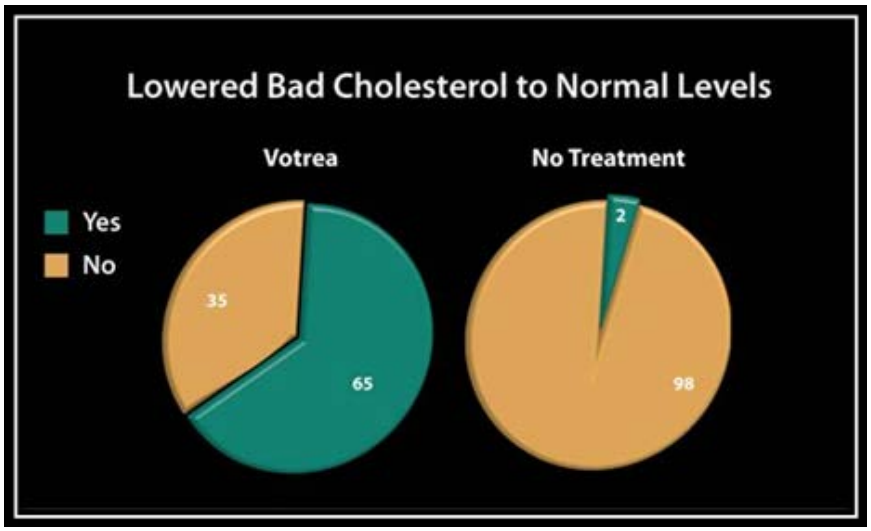


Figure 2. Bar Chart.

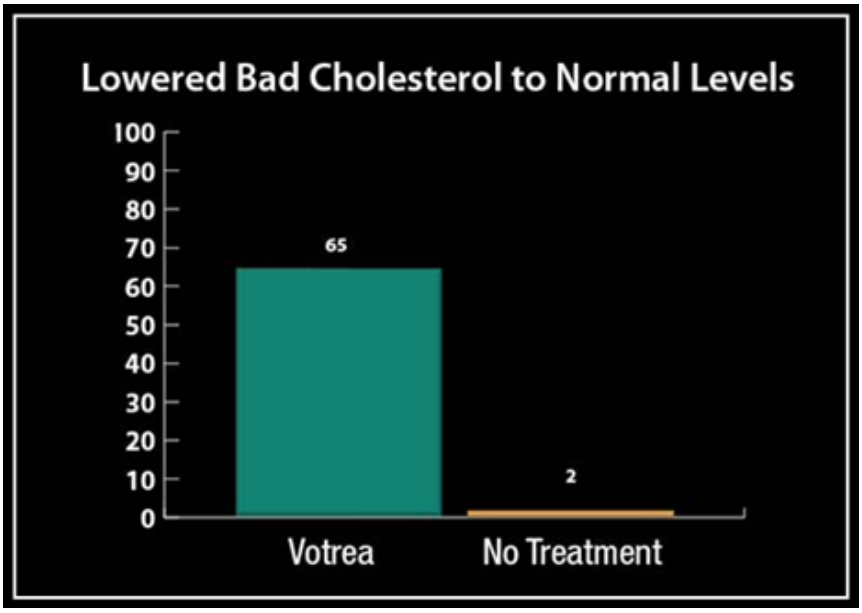


Figure 3. Pictograph.

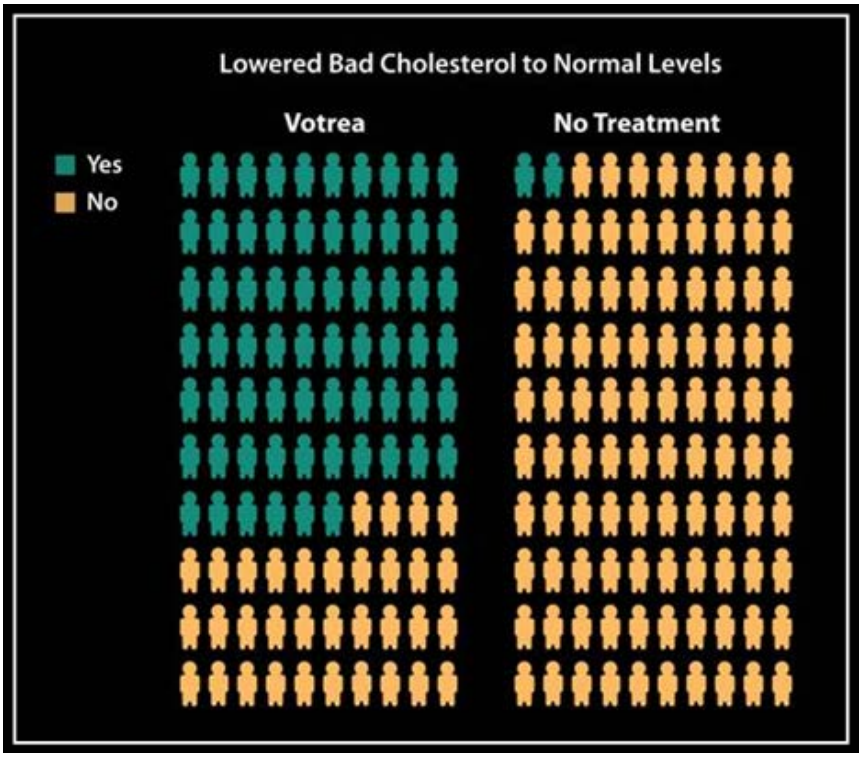


Figure 4. Table.

Lowered Bad Cholesterol to Normal Levels			
	Yes	No	Total
Votrea	65	35	100
No Treatment	2	98	100

Drug Efficacy Level. Because the drug's efficacy level could influence how people process and use the information, the ads were tailored to present one of three drug efficacy levels: high-efficacy, low-efficacy, and no-efficacy information (control). The particular drug efficacy rates were determined by examining existing cholesterol medications in the class and consulting with the consumer safety reviewers responsible for the metabolic and endocrine classes of prescription drugs in the Division of Drug Marketing, Advertising, and Communications (DDMAC) in FDA's Center for Drug Evaluation and Research.[5]

- **High Drug Efficacy**—65 out of 100 people who used Votrea lowered their bad cholesterol to normal levels (compared with 2 out of 100 who received no treatment).
- **Low Drug Efficacy**—10 out of 100 people who used Votrea lowered their bad cholesterol to normal levels (compared with 2 out of 100 who received no treatment).
- **Control**—No quantitative benefit information included in the ad (i.e., "Votrea reduces bad cholesterol for people with several common risk factors for heart disease").

[5] DDMAC is now the Office of Prescription Drug Promotion (OPDP). See http://www.fda.gov/AboutFDA/CentersOffices/OfficeofMedicalProductsandTobacco/CDER/ucm090142.htm, last accessed September 26, 2012.

Measures

Benefit Accuracy. We measured whether participants could accurately report how well the drug worked. To do so we used four separate questions (see Table 2). We scored responses as correct (see Table 2 for correct responses) or incorrect. For the control condition, participants could either report the low-efficacy or high-efficacy correct response to be considered correct.

Table 2. Accuracy Measures

Format	**Question**	**Correct Response**	
		Low Drug Efficacy	**High Drug Efficacy**
Absolute Frequency	If 100 people take Votrea, how many will lower their bad cholesterol to normal levels?	10	65
Percentage	What percentage (%) of people who take Votrea will lower their bad cholesterol to normal levels?	10	65
Relative Frequency	How many times more effective is Votrea than no treatment in lowering bad cholesterol?	5	32-33
Placebo	If 100 people receive *no treatment*, how many will lower their bad cholesterol to normal levels?	2	2

Perceived Benefit. The measure of perceived benefit was the average of two items: "Based on the information in the ad, how effective would Votrea be for you?" and "Based on the information in the ad, how well would Votrea work for you?" The correlation between these two items was high and ranged from .95 - .97 across the four designs. The measure used the following scale: 1 = *not at all effective/well*, 7 = *very effective/well*.

Perceived Risk. The measure of perceived risk was the average of two items: "Based on the information in the ad, how safe would Votrea be for you?" and "Based on the information in the ad, how risky would Votrea be for you?" The correlation between these two items was high

and ranged from 0.78 - 0.84 across the four designs. The measure used the following scale: 1 = *very safe/ not at all risky*, 7 = *not at all safe/ very risky*.

Behavioral Intention. Participants' behavioral intentions were measured by asking participants how likely they were to do three behaviors: "Talk to your doctor about Votrea," "Look for more information about Votrea," and "Take Votrea if prescribed." Participants responded using the following scale: 1 = *not at all likely*, 4 = *extremely likely*. These three items did not form a scale (Cronbach's alpha ranged from 0.73 to 0.77 across the four designs) and therefore were examined separately.

Risk-Benefit Tradeoff. Participants were asked to respond to the statement, "Thinking overall about the risk and benefits, would you say Votrea has:" using a 7-point scale (1 = *more risks than benefits*, 7 = *more benefits than risks*).

Attitude Toward the Ad. Participants were asked to respond to the statement, "How good or bad do you feel about this product?" using a 5-point scale (1 = *very bad*, 5 = *very good*).

Risk Recall. We had two measures of risk recall. The first was an open-ended, unaided, question: "What are the risks of Votrea?" Independent raters coded responses into (1) no risks recalled, (2) incorrect risks recalled only, or (3) one or more correct risks recalled (inter-rater reliability across questions was 0.96). The second measure was a closed-ended, aided measure. Participants saw eight statements about Votrea, such as "People with liver problems should not take Votrea," and were asked to report whether each statement was mentioned in the ad as a risk of taking Votrea (yes/no). Half of these statements were correct and half were incorrect. We summed the number of correct responses for the eight items to create the aided risk recall measure.

Numeracy. We had two measures of numeracy. The first measure consisted of three fill-in-the-blank items (for example, "In the Big Bucks Lottery, the chance of winning a $10 prize is 1%. What is your best guess about how many people would win a $10 prize if 1,000 people each buy a single ticket to the Big Bucks Lottery?"). We summed the number of correct responses for the three items to create an *objective* numeracy measure (Schwartz, Woloshin, et al. 1997). The second measure consisted of eight items measured on 6-point scales (for instance, "How good are you at calculating a 15% tip?"; 1 = *not at all good*, 6 = *extremely good*). We used the average of these six items to create the *subjective* numeracy measure (Fagerlin, Zikmund-Fisher, et al. 2007).

Demographic and Health Characteristics. The demographic variables used in this study were age, gender, race, ethnicity, and educational level. We also asked participants to report their total cholesterol level.

Additional Variables. We also included several other measures that are not discussed in this report because we did not have hypotheses about these variables or they did not show a relation with our variables of interest. These included questions about benefit recall, the severity of participants' high cholesterol, whether they were currently taking medication for high cholesterol, how much they felt they knew about high cholesterol and treatments for high cholesterol, how skeptical they were of the ad, and how clearly they thought the risks and benefits were presented in the ad.

Procedure

This study underwent IRB review and was granted an exemption by FDA's Research Involving Human Subjects Committee (RIHSC). The study sample included U.S. adults who self-reported that they had been medically diagnosed with high cholesterol (n = 4,805) and who

were recruited from an existing online consumer panel representative of the U.S. adult population. A priori power analyses were conducted to ensure that the sample size was sufficient to detect medium to small effect sizes with power = 0.90 and alpha = 0.05. A mock prescription drug for high cholesterol (Votrea) and campaign elements, such as a logo and a benefit-risk profile, were created. Print and television ads for the drug were created, and those ads were tailored into the study conditions. The quantitative benefit information was delivered as a written statement (print) or voiceover (television) within the ad (see Appendix A for a sample print ad). Participants were randomly assigned to a study arm, exposed to a single DTC ad, and asked to complete a short Web-based survey.

Sampling Frame and Recruitment. The study was conducted with U.S. adults aged 18 or older who were medically diagnosed with high cholesterol so that the mock drug (Votrea) and the DTC ads would be relevant to the participants. Knowledge Networks' (KN) online consumer panel of U.S. adults (KnowledgePanel [www.knowledgenetworks.com/knpanel/index.html]) was the study's sampling frame. KnowledgePanel is representative of the U.S. adult population and is the only probability-based consumer panel offering online data collection capabilities. To ensure full representation of U.S. adults, KN equips panelists from non-Internet households with computer hardware and dial-up Internet access. KN recruited KnowledgePanel members through two random sampling strategies: random digit dialing (RDD) and address-based sampling (ABS).

Study Sampling, Eligibility, and Recruitment. For this study, participants from KnowledgePanel were randomly selected if they met three eligibility criteria: (1) adult aged 18 or older, (2) diagnosed with high cholesterol by a health care professional, and (3) capable of

viewing and listening to DTC ads on a computer or survey device. To eliminate cross-contamination of study arms, only one person per household was sampled.

First, potential participants were identified by examining prescreening data and eliminating panelists that had not been diagnosed with high cholesterol. Second, a random sample of panelists who had been diagnosed with high cholesterol was selected. Third, an email invitation (via KN) was sent to these panelists that invited them to participate in the study. The email invitation contained a unique link, specifically for that panelist, to the study survey.

Reminder invitations and a reminder telephone call were used to try to convert nonresponders. Completed participants were offered 5,000 survey points (equivalent to $5) and the chance to win an in-kind prize through a monthly KN sweepstakes. When panelists were still unresponsive, additional panelists were randomly sampled to replace the nonresponders.

Data Collection

The study was administered to participants in four steps: informed consent, study arm assignment, DTC ad exposure, and questionnaire. Both an initial pretest and a main study were conducted. The goals of the pretest were to: (a) identify glitches in study administration, (b) preview the quality of participant responses, and (c) create codes for open-ended questions based on pretest responses. The data collection processes were identical for the pretest and main study. The pretest participants could not participate in the main study and data from the pretest was analyzed separately from the main study data. The pretest was fielded November 5-8, 2010, and the main study was fielded November 12-December 7, 2010.

Study Arm Assignment. The study was programmed to assign participants to one of 38 study arms. Each arm contained one different version of the Votrea DTC print or television ad.

Participants with broadband Internet connections were eligible for assignment to either the print or television arms. However, because of bandwidth and speed restrictions identified during pretesting, participants with dial-up or non-broadband Internet were assigned to the print ad arms only. This ensured that participants could view the television ad properly without start-and-stop video or major time delays. Participants' Internet connection speed was determined by KN programming at the time of consent. Within ad type (print or TV), participants were randomly assigned to study arm.

Ad Exposure. Once assigned to a study condition, participants viewed either a two-page print ad or a 90-second television ad for Votrea. Exposure to the ads was controlled and participants were not able to proceed to the survey without viewing the ad. Participants could not view the television or print ad again after moving beyond that part of the survey.

The print ad was a two-page, full-color magazine spread, and three different views of the ad were presented: two-page spread, first page only (display page), and second page only (important information/brief summary). Participants were allowed to navigate back and forth between the pages without time restrictions.

The television ad was a 90-second spot with two on-screen actors, one narrator (voiceover), and some on-screen text (supers). Based on pretest findings, participants were required to view the ad twice in succession. This ensured that video/audio problems did not interfere with ad exposure.

Questionnaire. Finally, participants were directed to a questionnaire about the ad and their perceptions of Votrea (see Appendix B for questionnaire). Once participants accessed the

questionnaire, they could no longer return to the ad. The questionnaire consisted of 40 items (including several sub-items) and presented one question per screen. Questionnaire navigation was disabled so that participants were not able to return to previous questions or amend previous answers. The questionnaire was cognitively tested prior to the pretest and refined accordingly.

Analyses

Study Sample Weighting. Weighting is a common procedure in survey studies that ensures the study data more accurately represent the target population of U.S. adults and helps to account for nonresponse, noncoverage, underrepresentation of minority groups, and other types of sampling and survey error. For example, if the study sample has fewer African Americans than the target population, weights would be applied to correct this discrepancy. African Americans' responses would be given more importance relative to other participants' responses and would have a relatively larger influence on statistical output, such as means and percentages.

For this study, three types of weights were constructed and applied to the dataset:

- *Base Weight*—Accounts for deviations from pure probability sampling in KnowledgePanel's sampling procedures. This weight was constructed prior to study sampling.
- *Panel Post-Stratification Weight*—Accounts for nonresponse and noncoverage effects in the entire KnowledgePanel. This weight was constructed prior to study sampling.
- *Study Post-Stratification Weight*—Accounts for nonresponse and undersampling or oversampling in the study sample. This weight was constructed at the conclusion of data collection.

First, although KN randomly sampled KnowledgePanel members through RDD and ABS methods, it incorporated several strategies to improve efficiency and boost representation of age

groups and racial minorities. These strategies, while valuable, were a deviation from equal probability sampling. Consequently, statistical weighting adjustments were constructed to offset these sampling deviations. These adjustments were the sample's *base weight*.

Second, as with any survey, KnowledgePanel was subject to several types of survey error. These errors occurred because not all households could be reached (for instance, because not all households have telephones; this is called noncoverage), not all households accepted the panel invitation (nonresponse), and not all panelists chose to continue participating after enrollment (attrition). These sources of error were addressed by creating a *panel post-stratification weight*.

Finally, this study sampled and invited a number of panelists who either did not respond (nonresponse) or declined to participate (nonconsent) in the study. While the nonresponse analysis suggested that participants and nonrespondents were not significantly different, nevertheless a *study post-stratification weight* was constructed to account for minor discrepancies.

Hypothesis Testing. We examined Analyses of Variance (ANOVAs) to test whether perceived benefit, attitude toward the ad, behavioral intentions, risk-benefit tradeoff, risk recall (aided), and perceived risk differed by format (statistical or visual) and drug efficacy level. Logistic regression models and chi-square tests were examined to test whether format (statistical or visual) and drug efficacy level affected benefit accuracy and risk recall (unaided). Post-hoc contrasts were examined to test hypotheses comparing formats (statistical or visual). For contrasts, we used Bonferroni adjusted *p* values, defining significance as follows:

- $p < 0.003$ for comparisons among statistical formats,
- $p < 0.005$ for comparisons among visual formats, and

- $p < 0.02$ for comparisons between the drug efficacy conditions and the control condition.

For all other tests we defined significance as $p < 0.05$.[6]

To test our hypotheses about numeracy as a moderator of format effects (Statistical Format Hypothesis 5 and Visual Format Hypothesis 2), we repeated the relevant analyses with numeracy included as a covariate and an interaction term between numeracy and format (statistical or visual). To test our hypotheses about numeracy as a moderator of drug efficacy level effects (Drug Efficacy Hypothesis 4), we repeated the relevant analyses with numeracy included as a covariate and an interaction term between numeracy and drug efficacy level. Objective and subjective numeracy were examined in separate models.

To test the hypothesis that perceived benefit mediates the relation between drug efficacy level and perceived risk (Drug Efficacy Hypothesis 6), we examined a series of linear regression models predicting: (1) perceived benefit from drug efficacy level, (2) perceived risk from drug efficacy level, and (3) perceived risk from drug efficacy level and perceived benefit (Baron and Kenny, 1986). As an additional way to test the meditational hypothesis, we conducted a Sobel test (MacKinnon, Warsi, et al. 1995) to determine whether the indirect effect of drug efficacy level on perceived risk via perceived benefit was significantly different from zero.

All analyses were conducted with weighted data. In addition, non-response bias analyses suggested that the study did not suffer from non-response bias, and outlier analyses suggested that outliers had no significant effect on the study results. Consequently, all participants were included in the hypothesis testing analyses.

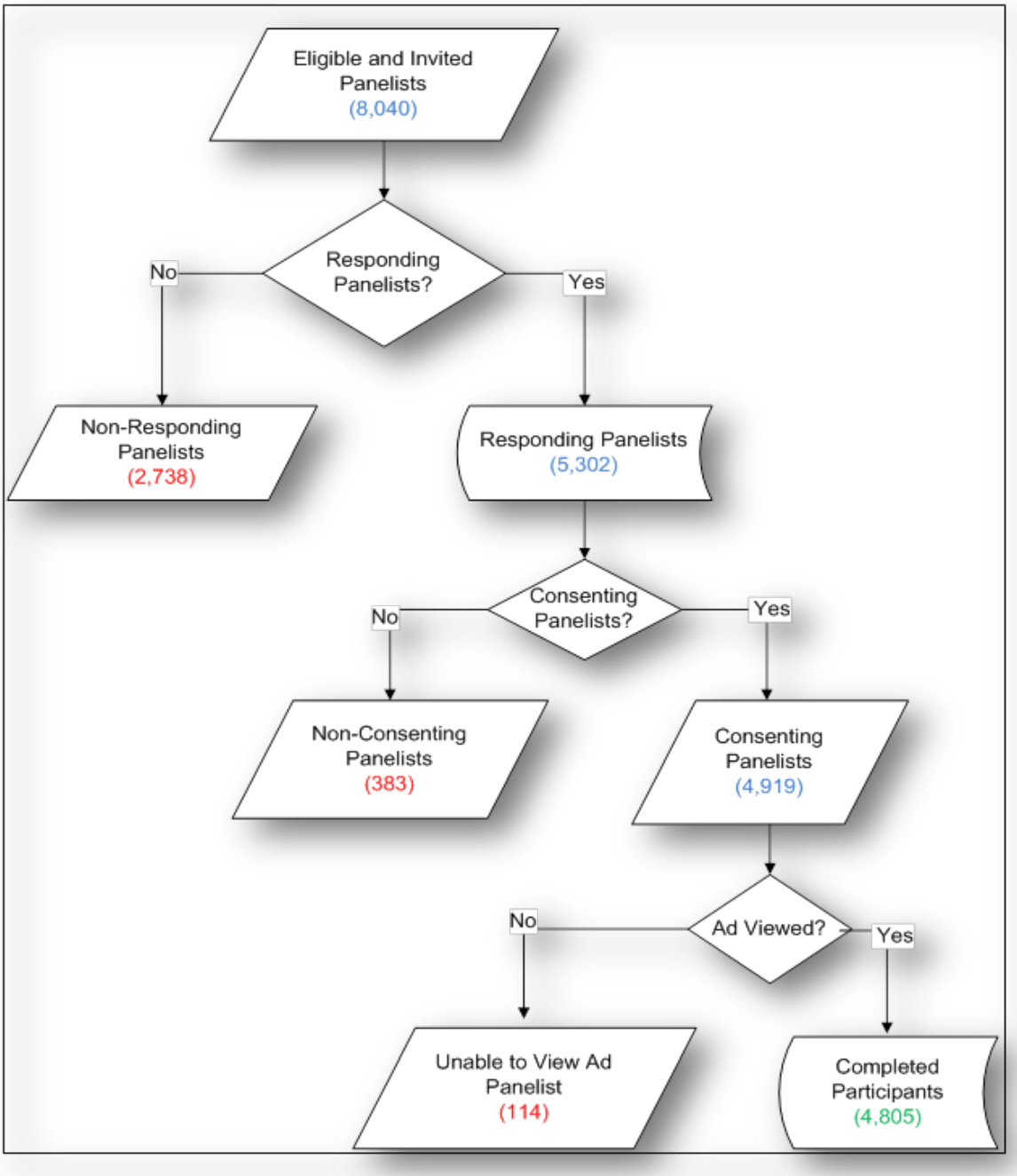
Results and Discussion

[6] Multiple *p* values are reported as *p*s, as in "Several differences were found between the control condition and the statistical format conditions, *p*s < 0.003."

Participants

A total of 8,040 panelists were invited to participate in the study. Of those panelists, 4,805 ultimately consented to and completed the study, resulting in an overall response rate of 60% (Figure 5). All participants self-reported that they had been medically diagnosed with high cholesterol. About half were women and the majority was white, over 55 years of age, and had some college education or more (see Table A, Appendix C, for demographic characteristics).

Figure 5. Study Sample Flowchart.



Statistical Format Results

To facilitate this section, Table 3 restates the statistical format conditions in the design and Table 4 gives an overview of the hypotheses and study findings.

Table 3. Statistical Formats and Their Wording

Statistical Format	Wording
[Absolute Frequency]	With Votrea, 65 out of 100 people lowered their bad cholesterol to normal levels versus 2 out of 100 people with no treatment.
[Percent]	With Votrea, 65% of people lowered their bad cholesterol to normal levels versus 2% of people with no treatment.
[Relative Frequency]	With Votrea, people were 33 times more likely to lower their bad cholesterol to normal levels compared to no treatment.
[Absolute Frequency + Percent]	With Votrea, 65 out of 100 people (65%) lowered their bad cholesterol to normal levels versus 2 out of 100 people (2%) with no treatment.
[Relative Frequency + Percent]	With Votrea, 65% of people lowered their bad cholesterol to normal levels versus 2% of people with no treatment—that's 33 times more effective.
Control	Votrea lowers bad cholesterol to normal levels.

Table 4. List of Statistical Format Hypotheses and Findings

Statistical Format Hypothesis	Was the hypothesis supported?	
	Print	**Television**
(S1) We will test competing hypotheses for behavioral intentions, attitude toward drug, and perceived benefit.		
(S1a) **Overestimation hypothesis**: The first hypothesis rests on the assumption that, in the absence of any quantitative information, participants overestimate the effectiveness of drugs. Accordingly, we would predict that behavioral intentions, attitude toward drug, and perceived benefit will	No	No

be **higher** for participants in the *no* statistical format condition, compared to *all other* statistical format conditions. Support for this interpretation will be found if estimates of the benefits are higher in the *no* statistical format condition than in *all other* statistical format conditions.		
(S1b). **Peripheral cue hypothesis**: The competing hypothesis rests on the assumption that any statistical information will be used as a peripheral cue; that is, participants will not process the quantitative information provided in the various statistical formats but will rather view it as scientific proof of the drug's benefit. Accordingly, we would predict that behavioral intentions, attitude toward drug, and perceived benefit will be **lower** for participants in the *no* statistical format condition, compared to *all other* statistical format conditions. Support for this interpretation will be found if, in addition to perceived benefit effects, estimates on attitude toward the ad peripheral cue measures—ratings of how believable, persuasive, informative, etc, the ad is—are lower in the *no* statistical format condition than in *all other* statistical format conditions.	No	No
(S2) Based on the risk communication literature, we predict that: • [Absolute Frequency], [Percent], and [Absolute Frequency + Percent] conditions may not differ on behavioral intentions, attitude toward drug, or perceived benefit. • Behavioral intentions, attitude toward drug, and perceived benefit will be higher in the [Relative Frequency] condition than in the [Absolute Frequency], [Percent], [Absolute Frequency + Percent], and [Relative Frequency + Percent] conditions.	Yes No	Yes No
(S3) The effects tested in hypotheses (S1) & (S2) will be modified by numeracy.	Not tested due to nonsignificant findings in (S2)	Not tested due to nonsignificant findings in (S2)
(S4) Benefit accuracy will be lowest in the *no* statistical format condition and highest in the [Absolute Frequency] condition. • Tests of other relations between statistical formats were exploratory.	Partially	Partially
(S5) The effects tested in (S4) will be modified by numeracy, such that low numeracy participants will show greater differences in	No	No

benefit accuracy across statistical formats than will high numeracy participants.		
(S6) We expect that risk recall will *not* differ by statistical format, but we will conduct exploratory analyses to determine whether information overload impedes risk recall.	Yes	Yes
(S7) We expect that perceived risk will be lowest in the [Relative Frequency] condition if perceived benefit is indeed highest in this condition (Slovic and Peters, 2006).	Yes	No

Note. This table represents a simplified presentation of hypotheses and support. Please refer to the text for complete details.

Did statistical format affect behavioral intentions, attitude toward the drug, and perceived benefit? (Statistical Format Hypotheses S1 and S2)

Statistical Print Conditions. There was no significant difference among statistical formats on behavioral intentions, attitude toward the drug, or perceived benefit, *p*s > 0.05 (see Table B, Appendix C, for means).

Statistical Television Conditions. There was no significant difference among statistical formats on behavioral intentions, attitude toward the drug, or perceived benefit, *p*s > 0.05 (see Table C, Appendix C, for means).

Discussion. We predicted that either behavioral intentions, attitude toward drug, and perceived benefit would be **higher** (because participants may overestimate the effectiveness of drugs in the absence of any information) or **lower** (because the statistical information may be used as a peripheral cue that the drug has scientific proof) for participants in the *no* statistical format condition, compared to *all other* statistical format conditions. We did not find support for either hypothesis. It appears that participants neither overestimated the benefit of the drug when they viewed a drug ad with no quantitative benefit information nor did they use the quantitative benefit information as a peripheral cue when given this information in various statistical formats.

Because we did not find significant results, we did not test the role of numeracy as a moderator, as our third statistical format hypothesis (S3) proposed.

Did statistical format affect benefit accuracy? (Statistical Format Hypotheses S4)

Statistical Print Condition. Several differences were found between the control condition and the statistical format conditions, *p*s < 0.003 (see Table D for percents and Table E for test statistics in Appendix C). Recall that each participant was asked accuracy questions in four different ways: absolute frequency, percentage, relative frequency, and placebo (see Table 2). Participants in the [Absolute Frequency], [Percent], [Absolute Frequency + Percent], and [Relative Frequency + Percent] conditions (i.e., all statistical formats except relative frequency), compared with participants in the control condition, were more likely to accurately report the benefits of the drug when asked to report it in the absolute frequency, percentage, and placebo formats. In addition, participants in the [Relative Frequency] and [Relative Frequency + Percent] conditions, compared with participants in the control condition, were more likely to accurately report the benefits of the drug when asked to report it in the relative frequency format.

There were also several differences found when comparing among the statistical formats, *p*s < 0.003 (see Table D for percents and Table F for test statistics in Appendix C). Participants in the [Absolute Frequency], [Percent], [Absolute Frequency + Percent], and [Relative Frequency + Percent] conditions, compared with participants in the [Relative Frequency] condition, were more likely to accurately report the benefits of the drug when asked to report it in the absolute frequency, percent, and placebo formats.

In addition, participants in the [Absolute Frequency] condition, compared with participants in the [Relative Frequency + Percent] condition, were more likely to accurately report the benefits of the drug when asked to report it in the absolute frequency format.

Participants in the [Absolute Frequency + Percent] conditions, compared with participants in the [Relative Frequency + Percent] condition, were more likely to accurately report the benefits of the drug when asked to report it in the absolute frequency and percentage formats. Finally, participants in the [Relative Frequency] and [Relative Frequency + Percent] conditions, compared with participants in the [Absolute Frequency] and [Absolute Frequency + Percent] conditions, were more likely to accurately report the benefits of the drug when asked to report in the relative frequency format. No other comparisons were significant. These results are depicted graphically in Figure 6.

Statistical television condition. Several differences were found between the control condition and the statistical format conditions, *p*s < 0.003 (see Table G for percentages and Table H for test statistics in Appendix C). Participants in the [Absolute Frequency], [Percent], [Absolute Frequency + Percent], and [Relative Frequency + Percent] conditions (i.e., all statistical formats except [Relative Frequency]), compared with participants in the control condition, were more likely to accurately report the benefits of the drug when asked to report it in the absolute frequency, percentage, and placebo formats. In addition, participants in the [Relative Frequency] and [Relative Frequency + Percent] conditions, compared with participants in the control condition, were more likely to accurately report the benefits of the drug when asked to report it in the relative frequency format.

There were also several differences found when comparing among the statistical formats *p*s < 0.003 (see Table G for percentages and Table I for test statistics in Appendix C). Participants in the [Absolute Frequency], [Percent], [Absolute Frequency + Percent], and [Relative Frequency + Percent] conditions, compared with participants in the [Relative

Frequency] condition, were more likely to accurately report the benefits of the drug when asked to report it in the absolute frequency, percentage, and placebo formats.

In addition, participants in the [Absolute Frequency + Percent] conditions, compared with participants in the [Relative Frequency + Percent] condition, were more likely to accurately report the benefits of the drug when asked to report it in the absolute frequency and percentage formats. Participants in the [Absolute Frequency + Percent] conditions, compared with participants in the [Absolute Frequency] condition, were more likely to accurately report the benefits of the drug when asked to report it in the [Absolute Frequency] format. Finally, participants in the [Relative Frequency + Percent] condition, compared with participants in the [Absolute Frequency], [Percent], and [Absolute Frequency + Percent] conditions, were more likely to accurately report the benefits of the drug when asked to report it in the relative frequency format. No other comparisons were significant. These results are depicted graphically in Figure 6.

Figure 6. Description of Statistical Format Effects on Benefit Accuracy.

IN PRINT:

THESE FORMATS	were better than	THESE FORMATS	when asked	THESE QUESTIONS
Absolute Frequency Percent Absolute Frequency + Percent Relative Frequency + Percent		Control Relative Frequency		Absolute Frequency Percent Placebo
Relative Frequency Relative Frequency + Percent		Control		Relative Frequency
Absolute Frequency Absolute Frequency + Percent		Relative Frequency + Percent		Absolute Frequency
Absolute Frequency + Percent		Relative Frequency + Percent		Percent
Relative Frequency Relative Frequency + Percent		Absolute Frequency Absolute Frequency + Percent		Relative Frequency

IN TELEVISION:

THESE FORMATS	were better than	THESE FORMATS	when asked	THESE QUESTIONS
Absolute Frequency Percent Absolute Frequency + Percent Relative Frequency + Percent		Control Relative Frequency		Absolute Frequency Percent Placebo

Relative Frequency Relative Frequency + Percent		Control		Relative Frequency
Absolute Frequency + Percent		Relative Frequency + Percent		Absolute Frequency Percent
Absolute Frequency + Percent		Absolute Frequency		Absolute Frequency
Relative Frequency + Percent		Absolute Frequency Percent Absolute Frequency + Percent		Relative Frequency

Discussion. As predicted, in both print and television ads, participants who saw any type of statistical format were more accurate in their recall of the benefit information than those in the control condition who did not see any quantitative benefit information at all. This finding supports the concept that people perform better when they are given the information about which they are asked than if they are given no information (e.g., Schwartz et al. 2009). Participants who received quantitative information in the current study appear to have processed the information.

There was also an expected match between the format participants viewed and the type of information they were able to accurately report. Participants who viewed the benefit information in an [Absolute Frequency] and/or [Percent] format were more likely to accurately report the absolute frequency and percentage information than participants in the control condition, but they were not more likely to accurately report the relative frequency information. Participants who viewed the benefit information in the [Relative Frequency] or [Relative Frequency + Percent] format were more likely to accurately report the relative frequency information than participants in the control condition, although participants who only viewed the [Relative Frequency] benefit information were not more likely to accurately report the absolute frequency or percentage information. Thus, participants could report the numbers they saw, but they were not necessarily capable of transforming the numbers into different formats. This, again, suggests that easy information was useful for participants, but information that requires calculation or added thought is not necessarily useful in this context.

Keeping in mind the match between statistical format and question, we found differences between certain statistical formats on benefit accuracy. We predicted that absolute frequency

would be the most easily processed format based on previous research (Slovic, Monahan, et al. 2000). We also suspected that some formats, such as the one that contained absolute frequency and percentage information, might cause information overload (Chandler and Sweller, 1991; Sweller 1988), leading to a reduction in benefit accuracy. We found, instead, that participants in the [Absolute Frequency] condition did not have the highest benefit accuracy. Descriptively, the combination of the two formats, absolute frequency and percentage ([Absolute Frequency + Percent]), actually resulted in the highest benefit accuracy on absolute frequency, percentage, and placebo accuracy. Conversely, [Relative Frequency + Percent] resulted in the most accurate reporting for relative frequency accuracy.

Rather than overload readers or viewers with information, the combination of [Absolute Frequency + Percent] resulted in better recall of the information. There are a number of possible reasons for this. First, the combination of formats may meet the numerical preferences and skills of a larger audience. The more numerical formats used, the more likely at least one of those formats will resonate with audience members. Second, for some participants, the two formats together may have served a repetitive function, further strengthening the memory trace (e.g., Thompson 2005). Some participants may respond well to one or the other format, and having both available may have allowed participants to use both information concepts together to recall the information. Finally, the combination of formats means that the quantitative information requires more space and more time to process. This extra time may make the quantitative information in these cases more noticeable to audience members, leading to increased accuracy.

It should be noted that these results are inconsistent with those of Woloshin and Schwartz (2011). They found that participants in their percent only condition performed as well on a comprehension measure as those in their percent + frequency condition, concluding that the

combination of formats is not worth the extra presentation of numbers. However, in their study, the format participants saw in the ad was the same format used in the comprehension measure. For example, participants who saw a percent format answered questions about percentages whereas participants who saw an absolute frequency format answered questions about absolute frequencies. Multiple formats may be useful when people need to translate numbers from one format to another (for example, understanding both percentages and relative frequencies). Further research should follow these studies to establish the boundary conditions of when multiple formats are and are not helpful. In any case, both our study and Woloshin and Schwartz's demonstrated that the absolute frequency alone condition does not result in the best performance.

Note specifically that adding percent to relative frequency improved benefit accuracy. Thus, if relative frequencies are included in presentations of numbers, adding an absolute frequency does increase the likelihood of accurate recall of the numbers. Whether the effect on accuracy translates to effects on the perceptions and actions related to this accuracy will be discussed in the General Discussion.

Was numeracy a moderator? (Statistical Format Hypothesis S5)

Statistical Print Conditions. Numeracy did not moderate the effect of statistical format on benefit accuracy, *p*s > 0.05. There was one exception: the interaction between statistical format and numeracy (objective and subjective) was significant when predicting placebo accuracy ($p = 0.005$ for both objective and subjective numeracy); however, this was likely an artifact of the control condition cell count (0).

Statistical Television Conditions. Numeracy did not moderate the effect of statistical format on benefit accuracy, *p*s > 0.05. There was one exception: the interaction between

statistical format and numeracy (objective and subjective) was significant when predicting placebo accuracy ($p < 0.001$ for both objective and subjective numeracy); however, this was likely an artifact of the control condition cell count (1).

Discussion. We expected that the effects of statistical format on benefit accuracy might be moderated by numeracy, such that some formats would result in higher accuracy, depending on how well participants could work with numbers. We did not find this to be the case. Thus, it appears that none of the statistical formats differentially helped or hindered participants with low numeracy.

Did statistical format affect risk recall and perceived risk? (Statistical Hypotheses S6 and S7).

Statistical Print Conditions. Statistical format did not affect aided or unaided risk recall, $p = 0.22$ and $p = 0.45$, respectively (see Table B for means and percents). As predicted, participants in the [Relative Frequency] condition had lower perceived risk than participants in the control condition, $p = 0.001$. However, although the perceived risk was lower in the [Relative Frequency] condition than in all other conditions, these differences did not reach significance when we used a more conservative statistic to adjust for the multiple comparisons we made (the Bonferroni-adjusted significance level; *p*s > 0.008; see Table B, Appendix C, for means).

Statistical Television Conditions. Statistical format did not affect aided or unaided risk recall, $p = 0.93$ and $p = 0.32$, respectively (see Table C, Appendix C, for means and percentages), or perceived risk, $p = 0.40$ (see Table C, Appendix C, for means and percentages).

Discussion. As expected, risk recall did not differ as a function of the statistical format of the benefit information. For perceived benefit, we hypothesized that if participants in the

[Relative Frequency] condition had the highest perceived benefit, they would show the lowest perceived risk, as suggested by the work of Slovic and Peters (2006). Although there were no statistically significant differences in perceived benefit across statistical format conditions, we did find that participants in the [Relative Frequency] condition who saw a print ad reported significantly lower perceived risk than those in the control condition. This was not the case for those who saw a television ad. This was our one finding in which print ad results were stronger than those for television ads.

Visual Format Results

To facilitate this section, Table 5 restates the visual format conditions in the design and Table 6 gives an overview of the hypotheses and findings.

Table 5. Visual Format Conditions

Pie Chart
Bar Graph
Pictograph
Table
Control

Table 6. List of Visual Format Hypotheses and Findings

Visual Format Hypotheses	**Was the hypothesis supported?**	
	Print	**Television**
(V1) We will test competing hypotheses for benefit accuracy, behavioral intentions, attitude toward drug, and perceived benefit.		
(V1a) **Visual information facilitation hypothesis**: The first hypothesis rests on the assumption that participants will, to the extent possible, process and use the information in the visual	Yes	Yes

formats. The risk communication literature suggests that visual representations of risk can increase understanding and that participants have a more difficult time processing this kind of information in pie charts, as compared to other visual formats. Therefore, our first hypothesis is that benefit accuracy will be higher in the *bar chart, table,* and *pictograph* conditions—but not necessarily the *pie chart* condition—than in the *no* visual format condition. Tests of other relations between visual formats will be exploratory.		
(V1b) **Information overload hypothesis**: Alternatively, there may be no differences across visual formats on behavioral intentions, attitude toward drug, perceived benefit, or benefit accuracy if the visual serves as a distraction or is too much information to process.	No	No
(V1c) **Peripheral cue hypothesis**: Behavioral intentions, attitude toward drug, and perceived benefit—but not benefit accuracy—may be higher in *all visual* conditions than in the *no* visual condition if the visual information serves as a peripheral cue.	No	No
(V2) The effects tested in (1) will be modified by numeracy.	No	No
(V3) We expect that perceived risk and risk recall will *not* differ by visual format, but we will conduct exploratory analyses to determine whether information overload impedes risk recall.	Yes	Yes

Note. This table represents a simplified presentation of hypotheses and support. Please refer to the text for complete details.

Did visual format affect behavioral intentions, attitude toward the drug, and perceived benefit? (Visual Format Hypothesis V1)

Visual print conditions. Visual format did not affect behavioral intentions, attitude toward the drug, or perceived benefit, *p*s > 0.05 (see Table J, Appendix C, for means and percents).

Visual Television Conditions. Visual format did not affect behavioral intentions, attitude toward the drug, or perceived benefit, *p*s > 0.05 (see Table K, Appendix C, for means and percents).

Discussion. One of our competing hypotheses (V1c) was that if the visual information serves as a peripheral cue, behavioral intentions, attitude toward drug, and perceived benefit—but not benefit accuracy—may be higher in *all visual* conditions than in the *no* visual condition. We found no significant differences on these outcomes across visual formats; therefore, this peripheral cue hypothesis was not supported. Because hypotheses V1a and V1b rest on our benefit accuracy findings, these competing hypotheses will be discussed next.

Did visual format affect benefit accuracy? (Visual Format Hypothesis V1)

Visual Print Conditions. Several differences were found between participants in the visual format conditions and participants who viewed the absolute frequency information with no accompanying visual (the no visual condition), *p*s < 0.005 (see Table L for percents and Table M for test statistics in Appendix C). Recall that each participant was asked accuracy questions in four different ways: absolute frequency, percentage, relative frequency, and placebo (see Table 2). Participants in the bar chart and table conditions, compared with participants in the no visual condition, were more likely to accurately report the benefits of the drug when asked to report it in the percentage, relative frequency, and placebo formats. In addition, participants in the bar chart condition, compared with participants in the no visual condition, were more likely to accurately report the benefits of the drug when asked to report it in the absolute frequency format. No other differences were significant.

There were also several differences found when comparing among the visual formats, *p*s < 0.005 (see Table L for percents and Table N for test statistics in Appendix C). Participants in

the bar chart condition, compared to participants in the pictograph condition, were more likely to accurately report the benefits of the drug when asked to report it in the absolute frequency and percentage formats. Participants in the table condition, compared with participants in the pie chart condition, were more likely to accurately report the benefits of the drug when asked to report it in the placebo format. No other comparisons were significant. These results are depicted graphically in Figure 7.

Visual Television Conditions. Several differences were found between the visual format conditions and the no visual condition), *p*s < 0.005 (see Table O for percents and Table P for test statistics in Appendix C). Participants in the bar chart, table, pie chart, and pictograph conditions (i.e., all conditions that contained a visual), compared with participants in the no visual condition, were more likely to accurately report the benefits of the drug when asked to report it in the absolute frequency, percentage, and placebo formats. Participants in the bar chart condition, compared with participants in the no visual condition, were more likely to accurately report the benefits of the drug when asked to report it in the relative frequency format.

There were also several differences found when comparing among the visual formats, *p*s < 0.005 (see Table O for percents and Table Q for test statistics in Appendix C). Participants in the bar chart condition, compared with participants in the pictograph condition, were more likely to accurately report the benefits of the drug when asked to report it in the absolute frequency, percentage, and relative frequency formats. Participants in the bar chart condition, compared with participants in the table condition, were more likely to accurately report the benefits of the drug when asked to report it in the absolute frequency format. These results are depicted graphically in Figure 7.

Figure 7. Description of Visual Format Effects on Benefit Accuracy.

IN PRINT:

THESE FORMATS	**were better than**	**THESE FORMATS**	**when asked**	**THESE QUESTIONS**
Bar Chart Table		No Visual		Percent Relative Frequency Placebo
Bar Chart		No Visual		Absolute Frequency
Bar Chart		Pictograph		Absolute Frequency Percent
Table		Pictograph		Percent
Table		Pie Chart		Placebo

IN TELEVISION:

THESE FORMATS	**were better than**	**THESE FORMATS**	**when asked**	**THESE QUESTIONS**
Bar Chart Table Pie Chart Pictograph		No visual		Absolute Frequency Percent Placebo
Bar Chart		No visual		Relative frequency
Bar Chart		Pictograph		Absolute Frequency Percent Relative Frequency
Bar chart		Table		Absolute Frequency

Discussion. Our visual information facilitation hypothesis (V1a) rested on the assumption that participants would process and use the information in the visual formats. Based on the risk communication literature, we predicted that benefit accuracy would be higher in the *bar chart*, *table*, and *pictograph* conditions—but not necessarily the *pie chart* condition—than in the *no* visual format condition. Our competing hypotheses were that: (1) there may be no

difference across conditions on any of the outcome measures (including benefit accuracy) if the visuals create information overload; or (2) if the visual information serves as a peripheral cue (V1b), behavioral intentions, attitude toward drug, and perceived benefit (but not benefit accuracy) may be higher in *all visual* conditions than in the *no* visual condition. Our results do not support the information overload (V1b) or peripheral cue hypotheses (V1c). We did find support for the visual information facilitation hypothesis (V1a). In the print conditions, participants who viewed the ad with the bar chart and participants who viewed the ad with the table were more accurate than those who viewed the ad with no visual. In the television conditions, participants who viewed any visual aid were more accurate than those who did not see any visual information.

Although these results support the idea that visuals increase participants' ability to understand quantitative benefit information, the pattern of results was not as expected based on the risk communication literature. In particular, the pie chart was not the least helpful nor was the pictograph the most helpful. In fact, the bar chart and table were the most consistently helpful, and in some cases the bar chart was more helpful than the pictograph. These findings are surprising given previous research on pictographs that suggests that pictographs help individuals understand numerical or medical information (Fagerlin, Wang, et al. 2011; Houts, Witmer, et al. 2001; Zkmund-Fisher, Ubel, et al. 2008). In particular, at least two previous studies compared pictographs to other display formats and found that pictographs improved risk communication (Zikmund-Fisher et al. 2008) and eliminated the effects of anecdotal information on treatment decisions (Fagerlin et al. 2005). It is possible that our findings are the result of the pictograph being the only format used in the study without actual numbers included as part of the visual, although all participants saw the absolute frequency information in the text. These

differences in results may also be a function of the various outcome measures that were used—whereas we tested participants' accuracy in recalling the numbers, the previous studies examined decisions. We did find, however, that only participants in the bar chart and table in print and the bar chart in television were more accurate on the relative frequency question compared with the no visual condition. This provides evidence that the bar chart and table are providing information that may be useful beyond simple accuracy; they may have influences on the transformation of numbers. This numerical transformation is something that pictographs in our study did not appear to help.

Was numeracy a moderator? (Visual Format Hypothesis V2)

Visual print conditions. Objective and subjective numeracy did not moderate the relation between visual format and benefit accuracy, *p*s > 0.05.

Visual television conditions. Objective and subjective numeracy did not moderate the relation between visual format and benefit accuracy, *p*s > 0.05.

Discussion. We expected the effects of visual format to be moderated by numeracy, such that some formats would result in higher accuracy, for instance, depending on how well participants could work with numbers. As with the drug efficacy level manipulations, we did not find support for this hypothesis. Thus, it appears that none of the visual formats especially helped or hindered participants with low numeracy.

Did visual format affect risk recall and perceived risk? (Visual Format Hypothesis V3)

Visual print conditions. Visual format did not affect risk recall or perceived risk, *p*s > 0.05 (see Table J for means and percents).

Visual television conditions. Visual format did not affect risk recall or perceived risk, *p*s > 0.05 (see Table K for means and percents).

Discussion. As expected, we did not find differences among visual formats in risk recall or risk perceptions. This suggests that adding specific information in clear visual forms about the benefit side of the prescription drug equation does not interfere with the processing of the risk information for the drug.

Drug Efficacy Level Results

To facilitate this section, Table 7 gives an overview of the hypotheses and findings.

Table 7. List of Efficacy Hypotheses and Findings

Drug Efficacy Hypothesis	**Was the hypothesis supported?**			
	Statistical Print	**Statistical Television**	**Visual Print**	**Visual Television**
(E1) Behavioral intentions, attitude toward drug, and perceived benefit will be higher in *high* drug efficacy conditions than in *low* drug efficacy conditions.	No	Yes	Yes	Yes
(E2) We will explore whether there are differences between the *no* efficacy condition (control) and the *low* and *high* drug efficacy condition on behavioral intentions, attitude toward drug, and perceived benefit.	No	Yes	Yes	Yes
(E3) Benefit accuracy will be higher in the *low* and *high* drug efficacy conditions than in the *no* efficacy condition. There will be no difference between the *low* and *high* drug efficacy conditions.	Yes	Yes	Yes	Yes
(E4) The effects tested in 1 & 3 will be modified by numeracy, such that high numeracy participants will be more likely to show these effects than will low numeracy participants.	No	Yes	No	Yes
(E5) Risk recall will not differ by drug efficacy level (no, low, high).	Yes	Yes	Yes	Yes
(E6) Perceived risk will be lower in the high drug efficacy condition compared with the low drug efficacy condition	No	Yes	Yes	Yes

Note. This table represents a simplified presentation of hypotheses and support. Please refer to the text for complete details.

Did drug efficacy affect behavioral intentions, attitude toward the drug, and perceived benefit? (Drug Efficacy Hypotheses E1 and E2)

Statistical Print Conditions. There were no differences on behavioral intentions, attitude toward the drug, or perceived benefit among the drug efficacy conditions, *p*s > 0.05 (see Table R, for means and Table S for test statistics in Appendix C). However, participants who viewed the high-efficacy drug ad (high-efficacy condition) were more likely to rate the drug as having more benefits than risks compared with participants who viewed the low-efficacy drug ad (low-efficacy condition), $p = 0.01$.

Statistical Television Conditions. Participants in the high-efficacy condition had marginally higher perceived benefit, $p = 0.055$, and a more positive attitude toward the drug than did participants in the low-efficacy condition, $p = 0.002$ (see Table R for means and Table S for test statistics in Appendix C). In addition, participants in the high-efficacy condition were more likely to rate the drug as having more benefits than risks compared with participants in the low-efficacy condition, $p = 0.002$. They were also more likely to intend to take the drug if prescribed, $p = 0.03$, and search for more information, $p = 0.04$. This last effect was qualified by an interaction with statistical format, $p = 0.01$. A visual inspection of the weighted means suggests that this finding was greatest in the [Relative Frequency] condition (low drug efficacy: $M = 1.82$, $SE = 0.09$; high drug efficacy: $M = 2.37$, $SE = 0.13$), and was reversed in the [Absolute Frequency + Percent] condition (low drug efficacy: $M = 2.20$, $SE = 0.12$; high drug efficacy: $M = 1.98$, $SE = 0.10$). There was no significant difference for intention to talk to the doctor between high- and low-efficacy conditions, $p > 0.05$.

Participants in the high-efficacy condition had a more positive attitude toward the drug than participants who viewed the control drug ad with no quantitative benefit information (control condition), $p = 0.01$. There were no significant differences on perceived benefit or behavioral intentions between participants in the high-efficacy and control conditions, and there were no significant differences on any of these measures between participants in the low-efficacy and control conditions, ps > 0.025.

Visual Print Conditions. Participants in the high-efficacy condition had higher perceived benefit, $p < 0.001$, and a more positive attitude toward the drug than participants in the low-efficacy condition, $p < 0.001$ (see Table T for means and Table U for test statistics in Appendix C). In addition, participants in the high-efficacy condition were more likely to rate the drug as having more benefits than risks compared with participants in the low-efficacy condition, $p < 0.001$. There were no significant differences between the high- and low-efficacy conditions for behavioral intentions, ps > 0.05.

Participants in the high-efficacy condition had a more positive attitude toward the drug than participants in the control condition, $p = 0.002$. There were no significant differences on perceived benefit or behavioral intentions between participants in the high-efficacy and control conditions, and there were no significant differences on any of these measures between participants in the low-efficacy and control conditions, ps > 0.025.

Visual Television Conditions. Participants in the high-efficacy condition had higher perceived benefit, $p < 0.001$, and a more positive attitude toward the drug, $p < 0.001$, than participants in the low-efficacy condition (see Table T for means and Table U for test statistics in Appendix C). In addition, participants in the high-efficacy condition were significantly more likely to rate the drug as having more benefits than risks than were participants in the low-

efficacy condition, $p < 0.001$. They were also more likely to intend to ask their doctor about the drug, $p = 0.02$, and to take the drug if prescribed, $p < 0.001$. There was no difference on the intention to seek out more information about the drug between the high- and low-efficacy conditions, $p > 0.05$.

Participants in the high-efficacy condition had higher perceived benefit, $p = 0.01$, and a more positive attitude toward the drug, $p < 0.001$, than participants in the control condition. There were no significant differences on behavioral intentions between participants in the high-efficacy and control conditions, and there were no significant differences on any of these measures between participants in the low-efficacy and control conditions, *p*s > 0.025.

Discussion. Our first hypotheses concerned whether the efficacy of the drug would influence perceived benefit, attitude toward the drug, and intention to pursue several behaviors including asking a doctor about the drug and looking for more information on the drug. We compared participants who saw high and low drug efficacy (Drug Efficacy Hypothesis E1) and participants who saw any numerical efficacy with participants in the control condition, who did not receive any numerical information about the efficacy of Votrea (Drug Efficacy Hypothesis E2).

These findings provide support that our drug efficacy manipulation was noticed and did cause expected responses. Overall, it appears that participants who viewed the high-efficacy drug ad thought the drug was more effective and had more positive attitudes toward the drug, compared with participants who viewed the low-efficacy drug ad. This was the case for participants in the statistical television, visual print, and visual television conditions, but not for the statistical print condition. This may be because text does not attract attention and aid processing of information as well as graphics do (Smerecnik, Mesters, et al. 2010).

In addition, participants in the high-efficacy conditions of both the statistical and visual television groups had greater behavioral intentions than participants in the low-efficacy conditions. The television ad was set to play in its entirety before proceeding to the survey; therefore, participants who were in the television ad conditions could not self-pace exposure to the ad whereas the print participants could pass through the print pages as quickly as they wanted to on their way to the rest of the study. This could explain the strength of the television findings. It is also possible that the multiple modes in a television ad (visual and audio) that are absent in a print ad create additional memory traces that then translate into deeper processing of the information, as demonstrated by behavioral intentions (e.g., Lang 1995; Paivio 1983).

Participants in all except the statistical print condition demonstrated greater positive attitudes when they viewed the high-efficacy drug ad compared with participants who viewed the control drug ad with no quantitative benefit information. In addition, those in the visual television conditions who viewed the high-efficacy drug ad thought the drug was more effective compared with those in the control condition. In general, those who did not see any quantitative benefit information responded more similarly to those in the low-efficacy condition rather than the high-efficacy condition on these variables. Thus, we did not find support for the concern that participants assume a high level of drug efficacy when not presented with specific quantitative benefit data about a prescription drug. This finding is counter to previous studies that found that, in the absence of quantitative information, participants assume a drug is highly effective (e.g., Schwartz et al. 2009).

Did drug efficacy affect benefit accuracy? (Drug Efficacy Hypothesis E3)

Statistical Print Conditions. As expected, participants in the high- and low-efficacy conditions, compared with participants in the control condition, were more likely to accurately

report the benefits of the drug when asked to report it in the absolute frequency, percentage, and placebo formats, *p*s < 0.001 (see Table V for percents and Table W for test statistics). In addition, participants in the low-efficacy condition, compared with participants in the control condition, were more likely to accurately report the benefits of the drug when asked to report it in the [Relative Frequency] format. Participants in the high-efficacy condition did *not* differ from participants in the control condition on the relative frequency accuracy measure.

Unexpectedly, participants in the low-efficacy condition, compared with participants in the high-efficacy condition, were more likely to accurately report the benefits of the drug when asked to report it in the relative frequency format. This was qualified by an interaction between drug efficacy level and statistical format, $p = 0.04$. A visual inspection of the weighted means suggests that the difference between the low- and high-efficacy conditions was smallest in the [Relative Frequency] (high drug efficacy: 17.54%; low drug efficacy: 29.7%) and [Relative Frequency + Percent] (high drug efficacy: 21.15%; low drug efficacy: 31.99%) conditions. There was also an interaction between drug efficacy level and statistical format when predicting absolute frequency accuracy, $p = 0.04$. An examination of the weighted means suggests that participants in the low-efficacy conditions, compared with participants in the high-efficacy conditions, were more likely to accurately report the benefits of the drug when asked to report it in the absolute frequency format in the [Absolute Frequency] (high drug efficacy: 30.97%; low drug efficacy: 43.55%), [Relative Frequency] (high drug efficacy: 0.23%; low drug efficacy: 2.79%), and [Relative Frequency + Percent] (high efficacy: 18.80%; low efficacy: 26.99%) conditions, but not in the [Percent] (high drug efficacy: 34.57%; low drug efficacy: 27.24%) or [Absolute Frequency + Percent] (high drug efficacy: 43.94%; low drug efficacy: 39.74%)

conditions. Participants in the high- efficacy condition did *not* differ from participants in the low-efficacy condition on the percentage or placebo accuracy measures.

Statistical Television Conditions. As expected, participants in the high- and low-efficacy conditions, compared with participants in the control condition, were more likely to accurately report the benefits of the drug when asked to report it in the absolute frequency, percentage, and placebo formats, *p*s < 0.001 (see Table V for percents and Table W for test statistics in Appendix C). In addition, participants in the low-efficacy condition, compared with participants in the control condition, were more likely to accurately report the benefits of the drug when asked to report it in the relative frequency format. Participants in the high-efficacy condition did *not* differ from participants in the control condition on the relative frequency accuracy measure.

Unexpectedly, participants in the low-efficacy condition, compared with participants in the high-efficacy condition, were more likely to accurately report the benefits of the drug when asked to report it in the relative frequency format. This was qualified by an interaction between drug efficacy level and statistical format, $p = 0.003$. An examination of the weighted means suggests that the effect was smallest in the [Relative Frequency + Percent] condition (high drug efficacy: 32.53%; low drug efficacy: 33.55%). There was also an interaction between efficacy and statistical format when predicting placebo accuracy, $p = 0.01$. An examination of the weighted means suggests that participants in the low-efficacy condition, compared with participants in the high-efficacy condition, were more likely to accurately report the benefits of the drug when asked to report it in the placebo format in the [Absolute Frequency] (high efficacy: 25.48%; low efficacy: 37.36%), [Percent] (high drug efficacy: 25.11%; low drug efficacy: 40.29%), and [Relative Frequency] (high drug efficacy: 0.00%; low drug efficacy:

0.82%) conditions, but that the effect was reversed in the [Absolute Frequency + Percent] (high drug efficacy: 43.08%; low drug efficacy: 30.85%) and [Relative Frequency + Percent] (high drug efficacy: 29.86%; low drug efficacy: 20.25%) conditions. Participants in the high-efficacy condition did *not* differ from participants in the low-efficacy condition on the absolute frequency or percentage accuracy measures.

Visual Print Conditions. As expected, participants in the high- and low-efficacy conditions, compared with participants in the control condition, were more likely to accurately report the benefits of the drug when asked to report it in the absolute frequency, percentage, and placebo formats, *p*s < 0.001 (see Table X for percents and Table Y for test statistics). In addition, participants in the low-efficacy condition, compared with participants in the control condition, were more likely to accurately report the benefits of the drug when asked to report it in the relative frequency format. Participants in the high-efficacy condition did *not* differ from participants in the control condition on the relative frequency accuracy measure.

Unexpectedly, participants in the low-efficacy condition, compared with participants in the high-efficacy condition, were more likely to accurately report the benefits of the drug when asked to report it in the absolute frequency, percentage, and relative frequency formats. Participants in the high-efficacy condition did *not* differ from participants in the low-efficacy condition on the placebo accuracy measure.

Visual Television Conditions. As expected, participants in the high- and low-efficacy conditions, compared with participants in the control condition, were more likely to accurately report the benefits of the drug when asked to report it in the absolute frequency, percentage, and placebo formats, *p*s < 0.001 (see Table X for percentages and Table Y for test statistics in Appendix C). In addition, participants in the low-efficacy condition, compared with participants

in the control condition, were more likely to accurately report the benefits of the drug when asked to report it in the relative frequency format. Participants in the high-efficacy condition did *not* differ from participants in the control condition on the relative frequency accuracy measure.

Unexpectedly, participants in the low-efficacy condition, compared with participants in the high-efficacy condition, were more likely to accurately report the benefits of the drug when asked to report it in the absolute frequency, percentage, and relative frequency formats. Participants in the high-efficacy condition did *not* differ from participants in the low-efficacy condition on the placebo accuracy measure.

Discussion. We investigated whether drug efficacy level would influence whether participants could accurately report the quantitative benefit information provided in the ad. We hypothesized that benefit accuracy would be higher for participants who viewed either the high- or low-efficacy drug ad compared with those who viewed the control drug ad with no quantitative benefit information, and this was supported. This finding—that participants perform better when they are given the information about which they are asked than if they are given no information—is consistent with prior research (e.g., Schwartz et al. 2009). Participants who received quantitative information in the current study appear to have processed the information. This was the case for both print and television ads.

There is one caveat to the conclusion above: participants in the high drug efficacy condition were *not* more likely, compared with participants in the control condition, to accurately report the benefits of the drug when asked to report it in the relative frequency format. None of the participants in the visual conditions saw relative frequency information, and only participants in two of the statistical conditions saw relative frequency information. This meant that most participants were required to perform a calculation (for instance, 10/2 = 5 times more likely) to

report the relative frequency. The low drug efficacy numbers may have been easier to work with (2 vs. 10) than the high-efficacy numbers (2 vs. 65), and therefore the drug efficacy manipulation may have been confounded with ease of calculation. This may have eliminated any advantage the high drug efficacy participants had in this case over the control condition participants. In other words, providing easy information seems to inform participants relative to no information, but providing difficult information does not necessarily have the same effect.

We also hypothesized that participants in the high-efficacy conditions and those in the low-efficacy conditions would not differ from each other in benefit accuracy. Instead, we found that benefit accuracy was often higher in the low-efficacy condition compared with the high-efficacy condition. This likely relates to the confound mentioned above. Although calculations were not required in all cases, it is possible that the number 10 (which is a relatively small number and, obviously, a multiple of 10), is easier to remember than 65 (which is a larger number with no prior meaning attached to it).

Was numeracy a moderator? (Drug Efficacy Hypothesis E4)

Statistical Print Conditions. Objective and subjective numeracy did not moderate the relation between drug efficacy level and behavioral intentions, attitude toward the drug, or perceived benefit, *p*s > 0.05.

Although objective and subjective numeracy appear to be direct predictors of reporting quantitative benefit information accurately (such that participants with greater numeracy skills were more accurate), numeracy was not a consistent moderator of the relation between drug efficacy level and benefit accuracy. There was a significant interaction between objective numeracy and drug efficacy level (low. vs. high vs. control) when predicting the relative frequency accuracy measure, $p = 0.04$. An examination of the unweighted means suggests that

participants higher in objective numeracy, compared with those with lower objective numeracy, were more likely to be accurate in the low-efficacy and high-efficacy conditions, but not in the control condition. When the control condition (which had only 13 correct responses) was not included in the model, this interaction was no longer significant, $p > 0.05$. There was also a significant interaction between objective numeracy and drug efficacy level (low vs. high) when predicting the placebo accuracy measure, $p = 0.04$. An examination of the unweighted means suggests that participants with higher objective numeracy were more likely to be accurate in the low drug efficacy, compared to high drug efficacy conditions; whereas participants with lower objective numeracy were less likely to show this effect. No other interactions were significant, *p*s > 0.05.

Statistical Television Conditions. Objective numeracy moderated the relation between drug efficacy level and perceived efficacy, $p = 0.01$. An examination of the unweighted means suggests that participants with higher objective numeracy, compared with lower objective numeracy, showed a larger difference in perceived benefit between the high- and low-efficacy conditions (that is, they were more likely to rate the drug in the high drug efficacy conditions as having higher efficacy than the drug in the low drug efficacy condition). Subjective numeracy showed the same effect, $p = 0.02$. In addition, subjective numeracy moderated the relation between drug efficacy level and attitude toward the drug. An examination of the unweighted means suggests that participants with higher subjective numeracy, compared with lower subjective numeracy, showed a larger difference in attitude toward the drug between the high- and low-efficacy conditions (that is, they were more likely to report a positive attitude toward the drug in the high-efficacy conditions than toward the drug in the low-efficacy condition), $p = 0.02$. Objective numeracy did not moderate the relation between drug efficacy level and attitude

toward the drug, and neither objective nor subjective numeracy moderated the relation between drug efficacy level and behavioral intentions, *p*s > 0.05.

Although objective and subjective numeracy appear to be direct predictors of reporting quantitative benefit information accurately (such that participants with greater numeracy skills were more accurate), numeracy was not a consistent moderator of the relation between drug efficacy level and benefit accuracy. There was a significant interaction between objective numeracy and drug efficacy level (low vs. high vs. control), $p < 0.001$, and between subjective numeracy and drug efficacy level (low vs. high vs. control), $p < 0.001$, when predicting the placebo accuracy measure, but these effects were no longer significant when the control condition was removed from the model, *p*s > 0.05. Because there was only one correct response in the control condition, this finding is likely an artifact. No other interactions were significant, *p*s > 0.05.

Visual Print Conditions. Objective and subjective numeracy did not moderate the relation between drug efficacy level and behavioral intentions, attitude toward the drug, and perceived benefit, *p*s > 0.05.

Although objective and subjective numeracy appear to be direct predictors of reporting quantitative benefit information accurately (such that participants with greater numeracy skills were more accurate), numeracy was not a consistent moderator of the relation between drug efficacy level and benefit accuracy. There was a significant interaction between objective numeracy and drug efficacy level (low vs. high vs. control) when predicting the relative frequency accuracy measure, $p = 0.01$. An examination of the unweighted means suggests that participants higher in objective numeracy, compared with those with lower objective numeracy, were more likely to be accurate in the low-efficacy and high-efficacy conditions but not in the

control condition. When the control condition (which had only 13 correct responses) was not included in the model, this interaction was no longer significant, $p > 0.05$. There was also a significant interaction between objective numeracy and drug efficacy level (low vs. high) when predicting the placebo accuracy measure, $p = 0.04$. An examination of the unweighted means suggests that participants with higher objective numeracy were more likely to be accurate in the low-efficacy, compared with high-efficacy conditions; whereas participants with lower objective numeracy were less likely to show this effect. No other interactions were significant (*p*s > 0.05).

Visual Television Conditions. Objective numeracy moderated the relation between drug efficacy level (low vs. high) and perceived benefit, $p < 0.001$. An examination of the unweighted means suggests that participants with higher objective numeracy, compared with lower objective numeracy, showed a larger difference in perceived benefit between the high- and low drug efficacy conditions (that is, they were more likely to rate the drug in the high-efficacy conditions as having greater benefits than the drug in the low-efficacy condition). Subjective numeracy showed the same effect, $p < 0.001$. In addition, objective numeracy moderated the relation between drug efficacy level (low vs. high) and attitude toward the drug, $p = 0.001$. An examination of the unweighted means suggests that participants with higher subjective numeracy, compared with lower subjective numeracy, showed a larger difference in attitude toward the drug between the high and low drug efficacy conditions (that is, they were more likely to report a positive attitude toward the drug in the high-efficacy conditions than toward the drug in the low-efficacy condition). Finally, objective numeracy moderated the relation between drug efficacy level (low vs. high) and intention to search for more information, $p = 0.01$, intention to request the drug, $p = 0.004$, and intention to take the drug if prescribed, $p = 0.001$. Subjective numeracy moderated the relation between drug efficacy level (low vs. high) and

intention to take the drug if prescribed, $p < 0.001$. An examination of the unweighted means suggests that these interactions were such that participants with higher numeracy, compared with lower subjective numeracy, showed a larger difference in intentions between the high- and low drug efficacy conditions (that is, they were more likely to report greater intentions in the high-efficacy condition than in the low-efficacy condition). Subjective numeracy did not moderate the relation between drug efficacy level and attitude toward the drug, intention to search for more information, or intention to request the drug, *p*s > 0.05.

Although objective and subjective numeracy appear to be direct predictors of reporting quantitative benefit information accurately (such that participants with greater numeracy skills were more accurate), numeracy was not a consistent moderator of the relation between drug efficacy level and benefit accuracy. There was a significant interaction between objective numeracy and drug efficacy level (low vs. high) when predicting the percent accuracy measure, $p = 0.04$. Subjective numeracy showed the same effect, $p = 0.02$. An examination of the unweighted means suggests that these interactions were such that participants with lower numeracy, compared with higher subjective numeracy, showed a larger difference on the percent accuracy measure between the high- and low drug efficacy conditions. There was a significant interaction between objective numeracy and drug efficacy level (low vs. high vs. control), $p < 0.001$, and between subjective numeracy and drug efficacy level (low vs. high vs. control), $p < 0.001$, when predicting the placebo accuracy measure, but these effects were no longer significant when the control condition was removed from the model, *p*s > 0.05. Because there was only one correct response in the control condition, this finding is likely an artifact. No other interactions were significant, *p*s > 0.05.

Discussion. We examined whether numeracy moderated the effects of drug efficacy level on a number of specific outcomes. In the print ad conditions, we found that numeracy did not moderate the relation between drug efficacy level and behavioral intentions, attitude toward the drug, or perceived benefit. There were no main effects of drug efficacy level on these measures in the statistical print conditions, so it is not surprising there was no moderation. In the visual print condition, there were main effects of drug efficacy level on perceived benefit and attitude toward the drug, but numeracy did not moderate these effects. In the statistical television condition, the main effects of drug efficacy level on perceived benefit and attitude toward the drug were moderated by numeracy, although the main effects for behavioral intentions were not. In the visual television condition, the main effects of drug efficacy level on perceived benefit, attitude toward the drug, and behavioral intentions were moderated by numeracy. In both television conditions these effects showed that participants with higher numeracy, compared with lower numeracy, were more likely to understand that the high-efficacy drug was better than the low-efficacy drug.

Although there were a few significant interactions between drug efficacy level and numeracy when predicting accuracy measures, these interactions did not follow any particular pattern. Several of these seemed to be driven by small numbers of participants in a condition. However, numeracy had a consistent main effect on accuracy. Thus, it appears that having the ability to work with numbers improves accurate reporting of quantitative information, regardless of the drug efficacy level of the drug.

Did drug efficacy affect risk recall and perceived risk? (Drug Efficacy Hypotheses E5 and E6)

Statistical Print Conditions. Aided and unaided risk recall and perceived risk did not significantly differ by drug efficacy level, *p*s > 0.05 (see Table R, Appendix C, for means and percents).

Statistical Television Conditions. Aided and unaided risk recall did not significantly differ by drug efficacy level, *p*s > 0.05 (see Table R for means and percents and Table S for test statistics in Appendix C). Participants in the high-efficacy condition had lower perceived risk than did participants in the low-efficacy condition, $p = 0.05$.

Following Baron and Kenny (1986), we examined four criteria to test whether perceived benefit mediated the relation between drug efficacy level and perceived risk. First, the direct link from drug efficacy level (low vs. high) to perceived risk was significant, $\beta = -0.19$, $p = 0.046$. As noted above, participants in the high-efficacy condition had lower perceived risk than did participants in the low-efficacy condition. Second, the link from drug efficacy level to perceived benefit (i.e., the mediator) was $\beta = 0.22$ $p = 0.05$. As noted above, participants in the high-efficacy condition had higher perceived benefit than did participants in the low-efficacy condition. Third, the link between the mediator and perceived risk was significant when controlling for drug efficacy level, $\beta = -0.45$, $p < 0.001$, such that higher perceived benefit was associated with lower perceived risk. Finally, when the mediator was included in the equation, the direct link from drug efficacy level (low vs. high) to perceived risk became non-significant, $\beta = -0.09$, $p = 0.28$. In addition, the Sobel test, another measure of mediation, was significant (2.30, $p = 0.02$).

Visual Print Conditions. Aided and unaided risk recall did not significantly differ by drug efficacy level, *p*s > 0.05 (see Table T for means and percents and Table U for test statistics

in Appendix C). Participants in the high-efficacy condition had lower perceived risk than did participants in the low-efficacy condition, $p = 0.01$.

Following Baron and Kenny (1986), we examined four criteria to test whether perceived benefit mediated the relation between drug efficacy level and perceived risk. First, the direct link from drug efficacy level (low vs. high) to perceived risk was significant, $\beta = -0.26$, $p = 0.01$. As noted above, participants in the high-efficacy condition had lower perceived risk than did participants in the low-efficacy condition. Second, the link from drug efficacy level to perceived benefit (i.e., the mediator) was significant, $\beta = 0.52$ $p < 0.001$. As noted above, participants in the high-efficacy condition had higher perceived benefit than did participants in the low-efficacy condition. Third, the link between the mediator and perceived risk was significant when controlling for drug efficacy level, $\beta = -0.48$, $p < 0.001$, such that higher perceived benefit was associated with lower perceived risk. Finally, when the mediator was included in the equation, the direct link from drug efficacy level (low vs. high) to perceived risk became non-significant, $\beta = -0.01$, $p = 0.89$. In addition, the Sobel test was significant, 5.22, $p < 0.001$.

Visual Television Conditions. Aided and unaided risk recall did not significantly differ by drug efficacy level, *p*s > 0.05 (see Table T for means and percents and Table U for test statistics in Appendix C). Participants in the high-efficacy condition had lower perceived risk than did participants in the low-efficacy condition, $p < 0.001$.

Following Baron and Kenny (1986), we examined four criteria to test whether perceived benefit mediated the relation between drug efficacy level and perceived risk. First, the direct link from drug efficacy level (low vs. high) to perceived risk was significant, $\beta = -0.41$, $p < 0.001$. As noted above, participants in the high-efficacy condition had lower perceived risk than did participants in the low-efficacy condition. Second, the link from drug efficacy level to

perceived benefit (i.e., the mediator) was significant, $\beta = 0.73$, $p < 0.001$. As noted above, participants in the high-efficacy condition had higher perceived benefit than did participants in the low-efficacy condition. Third, the link between the mediator and perceived risk was significant when controlling for drug efficacy level, $\beta = -0.42$, $p < 0.001$, such that higher perceived benefit was associated with lower perceived risk. Finally, when the mediator was included in the equation, the direct link from drug efficacy level (low vs. high) to perceived risk became non-significant, $\beta = -0.10$, $p = 0.25$. In addition, the Sobel test was significant, 6.85, $p < 0.001$.

Discussion. We were interested in whether adding quantitative benefit information to an ad would influence the recall and perceptions of the risk information in that ad. We did not expect the addition of quantitative benefit information to affect risk recall, and this was supported in all conditions. Thus, there is no evidence that including quantitative benefit information detracts from peoples' understanding of the risk information.

Risk recall is one way to look at the influence of quantitative benefit information on the understanding of risk; another is to look at risk perceptions. We hypothesized that perceived risk would be affected by drug efficacy level such that perceived risk would be lower in the high-efficacy condition compared with the low-efficacy condition, based on the affect heuristic hypothesis (Finucane, Alhakami, et al. 2000; Slovic and Peters, 2006). According to this hypothesis, people cannot avoid using affect in making decisions, resulting in perceptions that as benefits increase, risks decrease. Finucane et al. demonstrated this heuristic as it applies to perceptions of nuclear power plants. The pattern of results in the statistical television, visual print, and visual television conditions demonstrated that this hypothesis applies to the perceptions of prescription drugs as well. Again, the results in the statistical print conditions

may have been weaker because the text alone was not sufficient to attract attention (Smerecnik et al. 2010).

General Discussion

The purpose of this study was to investigate the role of quantitative information about prescription drug benefits in DTC print and television ads. Specifically, we looked at whether the addition of quantitative benefit information, in different statistical formats or with the addition of a visual aid, influenced participants' perceptions and behavioral intentions and their ability to accurately recall the drug's quantitative benefit information and risk information. We also examined which statistical and visual formats were the most helpful to people.

We conducted an experimental study via the Internet that consisted of 38 study arms, varying ad type, drug efficacy level, statistical format, and visual format between participants. A total of 4,805 men and women from across the country who had been diagnosed with high cholesterol responded to the survey. Participants in all parts of this study were introduced to a completely novel fictitious drug (Votrea) and were then asked questions about their reactions to it in one 20-minute session.

A major contribution of this research is that, to our knowledge, it is the first study to systematically examine the addition of quantitative information in television DTC ads. In fact, to our knowledge, the risk communication literature has focused only on print (or online text) modalities, making this the first study to examine the addition of quantitative information in a dynamic, television modality. It is important to study the addition of quantitative information in multiple modalities before applying the risk communication literature to DTC advertising more generally. Through the descriptive comparisons of our print and television designs, we were able to assess whether our effects were consistent or divergent, depending on whether the information

was read or viewed. In general, we found that our results were similar but stronger for participants who viewed the television ad. It is possible that the television ad captured participants' attention in ways that the print ad was not capable of because of the addition of peripheral factors such as music, dialog, and movement. It is also possible that the presence of two kinds of communication, audio and visual, created additional memory traces that allowed for better recall and processing of the television ad. However, it is important to note that the television and print ad conditions were not designed for comparison with one another and some differences existed in the administration of these two conditions. First, there were slight differences in the print and television samples (i.e., participants without broadband Internet access were excluded from television conditions due to loading time difficulties). Second, participants in the print conditions controlled how long they viewed the ad: they could examine the print ads for as little or as much time as they wanted. Participants in the television conditions, however, could only progress to the questionnaire when they had viewed the ads twice. Nonetheless, the information gleaned from the television conditions themselves provides novel data for the literature.

In general, we found that providing quantitative benefit information in DTC ads increases participants' ability to accurately report the benefits of the drug in quantitative terms. Moreover, there were specific differences in accurate recall of quantitative benefit information, depending on the way the information was presented, either statistically or visually. Whereas some combinations of absolute frequency and percent seemed to provide useful information to people, relative frequency was a difficult concept, particularly because it required calculation or transformation in many cases. Bar charts and tables seemed to provide more clarity for

individuals than pie charts and pictographs, at least when it came to the accuracy of benefit recall.

Overall, our results showed that benefit recall was low, regardless of the particular presentation of information. This is likely an effect of our procedure, in which participants were not able to refer back to the print ad or television ad as they were answering the questions. We were interested in recall as opposed to the ability to search and recover information, particularly because this was the first study to examine these issues in the context of a television ad. Other studies have examined such search and recover behaviors, such as comprehension and application of information in print formats (e.g., Woloshin and Schwartz, 2011). Even when participants could examine the document as long as they liked and while answering questions, approximately one third of the sample responded incorrectly, demonstrating that information is generally difficult to glean from ads. It is through the collection of various studies that we can arrive at an understanding of exactly what each presentation format can do to minimize misunderstanding of such information and in what situations. Future research with DTC ads that use moving images should address issues of comprehension and application.

Despite the benefit accuracy findings, we did not find differences to support our hypotheses regarding perceived benefits, attitudes toward the drug, or behavioral intentions. In teasing apart the effects of statistical format on the recall and processing of benefit information, one question we were most interested in addressing was the theoretical basis for the pattern of findings we predicted. Our first hypotheses explored the influence of statistical format on perceived benefit, attitude toward the drug, and behavioral intentions such as seeking more information about the drug. Based on previous research, we pitted two competing hypotheses against each other:

(1a) **Overestimation hypothesis**: The first hypothesis rested on the assumption that in the absence of any quantitative information participants overestimate the effectiveness of drugs (Schwartz et al. 2009). Accordingly, we predicted that behavioral intentions, attitude toward drug, and perceived benefit would be **higher** for participants in the *no* statistical format condition, compared to *all other* statistical format conditions. Support for this interpretation would be found if estimates of the benefits are higher in the *no* statistical format condition than in *all other* statistical format conditions.

(1b). **Peripheral cue hypothesis**: The competing hypothesis rested on the assumption that any statistical information will be used as a peripheral cue (e.g., Petty and Cacioppo, 1986); that is, participants will not process the details of the quantitative information provided in the various statistical formats but will rather view any quantitative information as scientific proof of the drug's efficacy. Accordingly, we predicted that behavioral intentions, attitude toward the drug, and perceived benefit would be **lower** for participants in the *no* statistical format condition, compared to *all other* statistical format conditions. Support for this interpretation would be found if, in addition to perceived benefit effects, estimates on attitude toward the ad peripheral cue measures—ratings of how believable, persuasive, informative, etc, the ad is—are lower in the *no* statistical format condition than in *all other* statistical format conditions.

We did not find support for either hypothesis. There were no significant differences between the statistical format conditions or between those who saw the numerical information and those in the control condition on any of these three outcome measures, for print or television ads. However, we did find differences depending on whether participants saw a high- or low-efficacy drug or did not receive any benefit information at all. Contrary to the overestimation hypothesis, we found that there were no differences between participants in the low-efficacy

conditions and those who did not see any information, whereas participants in the high-efficacy condition viewed the drug as significantly more effective than those who did not see any information. If participants had been overestimating the benefits of the drug in the absence of information, the opposite effect would have occurred—their ratings would have been indistinguishable from the high-efficacy participants. Thus, it seems that in the absence of information, participants did not overestimate the benefits of the drug. It should be noted, however, that the raw responses to the accuracy measures (that is, the numbers they guessed rather than whether they were correct; data not shown) of participants who did not see any information were closer to those of the high efficacy condition participants for some formats (i.e., absolute frequency and percent). The raw responses were also closer to those of the low efficacy condition participants for other formats (i.e., relative frequency).

We also suspected that participants might use numerical information as a peripheral cue or signal, somewhat akin to the idea that "science is credible and therefore believable" (e.g., Eagly and Chaiken, 1993; Petty and Cacioppo, 1986). Past research suggests that when faced with a problem, people may employ either effortful processing or engage in heuristics or mental shortcuts to guide their judgments, thereby avoiding the expenditure of cognitive effort (De Neys 2006). We did not find support for the idea that participants used quantitative information as a heuristic. There were no significant differences in perceived benefit, attitude toward the drug, or behavioral intentions between participants who saw statistical benefit information and those who did not see any information. This was not a simple failure of the manipulations because, as stated earlier, participants were able to accurately distinguish between levels of benefit as presented in the ad. Thus, we find no evidence that participants used the presence of quantitative benefit information as a heuristic to indicate the (high) efficacy of the product. Instead,

participants processed the information and recalled it differently, depending on which statistical format they saw.

We found similar results in the visual format designs. Benefit accuracy varied by visual format, but our data did not suggest that visual formats influence perceived risk, attitude toward the drug, or behavioral intentions. There are a number of possible reasons that we found significant benefit accuracy differences but did not find significant differences for behavioral intentions, attitude toward the drug, and perceived benefit. First, it could be a measurement issue; our measures might not have been sensitive enough to pick up differences. This explanation is unlikely, however, since, as in the statistical format design, we did see significant drug efficacy level differences on the same outcome measures. Moreover, if the visual aids themselves were too similar, we still may have seen a difference between those who saw any visual aid and those in the control group.

Second, it is possible that an artifact of the study protocol may have influenced our results. Although our findings demonstrated that the accuracy of recall of benefit information does vary depending on how it is presented, our findings do not demonstrate that this benefit accuracy necessarily translates into further processing. We saw that visual aids do have some effect on perceived benefit and behavioral intentions, but variations in statistical format did not have the same effect. This is likely due to the one-time administration of the survey. Because participants in all parts of this study were introduced to a completely novel fictitious drug (Votrea) and were then asked questions about their reactions to it in one 20-minute session, they may not have had a chance to develop strong beliefs about the drug. Attitudes, in particular, are generally long-standing beliefs about particular objects (e.g., Fazio 1986). It is possible that with repeated exposure to information about Votrea, as would occur over time during an advertising

campaign for a prescription drug, people may develop perceptions and attitudes about the drug and make decisions about specific behaviors they intend to pursue. Therefore, the lack of findings on the non-accuracy variables does not indicate that advertising campaigns do not have effects on these outcomes, merely that our particular procedure may have failed to tap into these constructs. In a more naturalistic setting, where participants are exposed to ads repeatedly over time, they may develop: attitudes toward the drug, have stronger differences in perceptions, and may show greater differences in intention to take further actions.

Another explanation that draws support from current psychological theory is that people do not use specific quantitative benefit information when making decisions. Recent research on fuzzy trace theory has demonstrated evidence for two distinct types of memory: *verbatim* memory, which records the details of an event or activity and fades within a few days, and *gist* memory, which involves an interpretation of the event or activity and is much longer lasting (e.g., Brainerd, Reyna, et al. 2008). Gist memory evolves when the actual event is compared with existing information in the brain and modeled to fit within this existing framework (e.g., Reyna and Brainerd, 1995). Given the limitations of our study paradigm—that is, showing participants an ad for a novel drug and soliciting responses to it in one 20-minute sitting—it seems likely that we tapped the verbatim memory in our questions but were unable to assess gist memory, which may not have formed yet. Reyna (2008) has shown that when making medical decisions, accuracy may not be the best predictor of good decisions. Instead, gist memory appears to predict sensible decision-making. Thus, our findings of greater benefit accuracy are a start and suggest that participants are reading and or hearing the information, but the information does not seem to have translated into longer lasting perceptions or attitudes that would result in changed behavior. Future research on the communication of prescription drug information in

DTC ads should work to further analyze the occurrences when verbatim memory and when gist memory come into play.

Other factors may also influence the role of drug efficacy level, statistical format, and visual format on recall, perception, attitudes, and intentions. Any time that quantitative information is presented, the numeracy level of individuals must be addressed, and we did so in the current study. Numeracy is the degree to which a person can work with numbers, including basic probability and mathematical concepts (Lipkus, Samsa, et al. 2001). Like literacy with prose, each person has a certain level of comfort with the presentation of numbers. Past research has shown that even highly educated individuals in the United States generally have great difficulty with numerical concepts (Lipkus et al. 2001). Many medical concepts contain numerical information and a deeper understanding of prescription drugs in particular requires a knowledge of how numbers work. Numeracy, therefore, is an important factor to measure and to correlate with the understanding of quantitative information.

We did indeed find that numeracy was relevant to the recall and perceptions of the information in the ads. However, the moderating effect that we hypothesized was not supported. In other words, presenting the information in different formats did not differentially help those with low numeracy versus those with high numeracy. A pictograph was no more helpful for low numerate individuals than it was for high numerate individuals, for example. Instead, numeracy had direct effects across the board. Unsurprisingly, participants who had good numerical skills were more likely to accurately report the quantitative benefits of the drug than were participants who had fewer numerical skills. Our finding that participants did not process different formats differently depending on their numeracy skills suggests that numeracy levels may have such a

large influence on the understanding of information that they cancel out any nuances between visual formats.

Limitations and Future Directions

As with any research study, the present study had several limitations. Although we recruited a large number of individuals who suffer from high cholesterol, the study was conducted via the Internet and thus is subject to many of the limitations inherent in this mode of administration, such as the lack of control over testing conditions (e.g., Heerwegh and Loosveldt, 2008; Klovning, Sandvik, et al. 2009). However, the Internet panel from which we drew our participants was not a typical online opt-in panel, but rather a panel specifically designed to be nationally representative. Therefore the sample suffers from fewer biases than would a typical opt-in panel (Chang and Krosnick, 2009; Dennis 2010).

One limitation of the administration mode was that only those individuals with broadband access could be selected for random assignment into the television ad conditions. Participants who had dial-up capabilities only were limited to one of the print ad conditions. This limitation was a logistical compromise after determining that dial-up methods could not support the reliable download of the television ad that was required to complete the questionnaire. Thus, our sample in the television ad conditions of the study may have skewed toward higher socioeconomic groups. Despite this limitation, conducting the study via the Internet enabled us to maximize the power needed to detect differences and allowed for greater geographical distribution. Future research could be conducted in-person to confirm the results with participants who do not have broadband access.

Our findings provide many rich areas for continued investigation. At the very least, future research should include measures of gist processing to assess the role of verbatim versus

gist processing. Recall and comprehension are key outcome measures, but future research should go beyond these measures to tap effects on informed decision-making. Research has been conducted in this vein (e.g., Woloshin and Schwartz, 2009), but no previous research on television DTC or the display page of print DTC advertising has combined a study of quantitative information with informed decision-making outcome measures. The addition of gist measures would complement and advance the field.

Finally, this study focused on one medical condition (high cholesterol) and one ad execution in print and television for a fictitious prescription drug. It is possible that different effects would be found with a different medical condition or with different ads. To ensure generalizability, these results should be confirmed in another medical condition with various ad executions.

Our findings demonstrated that participants can accurately recall quantitative benefit information from DTC prescription drug print and television ads. Overall, presenting information using absolute frequency and percent formats appears to produce the most successful recall. Presenting a visual aid also appears to assist in benefit accuracy, with bar charts and tables demonstrating advantages over other visual formats. In general, providing information to participants allows them to see the information and answer questions about it correctly, although it does not necessarily change their perceptions, attitudes, and behavioral intentions. At the same time, including quantitative benefit information did not have a detrimental effect on the recall of risk information. Thus, presented in nonmisleading ways, the inclusion of quantitative benefit information in DTC print and television ads may aid consumers.

References

Ancker, JS, Y Senathirajah, R Kukafka, and JB Starren, 2006, Design Features of Graphs in Health Risk Communication: A Systematic Review, Journal of the American Medical Information Association, 13:608-618.

Baron, RM, and DA Kenny, 1986, The Moderator-Mediator Distinction in Social Psychological Research: Conceptual, Strategic, and Statistical Considerations, Journal of Personality and Social Psychology, 51:1173-1182.

Beyth-Marom, R, 1982, How Probable is Probable? A Numerical Translation of Verbal Probability Expressions, Journal of Forecasting, 1:257-269.

Bowman, ML, 2002, The Perfidity of Percentiles, Archives of Clinical Neuropsychology, 17:295-303.

Brainerd, CJ, VF Reyna, SJ Ceci, 2008, Developmental Reversals in False Memory: A Review of Data and Theory, Psychological Bulletin, 134:343-382.

Chandler, P and J Sweller, 1991, Cognitive Load Theory and the Format of Instruction, Cognition and Instruction, 8:293-332.

Chang, L., and JA Krosnick, 2009, National Surveys Via RDD Telephone Interviews Versus the Internet: Comparing Sample Representativeness and Response Quality, Public Opinion Quarterly, 73:641-678.

Cohen, DJ, JM Ferrell, and N Johnson, 2002, What Very Small Numbers Mean, Journal of Experimental Psychology: General, 131:424-442.

De Neys, W, 2006, Dual Processing in Reasoning: Two Systems But One Reasoner, Psychological Science, 17:428-433.

Dennis, M, 2010, KnowledgePanel: Processes and Procedures Contributing to Sample

Representativeness and Tests for Self-Selection Bias, http://www.knowledgenetworks.com/ganp/docs/KnowledgePanelR-Statistical-Methods-Note.pdf, last accessed March 9, 2012.

Eagly, AH and S Chaiken, 1993, Process Theories of Attitude Formation and Change: The Elaboration Likelihood and Heuristic-Systematic Models. In: AH Eagly and S Chaiken, eds. The Psychology of Attitudes, Orlando: Harcourt Brace, 303-350.

Fagerlin, A, PA Ubel, DM Smith, and BJ Zikmund-Fisher, 2007, Making Numbers Matter: Present and Future Research in Risk Communication, American Journal of Health Behavior, 31, Suppl 1: S47-56.

Fagerlin, A, C Wang, and PA Ubel, 2005, Reducing the Influence of Anecdotal Reasoning on People's Health Care Decisions: Is a Picture Worth a Thousand Statistics? Medical Decision Making, 25:398-405.

Fagerlin, A, BJ Zikmund-Fisher, PA Ubel, A Jankovic, HA Derry, and DM Smith, 2007, Measuring Numeracy Without a Math Test: Development of the Subjective Numeracy Scale (SNS), Medical Decision Making, 27:672-680.

Fazio, RH, 1986, How Do Attitudes Guide Behavior? In: RM Sorrentino and ET Higgins, eds. The Handbook of Motivation and Cognition: Foundations of Social Behavior, New York: Guilford Press, 204-243.

Finucane, ML, A Alhakami, P Slovic, and SM Johnson, 2000, The Affect Heuristic in Judgments of Risks and Benefits, Journal of Behavioral Decision Making, 13:1-17.

Frosch, DL, PM Krueger, RC Hornik, PF Cronholm, and FK Barg, 2007, Creating Demand for Prescription Drugs: A Content Analysis of Television Direct-to-Consumer Advertising, Annals of Family Medicine, 5(1):6-13.

Hess, R, CHM Visschers, and M Siegrist, 2011, Risk Communication With Pictographs: The Role of Numeracy and Graph Processing, Judgment and Decision Making, 6:263-274.

Houts, PS, JT Witmer, HE Egeth, MJ Loscalzo, and JR Zabora, 2001, Using Pictographs to Enhance Recall of Spoken Medical Instructions, Patient Education and Counseling, 43:231-242.

Heerwegh, D, and G Loosveldt, 2008, Face-to-Face Versus Web Surveying in a High-Internet-Coverage Population: Differences in Response Quality, Public Opinion Quarterly, 72:836-846.

Klovning, A, H Sandvik, and S Hunskaar, 2009, Web-Based Survey Attracted Age-Biased Sample With More Severe Illness Than Paper-Based Survey, Journal of Clinical Epidemiology, 62:1068-1074.

Lang, A, 1995, Defining Audio/Video Redundancy From a Limited-Capacity Information Processing Perspective, Communication Research, 22(1):86-115.

Lipkus, I, 2007, Numeric, Verbal, and Visual Formats of Conveying Health Risks: Suggested Best Practices and Future Recommendations, Medical Decision Making, 27:697-713.

Lipkus, I, and JG Hollands, 1999, The Visual Communication of Risk, Journal of the National Cancer Institute Monographs, 25:149-163.

Lipkus, IM, G Samsa, and BK Rimer, 2001, General Performance on a Numeracy Scale Among Highly Educated Samples, Medical Decision Making, 21:37-44.

MacKinnon, DP, G Warsi, and JH Dwyer, 1995, A Simulation Study of Mediated Effect Measures, Multivariate Behavioral Research, 30(1):41-62.

Paivio, A, 1983, The Empirical Case for Dual Coding. In: Yuille, J, ed. Imagery, Memory, and Cognition: Essays in Honor of Allan Paivio, Hillside, NJ: Lawrence Erlbaum Associates, 307-332.

Peters, E, D Vastfjall, P Slovic, CK Mertz, K Massocco, and S Dickert, 2006, Numeracy and Decision Making, Psychological Science, 17:407-13.

Petty, RE, and JT Cacioppo, 1986, Communication and Persuasion: Central and Peripheral Routes to Attitude Change, New York: Springer-Verlag.

Reyna, VF, 2008, A Theory of Medical Decision Making and Health: Fuzzy Trace Theory, Medical Decision Making, 28:850-865.

Reyna, VF and CJ Brainerd, 1995, Fuzzy-Trace Theory: An Interim Synthesis, Learning and Individual Differences, 7:1-75.

Schwartz, L, S Woloshin, W Black, and HG Welch, 1997, The Role of Numeracy in Understanding the Benefit of Screening Mammography, Annals of Internal Medicine, 127(11):966-72.

Schwartz, LM, S Woloshin, and HG Welch, 2007, The Drug Facts Box: Providing Consumers With Simple Tabular Data on Drug Benefit and Harm, Medical Decision Making. 27:655-692.

Schwartz, LM, S Woloshin, and HG Welch, 2009, Communicating Drug Benefits and Harms With a Drug Facts Box: Two Randomized Trials, Annals of Internal Medicine, 150:516-527.

Slovic, P, J Monahan, and DG MacGregor, 2000, Violence Risk Assessment and Risk Communication: The Effects of Using Actual Cases, Providing Instruction, and Employing Probability Versus Frequency Formats, Law and Human Behavior, 24:271-96.

Slovic, P, and E Peters, 2006, Risk Perception and Affect, Current Directions in Psychological Science, 15:322-325.

Smerecnik, CMR, I Mesters, LTE Kessels, RAC Ruiter, NK De Vries, and H De Vries, 2010, Understanding the Positive Effects of Graphical Risk Information on Comprehension: Measuring Attention Directed to Written, Tabular, and Graphical Risk Information, Risk Analysis, 30:1287-1398.

Sweller, J, 1988, Cognitive Load During Problem Solving: Effects on Learning, Cognitive Science, 12:257-285.

Thompson, RF, 2005, In Search of Memory Traces, Annual Review of Psychology, 56:1-23.

Woloshin, S, and LM Schwartz, 2011, Communicating Data About the Benefits and Harms of Treatment: A Randomized Trial, Annals of Internal Medicine, 155:87-96.

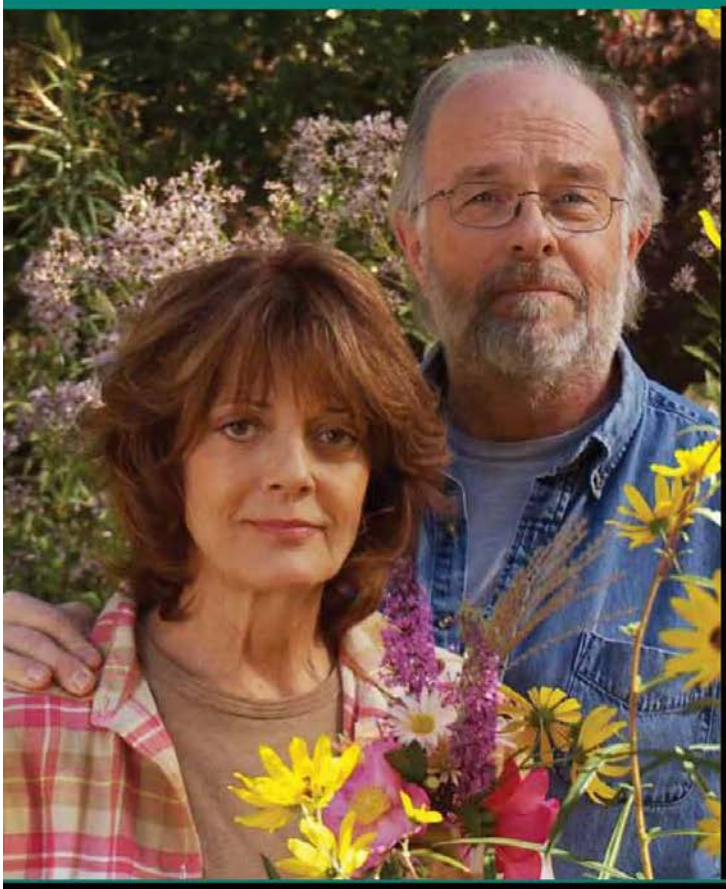
Woloshin, S, L Schwartz, J Tremmel, and HG Welch, 2001, Direct to Consumer Advertisements for Prescription Drugs: What Are Americans Being Told, Lancet, 358:1141-46.

Woloshin, S, LM Schwartz, and HG Welch, 2004, The Value of Benefit Data in Direct-to-Consumer Drug Ads, Health Affairs, Suppl Web Exclusives, W4:234-245.

Zikmund-Fisher, BJ, PA Ubel, DM Smith, HA Derry, JB McClure, A Stark, RK Pitch, and A Fagerlin, 2008, Communicating Side Effect Risks in a Tamoxifen Prophylaxis Decision Aid: The Debiasing Influence of Pictographs, Patient Education and Counseling, 73:209-214.

Appendix A

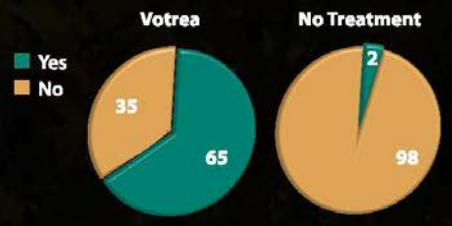
High cholesterol can be serious— it's a risk factor for heart disease.



Some people have increased risk factors for heart disease: high LDL (bad cholesterol), low HDL (good cholesterol), family history of heart disease, smoking, age, and obesity. Votrea is an enzyme blocker, also known as a "statin," which reduces the amount of bad cholesterol (LDL) your body makes normally each day. If used along with a proper diet and exercise program, Votrea may help lower your bad cholesterol.

Votrea reduces bad cholesterol for people with several common risk factors for heart disease (starting from an average bad cholesterol of 160 mg/dL). With Votrea, 65 out of 100 people lowered their bad cholesterol to normal levels versus 2 out of 100 people with no treatment.

Lowered Bad Cholesterol to Normal Levels



VOTREA
trevastatin calcium tablets
60 mg

Important risk information

Votrea is not for everyone, including people with liver problems and women who are nursing, pregnant, or may become pregnant. You need simple blood tests to check for liver problems. If you develop fever, unexplained weakness, or confusion, tell your doctor right away as these may be signs of a rare but potentially life threatening condition called TTP, which has been reported sometimes in less than 2 weeks after starting therapy. Also tell your doctor if you are taking other medications, or if you have any muscle pain or weakness, as this may be a sign of another rare but serious side effect. Common side effects include diarrhea, joint pain, and tiredness.

KAJ PHARMACEUTICALS
WARWICK, RI

You are encouraged to report negative side effects of prescription drugs to the FDA. Visit www.fda.gov/medwatch or call 1-800-FDA-1088.

Appendix B

Quantitative Study

Survey Questionnaire

INTRODUCTION

Thank you for agreeing to participate in this study today.

Make sure you are comfortable and can read the screen from where you sit. This study is about advertising for a new product. We will show you an ad for a new product, then ask you some questions about it. Your answers are private, and no one will ever connect your name with your answers. Your input is extremely valuable.

INSTRUCTIONS

Even though it is on a computer screen, please read this ad as if it were in a magazine. Imagine that this ad is for a product that you might be interested in for yourself. You can take as much time as you want to review it. We will ask you some questions about the ad later in the study.

The ad has two pages. You can enlarge and flip between the pages by clicking on the "Next" and "Back" buttons at the bottom of each page. Once you finish reading, please click "next" to move to the next part of the study.

INSTRUCTIONS

Even though it is on a computer screen, please watch this ad as if you saw it on television. Imagine that this ad is for a product that you might be interested in for yourself. We will ask you some questions about the ad later in the study. Once you finish viewing, please click "next" to move to the next part of the study.

SURVEY INSTRUCTIONS

Now please answer the following questions based on the ads you saw.

Q1. Were you able to view the ad for Votrea?

Yes
No ➔ [Terminate]
Not sure ➔ [Terminate]

Q2. What type of product is Votrea?

Over the counter drug
Prescription drug
Herbal supplement
Lens cleaner

Q3. What condition does Votrea treat?

High blood pressure
High cholesterol
Migraine headaches
Seasonal allergies

Q4. What are the benefits of Votrea?

(open ended)

Q5. Based on the information in the ad, does Votrea work better than *no treatment*?

Yes
No
Not sure

Q6. Based on the information in the ad, how effective would Votrea be for you?

1	2	3	4	5	6	7
Not at all effective			Moderately effective			Very effective

Q7. Based on the information in the ad, how well would Votrea work for you?

1	2	3	4	5	6	7
Not at all well			Moderately well			Very well

Please answer the following specific questions based on what you learned from the Votrea ad.

Q8. If 100 people take Votrea, how many will lower their bad cholesterol to normal levels?

_____ people

Q9. What percentage (%) of people who take Votrea will lower their bad cholesterol to normal levels?

_____ percent

Q10. How many times more effective is Votrea than no treatment in lowering bad cholesterol? Enter a number to show how effective Votrea is compared to no treatment.

_____ times more effective

Q11. If 100 people receive *no treatment*, how many will lower their bad cholesterol to normal levels?

_____ people

Q12. This ad had a chart, graph, or figure showing how well Votrea works.

Yes
No
Not sure

Q13. Please rate how likely or not likely you are to do each of the following behaviors using the scale on this page.

	Not at all likely	Somewhat likely	Very likely	Extremely likely
a. Talk to your doctor about Votrea				
b. Look for more information about Votrea				
c. Take Votrea if prescribed				

Q14. Please check which of the following were mentioned in the ad as benefits of taking Votrea. Even if you think a statement is true, please select it only if it was mentioned in the ad.

	Yes	No
a. Votrea works with diet and exercise.	x	
b. Votrea can lower bad cholesterol to normal levels.	x	
c. Votrea works for people with several common risk factors for heart disease.	x	
d. Votrea takes the place of diet and exercise.		x
e. Votrea can reduce the risk of diabetes.		x
f. Votrea is the #1 prescribed medication for high cholesterol.		x
g. High cholesterol is a risk factor for heart disease.	x	
h. Votrea can raise good cholesterol to normal levels.		x

Q15. What are the risks of Votrea?

(open ended)

Q16. Based on the information in the ad, how safe would Votrea be for you?

1	2	3	4	5	6	7
Not at all Safe			Moderately safe		Very	safe

Q17. Based on the information in the ad, how risky would Votrea be for you?

1	2	3	4	5	6	7
Not at all risky			Moderately risky			Very risky

Q19. Compared to other prescription treatments you could take for high cholesterol, how well do you think Votrea works?

Much better
Somewhat better
About the same
Somewhat worse
Much worse

Q20. Compared to other prescription treatments you could take for high cholesterol, how safe or risky do you think Votrea is?

Much safer
Somewhat safer
About the same
Somewhat riskier
Much riskier

Q21. How good or bad do you feel about this product?

Very bad
Somewhat bad
Neither bad nor good
Somewhat good
Very good

Q18. Please check which of the following were mentioned in the ad as risks of taking Votrea. Even if you think a statement is true, please select it only if it was mentioned in the ad.

	Yes	No
a. TTP is a risk when you take Votrea	x	
b. You need blood tests when taking Votrea.	x	
c. People with liver problems should not take Votrea.	x	
d. Votrea may cause muscle pain or weakness.	x	
e. A side effect of Votrea is blurry vision.		x
f. A side effect of Votrea is tiredness.	x	
g. A side effect of Votrea is dizziness.		x
h. People with kidney problems should not take Votrea.		x
i. A side effect of Votrea is nausea.		x

Q23. Thinking overall about the risks and benefits, would you say Votrea has:

1	2	3	4	5	6	7
More risks than benefits			Equal risks and benefits			More benefits than risks

Q22. Please rate your agreement or disagreement with each of the following statements.

	Strongly agree	Somewhat agree	Neither agree nor disagree	Somewhat disagree	Strongly disagree
a. The ad clearly communicated the *benefits* of Votrea.					
b. The ad clearly communicated the *risks* of Votrea.					
c. The ad evenly balanced the *risks and benefits* of Votrea.					

Q25. This ad had information about the *risks and side effects* of Votrea. When answering this question, please think only about the *risks and side effects* mentioned in the ad.

To what extent do you agree or disagree that the *risks and side effects* were:

	Strongly agree	Somewhat agree	Neither agree nor disagree	Somewhat disagree	Strongly disagree
a. Noticeable					
b. Persuasive					
c. Believable					
d. Distracting					
e. Confusing					
f. Scary					
g. Attractive					
h. Informative					

Q24. This ad had information about *the benefits of Votrea versus no treatment*. When answering this question, please think <u>only</u> about the *benefits* mentioned in the ad.

To what extent do you agree or disagree that the *benefits* were:

	Strongly agree	Somewhat agree	Neither agree nor disagree	Somewhat disagree	Strongly disagree
a. Noticeable					
b. Persuasive					
c. Believable					
d. Distracting					
e. Confusing					
f. Scary					
g. Attractive					
h. Informative					

Q26. How likely is it that the benefits of the drug presented in this ad are true?

1	2	3	4	5	6	7	8	9
Not at all Likely								Extremely Likely

Q27. How skeptical are you about the truth of the benefits of the drug presented in this ad?

1	2	3	4	5	6	7	8	9
Not at all Skeptical								Extremely Skeptical

Q28. How often do you need to have someone help you when you read instructions, pamphlets, or other written material from your doctor or pharmacy?

Never
Rarely
Sometimes
Often
Always

Q29. Now here are some questions that require you to use numbers to solve the problem. Some are easy, and others are more difficult. Please don't use a calculator: We'd like you to answer on your own. Remember, almost everyone will have trouble with these questions, so don't be upset if some are difficult—just do your best!

a. What is the correct answer:

17 – 8 + 4 = ?

a. 11
b. 13
c. 21
d. 23
e. 29
f. Don't know

b. What is the correct answer:

100 x 15 x 20 = ?

a. 1,500
b. 15,000
c. 30,000
d. 150,000
e. 300,000
f. Don't know

c. Imagine that you flip a fair coin 1,000 times. What is your best guess about how many times the coin would come up heads in 1,000 flips?

___ times out of 1,000

d. In the BIG BUCKS LOTTERY, the chance of winning a $10 prize is 1%. What is your best guess about how many people would win a $10 prize if 1,000 people each buy a single ticket to BIG BUCKS LOTTERY?

_____ people

e. **In ACME PUBLISHING SWEEPSTAKES, the chance of winning a car is 1 in 1,000. What percent of tickets to ACME PUBLISHING SWEEPSTAKES will win a car?**

___ percent

Q30. For each of the following questions, please check the box that best reflects how good you are at doing the following things:

a. **How good are you at working with fractions?**

❑1	❑2	❑3	❑4	❑5	❑6
Not at all good					Extremely good

b. **How good are you at working with percentages?**

❑1	❑2	❑3	❑4	❑5	❑6
Not at all good					Extremely good

c. **How good are you at calculating a 15% tip?**

❑1	❑2	❑3	❑4	❑5	❑6
Not at all good					Extremely good

d. **How good are you at figuring out how much a shirt will cost if it is 25% off?**

❑1	❑2	❑3	❑4	❑5	❑6
Not at all good					Extremely good

Q31. For each of the following questions, please check the box that best reflects your answer:

a. **When reading the newspaper, how helpful do you find tables and graphs that are part of a story?**

❑1	❑2	❑3	❑4	❑5	❑6
Not at all helpful					Extremely helpful

b. When people tell you the chance of something happening, do you prefer that they use words ("it rarely happens") or numbers ("there is a 1% chance")?

❑1	❑2	❑3	❑4	❑5	❑6
Always Prefer Words					Always Prefer Numbers

c. When you hear a weather forecast, do you prefer predictions using percentages ("there will be a 20% chance of rain") or predictions using only words ("there is a small chance of rain today")?

❑1	❑2	❑3	❑4	❑5	❑6
Always Prefer Words					Always Prefer Percentages

d. How often do you find numerical information to be useful?

❑1	❑2	❑3	❑4	❑5	❑6
Never					Very Often

Q32. Are you currently taking a prescription medicine for high cholesterol?

Yes
No
Don't know or uncertain

Q33. How severe is your high cholesterol now? Would you describe it as:

Very mild
Mild
Moderate
Serious
Very serious

Q34. In general, how much do you feel you know about *high cholesterol*? Would you say you know:

Nothing at all
Only a slight amount
Some
A good bit

A lot

Q35. In general, how much do you feel you know about *treatments for high cholesterol*? Would you say you know:

Nothing at all
Only a slight amount
Some
A good bit
A lot

Q36. Roughly, what is your total cholesterol level? Your best guess is OK.

_____

Q37. Have you ever seen any advertising for Votrea before today?

Yes
No
Not sure

Q38. What did you use to complete today's survey?

Laptop computer / Netbook computer
Desktop computer
Handheld phone (iPhone, Blackberry, Droid)
Handheld device other than a phone (iPod Touch, Palm Pilot, iPad)
Other: _____

Q39. Did you receive any help on the questions that had numbers in them from anything or anyone (e.g., calculator, other person, website)? It's okay if you did—we just want an honest answer.

Yes (write source_____)
No

Q43. How many years of education have you had?

Completed grade school
Completed middle school
Completed high school
Some college
Associate's degree (2-year)
Bachelor's degree (4-year)
Some postgraduate work
Postgraduate degree (M.A., Ph.D., M.D., J.D., etc.)

The purpose of this study is to learn about reactions to medical advertising. In order to get a real-life reaction, we used a pretend product in this study. Votrea is not a real product and is not available for sale. Please see your healthcare professional for questions about high cholesterol.

You have been very helpful. Thank you very much for your participation!

Appendix C

Table	Title	Page
A	Demographic Characteristics of Participants	100
B	Statistical Print Conditions: Weighted Means (Standard Errors) and Percents of Dependent Variables, by Statistical Format	101
C	Statistical Television Conditions: Weighted Means (Standard Errors) and Percents of Dependent Variables, by Statistical Format	102
D	Statistical Print Conditions: Weighted Percents of Accuracy Variables, by Statistical Format	103
E	Statistical Print Conditions: Test Statistics (*t*-tests and *p*-values) from Analyses Comparing Accuracy Measures for Each Statistical Format to the Control Condition	103
F	Statistical Print Conditions: Test Statistics (*t*-tests and *p*-values) from Analyses Comparing Accuracy Measures Among Statistical Formats	104
G	Statistical Television Conditions: Weighted Percents of Accuracy Variables, by Statistical Format	105
H	Statistical Television Conditions: Test Statistics (*t*-tests and *p*-values) from Analyses Comparing Accuracy Measures for Each Statistical Format to the Control Condition	105
I	Statistical Television Conditions: Test Statistics (*t*-tests and *p*-values) from Analyses Comparing Accuracy Measures Among Statistical Formats	106
J	Visual Print Conditions: Weighted Means (Standard Errors) and Percents of Dependent Variables, by Visual Format	107
K	Visual Television Conditions: Weighted Means (Standard Errors) and Percents of Dependent Variables, by Visual Format	108
L	Visual Print Conditions: Weighted Percents of Accuracy Variables, by Visual Format	108
M	Visual Print Conditions: Test Statistics (*t*-tests and *p*-values) from Analyses Comparing Accuracy Measures for Each Visual Format to the No Visual Condition	109
N	Visual Print Conditions: Test Statistics (*t*-tests and *p*-values) from Analyses Comparing Accuracy Measures Among Visual Formats	109
O	Visual Television Conditions: Weighted Percents of Accuracy Variables, by Visual Format	110
P	Visual Television Conditions: Test Statistics (*t*-tests and *p*-values) from Analyses Comparing Accuracy Measures for Each Visual Format to the No Visual Condition	110
Q	Visual Television Conditions: Test Statistics (*t*-tests and *p*-values) from Analyses Comparing Accuracy Measures Among Visual Formats	111
R	Statistical Print and Television Conditions: Weighted Means (Standard Errors) and Percents of Dependent Variables, by Drug Efficacy Level	112
S	Statistical Print and Television Conditions: Test Statistics (*F*-tests and *p*-values) from Analyses Comparing Dependent Measures Among Drug Efficacy Levels	113

T	Visual Print and Television Conditions: Weighted Means (Standard Errors) and Percents of Dependent Variables, by Drug Efficacy Level	114
U	Visual Print and Television Conditions: Test Statistics (*F*-tests and *p*-values) from Analyses Comparing Dependent Measures Among Drug Efficacy Levels	115
V	Statistical Print and Television Conditions: Weighted Percents of Accuracy Variables, by Drug Efficacy Level	116
W	Statistical Print and Television Conditions: Test Statistics (*F*-tests, *t*-tests, and *p*-values) from Analyses Comparing Accuracy Measures Among Drug Efficacy Levels	117
X	Visual Print and Television Conditions: Weighted Percents of Accuracy Variables, by Drug Efficacy Level	118
Y	Visual Print and Television Conditions: Test Statistics (*F*-tests, *t*-tests, and *p*-values) from Analyses Comparing Accuracy Measures Among Drug Efficacy Levels	119

Table A. Demographic Characteristics of Participants

	n (%)		
	Statistical Format Design	Visual Format Design	Total
N	2,807	2,504	4,805
Sex			
Male	1,358 (48.4%)	1,202 (48.0%)	2,318 (48.2%)
Female	1,449 (51.6%)	1,302 (52.0%)	2,487 (51.8%)
Age			
18–24	3 (0.2%)	6 (0.2%)	9 (0.2%)
25–34	32 (1.1%)	45 (1.8%)	70 (1.5%)
35–44	187 (6.7%)	171 (6.8%)	321 (6.7%)
45–54	535 (19.1%)	439 (17.5%)	880 (18.3%)
55–64	985 (35.1%)	911 (36.4%)	1,718 (35.8%)
65–74	793 (28.3%)	691 (27.6%)	1,340 (27.9%)
75+	272 (9.7%)	241 (9.6%)	467 (9.7%)
Race/Ethnicity			
White, Non-Hispanic	2,361 (84.1%)	2,122 (84.7%)	4,057 (84.4%)
Black, Non-Hispanic	176 (6.3%)	159 (6.3%)	302 (6.3%)
Other, Non-Hispanic	91 (3.2%)	70 (2.8%)	145 (3.0%)
Hispanic	111 (4.0%)	99 (4.0%)	191 (4.0%)
2+ Races, Non-Hispanic	68 (2.4%)	54 (2.2%)	110 (2.3%)
Education			
Less than High School	85 (3.0%)	68 (2.7%)	140 (2.9%)
High School	524 (18.7%)	429 (17.1%)	859 (17.9%)
Some College	1,002 (35.7%)	883 (35.3%)	1,703 (35.4%)
Bachelor’s degree or higher	1,196 (42.6%)	1,124 (44.9%)	2,103 (43.8%)

Note. Study Arms 1 and 2 (No Visual/Absolute Frequency: n = 506) were included in both designs; see Table 1.

Table B. Statistical Print Conditions: Weighted Means (Standard Errors) and Percents of Dependent Variables, by Statistical Format

	Statistical Format					
	Absolute Frequency	Percent	Absolute Frequency + Percent	Relative Frequency	Relative Frequency + Percent	Control
N	268	261	266	249	282	136
Intention: Talk to doctor	2.09 (0.09)	1.98 (0.07)	1.96 (0.08)	1.93 (0.07)	1.89 (0.07)	1.93 (0.11)
Intention: Search for information	2.11 (0.08)	2.04 (0.07)	2.01 (0.08)	2.08 (0.08)	1.97 (0.08)	2.03 (0.12)
Intention: Take drug	2.45 (0.08)	2.34 (0.08)	2.33 (0.08)	2.43 (0.08)	2.37 (0.08)	2.30 (0.11)
Attitude toward the drug	3.04 (0.06)	3.11 (0.06)	3.07 (0.07)	3.13 (0.06)	3.01 (0.06)	2.89 (0.10)
Risk/benefit tradeoff	3.95 (0.10)	4.08 (0.10)	3.96 (0.12)	4.12 (0.11)	3.92 (0.10)	3.60 (0.17)
Perceived benefit	4.25 (0.12)	4.21 (0.13)	4.37 (0.12)	4.50 (0.14)	4.31 (0.12)	4.03 (0.22)
Perceived risk	4.26 (0.11)	4.26 (0.12)	4.12 (0.12)	3.82 (0.11)	4.10 (0.11)	4.56 (0.20)
Risk recall (aided)	5.85 (0.11)	5.82 (0.12)	5.87 (0.11)	5.65 (0.11)	5.93 (0.10)	6.12 (0.15)
Risk recall (unaided)	83.42%	78.00%	83.46%	83.61%	81.40%	83.18%

Note. Intention measures were assessed on a scale (1 = not at all likely, 4 = extremely likely). Attitude toward the drug was assessed on a scale (1 = very bad, 5 = very good). Risk-benefit tradeoff was assessed on a scale (1= more risks than benefits, 7 = more benefits than risks). Perceived benefit was assessed on a scale (1 = not at all effective/well, 7 = very effective/well). Perceived risk was assessed on a scale (1 = very safe/not at all risky, 7 = not at all safe/very risky). Risk recall was the sum of correct responses (0 to 8). Participants in the [Relative Frequency] condition had lower perceived risk than participants in the control condition, $F(1,1457) = 10.67$, $p = 0.001$.

Table C. Statistical Television Conditions: Weighted Means (Standard Errors) and Percents of Dependent Variables, by Statistical Format

	Statistical Format					
	Absolute Frequency	Percent	Absolute Frequency + Percent	Relative Frequency	Relative Frequency + Percent	Control
N	238	244	238	243	258	124
Intention: Talk to doctor	1.89 (0.07)	1.69 (0.07)	1.90 (0.07)	1.93 (0.07)	1.72 (0.07)	1.79 (0.10)
Intention: Search for information	1.99 (0.08)	1.87 (0.08)	2.09 (0.08)	2.09 (0.09)	1.90 (0.07)	1.98 (0.11)
Intention: Take drug	2.39 (0.08)	2.21 (0.07)	2.37 (0.08)	2.30 (0.08)	2.30 (0.08)	2.47 (0.11)
Attitude toward the drug	3.08 (0.06)	2.94 (0.06)	3.02 (0.06)	3.02 (0.07)	2.93 (0.06)	2.88 (0.07)
Risk/benefit tradeoff	3.83 (0.11)	3.75 (0.10)	3.90 (0.10)	3.92 (0.11)	3.76 (0.11)	3.83 (0.14)
Perceived benefit	4.33 (0.13)	4.15 (0.12)	4.48 (0.12)	4.50 (0.12)	4.28 (0.13)	4.31 (0.16)
Perceived risk	4.20 (0.11)	4.36 (0.10)	4.08 (0.10)	4.18 (0.11)	4.30 (0.10)	4.07 (0.16)
Risk recall (aided)	6.40 (0.12)	6.50 (0.12)	6.38 (0.11)	6.49 (0.10)	6.47 (0.12)	6.36 (0.14)
Risk recall (unaided)	85.85%	80.88%	85.01%	87.25%	90.98%	87.83%

Note. Intention measures were assessed on a scale (1 = not at all likely, 4 = extremely likely). Attitude toward the drug was assessed on a scale (1 = very bad, 5 = very good). Risk-benefit tradeoff was assessed on a scale (1= more risks than benefits, 7 = more benefits than risks). Perceived benefit was assessed on a scale (1 = not at all effective/well, 7 = very effective/well). Perceived risk was assessed on a scale (1 = very safe/not at all risky, 7 = not at all safe/very risky). Risk recall was the sum of correct responses (0 to 8).

Table D. Statistical Print Conditions: Weighted Percents of Accuracy Variables, by Statistical Format

	Statistical Format					
	Absolute Frequency	Percent	Absolute Frequency + Percent	Relative Frequency	Relative Frequency + Percent	Control
N	268	261	266	249	282	136
Accuracy: absolute frequency	37.73%	30.81%	41.78%	1.52%	22.84%	3.36%
Accuracy: percent	34.12%	32.37%	37.79%	0.74%	23.38%	2.50%
Accuracy: relative frequency	11.12%	15.19%	11.42%	23.70%	26.50%	9.45%
Accuracy: placebo	33.99%	27.07%	40.96%	0.96%	27.71%	0.00%

Note. Accuracy (absolute frequency, percentage, relative frequency, and placebo) was measured as the percentage who reported the correct quantitative benefit information.

Table E. Statistical Print Conditions: Test Statistics (*t*-tests and *p*-values) from Analyses Comparing Accuracy Measures for Each Statistical Format to the Control Condition

	Statistical Format vs. Control				
	Absolute Frequency	Percent	Absolute Frequency + Percent	Relative Frequency	Relative Frequency + Percent
Accuracy: absolute frequency	8.91, $p < 0.001$	7.10, $p < 0.001$	9.64, $p < 0.001$	-1.03, $p = 0.30$	5.86, $p < 0.001$
Accuracy: percent	8.65, $p < 0.001$	7.87, $p < 0.001$	9.36, $p < 0.001$	-1.30, $p = 0.19$	6.46, $p < 0.001$
Accuracy: relative frequency	0.47, $p = 0.64$	1.41, $p = 0.16$	0.53, $p = 0.60$	3.21, $p = 0.001$	3.98, $p < 0.001$
Accuracy: placebo	9.64, $p < 0.001$	8.06, $p < 0.001$	11.08, $p < 0.001$	2.04, $p = 0.04$	8.39, $p < 0.001$

Note. Accuracy (absolute frequency, percentage, relative frequency, and placebo) was measured as the percentage who reported the correct quantitative benefit information. Bonferroni-adjusted significance was defined as $p < 0.003$. All *t*-tests had 1456 degrees of freedom.

Table F. Statistical Print Conditions: Test Statistics (*t*-tests and *p*-values) from Analyses Comparing Accuracy Measures Among Statistical Formats

	Absolute frequency vs.					Percent vs.				Absolute frequency + percent			Relative frequency vs.
	Percent	Absolute frequency + percent	Relative frequency	Relative frequency + percent		Absolute frequency + percent	Relative frequency	Relative frequency + percent		Relative frequency	Relative frequency + percent		Relative frequency + percent
Accuracy: absolute frequency	1.39, p = 0.17	-0.80, p = 0.43	9.96, p < 0.001	3.25, p = 0.001		-2.15, p = 0.03	8.03, p < 0.001	1.73, p = 0.08		10.67, p < 0.001	4.03, p < 0.001		-6.95, p < 0.001
Accuracy: percent	0.35, p = 0.72	-0.74, p = 0.46	9.68, p < 0.001	2.37, p = 0.02		-1.08, p = 0.28	8.80, p < 0.001	1.94, p = 0.05		10.38, p < 0.001	3.12, p = 0.002		-7.56, p < 0.001
Accuracy: relative frequency	-1.20, p = 0.23	-0.10, p = 0.92	-3.26, p = 0.001	-4.19, p < 0.001		1.07, p = 0.29	-1.97, p = 0.05	-2.73, p = 0.006		-3.09, p = 0.002	-3.97, p < 0.001		-0.62, p = 0.53
Accuracy: placebo	1.42, p = 0.16	-1.37, p = 0.17	9.29, p < 0.001	1.30, p = 0.19		-2.78, p = 0.005	7.70, p < 0.001	-0.14, p = 0.89		10.74, p < 0.001	2.67, p = 0.008		-8.02, p < 0.001

Note. Accuracy (absolute frequency, percentage, relative frequency, and placebo) was measured as the percentage who reported the correct quantitative benefit information. Bonferroni-adjusted significance was defined as $p < 0.003$. All *t*-tests had 1456 degrees of freedom.

Table G. Statistical Television Conditions: Weighted Percents of Accuracy Variables, by Statistical Format

	Statistical Format					
	Absolute Frequency	Percent	Absolute Frequency + Percent	Relative Frequency	Relative Frequency + Percent	Control
N	238	244	238	243	258	124
Accuracy: absolute frequency	27.85%	33.20%	43.69%	4.28%	21.55%	3.52%
Accuracy: percent	26.63%	32.32%	40.58%	4.34%	20.36%	4.09%
Accuracy: relative frequency	19.78%	15.18%	16.17%	26.68%	33.05%	8.51%
Accuracy: placebo	31.42%	32.95%	37.01%	0.42%	24.91%	0.56%

Note. Accuracy (absolute frequency, percentage, relative frequency, and placebo) was measured as the percentage who reported the correct quantitative benefit information.

Table H. Statistical Television Conditions: Test Statistics (*t*-tests and *p*-values) from Analyses Comparing Accuracy Measures for Each Statistical Format to the Control Condition

	Statistical Format vs. Control				
	Absolute Frequency	Percent	Absolute Frequency + Percent	Relative Frequency	Relative Frequency + Percent
Accuracy: absolute frequency	6.63, $p < 0.001$	7.41, $p < 0.001$	9.83, $p < 0.001$	0.30, $p = 0.76$	5.28, $p < 0.001$
Accuracy: percent	6.12, $p < 0.001$	7.10, $p < 0.001$	8.89, $p < 0.001$	0.10, $p = 0.92$	4.90, $p < 0.001$
Accuracy: relative frequency	2.78, $p = 0.005$	1.74, $p = 0.08$	1.98, $p = 0.05$	4.10, $p < 0.001$	5.61, $p < 0.001$
Accuracy: placebo	8.93, $p < 0.001$	9.02, $p < 0.001$	10.02, $p < 0.001$	-0.19, $p = 0.85$	7.71, $p < 0.001$

Note. Accuracy (absolute frequency, percentage, relative frequency, and placebo) was measured as the percentage who reported the correct quantitative benefit information. Bonferroni-adjusted significance was defined as $p < 0.003$. All *t*-tests had 1339 degrees of freedom.

Table I. Statistical Television Conditions: Test Statistics (*t*-tests and *p*-values) from Analyses Comparing Accuracy Measures Among Statistical Formats

	Absolute frequency vs.					Percent vs.				Absolute frequency + percent			Relative frequency vs.
	Percent	Absolute frequency + percent	Relative frequency	Relative frequency + percent		Absolute frequency + percent	Relative frequency	Relative frequency + percent		Relative frequency	Relative frequency + percent		Relative frequency + percent
Accuracy: absolute frequency	-1.10, p = 0.27	-3.20, p = 0.001	6.25, p < 0.001	1.43, p = 0.15		-2.02, p = 0.04	7.06, p < .001	2.48, p = 0.01		9.43, p < 0.001	4.65, p < 0.001		-4.90, p < 0.001
Accuracy: percent	-1.18, p = 0.24	-2.84, p = 0.005	6.02, p < 0.001	1.47, p = 0.14		-1.61, p = 0.11	7.01, p < .001	2.64, p = 0.008		8.79, p < 0.001	4.35, p < 0.001		-4.79, p < 0.001
Accuracy: relative frequency	1.15, p = 0.25	0.89, p = 0.37	-1.51, p = 0.13	-2.94, p = 0.003		-0.26, p = 0.80	-2.62, p = .009	-4.13, p < 0.001		-2.38, p = 0.02	-3.87, p < 0.001		-1.31, p = 0.19
Accuracy: placebo	-0.31, p = 0.76	-1.13, p = 0.26	9.02, p < 0.001	1.41, p = 0.16		-0.80, p = 0.42	9.10, p < .001	1.70, p = 0.09		10.11, p < 0.001	2.55, p = 0.01		-7.80, p < 0.001

Note. Accuracy (absolute frequency, percentage, relative frequency, and placebo) was measured as the percentage who reported the correct quantitative benefit information. Bonferroni-adjusted significance was defined as $p < 0.003$. All *t*-tests had 1339 degrees of freedom.

Table J. Visual Print Conditions: Weighted Means (Standard Errors) and Percents of Dependent Variables, by Visual Format

	Visual Format				
	Pie Chart	Bar Chart	Table	Pictograph	No Visual
N	269	262	251	259	268
Intention: Talk to doctor	2.07 (0.08)	1.85 (0.07)	2.07 (0.08)	1.92 (0.07)	2.09 (0.09)
Intention: Search for information	2.05 (0.07)	1.91 (0.07)	2.00 (0.08)	2.08 (0.08)	2.11 (0.08)
Intention: Take drug	2.48 (0.08)	2.34 (0.08)	2.31 (0.09)	2.35 (0.08)	2.45 (0.08)
Attitude toward the drug	3.15 (0.07)	2.95 (0.07)	3.03 (0.09)	3.12 (0.06)	3.04 (0.06)
Risk/benefit tradeoff	4.04 (0.12)	3.75 (0.12)	3.78 (0.12)	4.04 (0.11)	3.95 (0.10)
Perceived benefit	4.29 (0.14)	4.26 (0.13)	4.28 (0.16)	4.37 (0.15)	4.25 (0.12)
Perceived risk	4.11 (0.11)	4.25 (0.11)	4.24 (0.13)	4.25 (0.12)	4.26 (0.11)
Risk recall (aided)	5.94 (0.10)	5.84 (0.11)	6.11 (0.14)	5.93 (0.11)	5.85 (0.11)
Risk recall (unaided)	80.81%	84.54%	87.57%	83.56%	83.42%

Note. Intention measures were assessed on a scale (1 = not at all likely, 4 = extremely likely). Attitude toward the drug was assessed on a scale (1 = very bad, 5 = very good). Risk-benefit tradeoff was assessed on a scale (1= more risks than benefits, 7 = more benefits than risks). Perceived benefit was assessed on a scale (1 = not at all effective/well, 7 = very effective/well). Perceived risk was assessed on a scale (1 = very safe/not at all risky, 7 = not at all safe/very risky). Risk recall was the sum of correct responses (0 to 8).

Table K. Visual Television Conditions: Weighted Means (Standard Errors) and Percents of Dependent Variables, by Visual Format

	Visual Format				
	Pie Chart	Bar Chart	Table	Pictograph	No Visual
N	241	237	238	241	238
Intention: Talk to doctor	1.91 (0.07)	1.86 (0.08)	1.97 (0.07)	1.92 (0.07)	1.89 (0.07)
Intention: Search for information	1.90 (0.07)	2.02 (0.08)	1.99 (0.07)	2.06 (0.07)	1.99 (0.07)
Intention: Take drug	2.35 (0.08)	2.43 (0.09)	2.34 (0.08)	2.36 (0.08)	2.39 (0.08)
Attitude toward the drug	3.00 (0.06)	2.95 (0.06)	3.00 (0.06)	3.06 (0.06)	3.08 (0.06)
Risk/benefit tradeoff	3.78 (0.11)	3.88 (0.13)	3.90 (0.12)	3.98 (0.12)	3.83 (0.11)
Perceived benefit	4.27 (0.14)	4.41 (0.15)	4.40 (0.13)	4.54 (0.12)	4.33 (0.13)
Perceived risk	4.11 (0.11)	4.03 (0.12)	4.19 (0.09)	4.01 (0.10)	4.20 (0.11)
Risk recall (aided)	6.55 (0.11)	6.66 (0.10)	6.41 (0.13)	6.51 (0.10)	6.40 (0.12)
Risk recall (unaided)	85.55%	87.35%	85.26%	87.36%	85.85%

Note. Intention measures were assessed on a scale (1 = not at all likely, 4 = extremely likely). Attitude toward the drug was assessed on a scale (1 = very bad, 5 = very good). Risk-benefit tradeoff was assessed on a scale (1= more risks than benefits, 7 = more benefits than risks). Perceived benefit was assessed on a scale (1 = not at all effective/well, 7 = very effective/well). Perceived risk was assessed on a scale (1 = very safe/not at all risky, 7 = not at all safe/very risky). Risk recall was the sum of correct responses (0 to 8).

Table L. Visual Print Conditions: Weighted Percents of Accuracy Variables, by Visual Format

	Visual Format				
	Pie Chart	Bar Chart	Table	Pictograph	No Visual
N	269	262	251	259	268
Accuracy: absolute frequency	41.51%	52.75%	51.90%	36.53%	37.73%
Accuracy: percent	38.51%	50.98%	48.12%	35.06%	34.12%
Accuracy: Relative frequency	17.40%	20.96%	21.97%	15.80%	11.12%
Accuracy: placebo	38.52%	49.17%	58.00%	46.83%	33.99%

Note. Accuracy (absolute frequency, percentage, relative frequency, and placebo) was measured as the percent who reported the correct quantitative benefit information.

Table M. Visual Print Conditions: Test Statistics (*t*-tests and *p*-values) from Analyses Comparing Accuracy Measures for Each Visual Format to the No Visual Condition

	Visual Format vs. No Visual			
	Pie Chart	Bar Chart	Table	Pictograph
Accuracy: Absolute frequency	0.80, $p = 0.42$	2.95, $p = 0.003$	2.74, $p = 0.006$	-0.10, $p = 0.92$
Accuracy: percent	0.96, $p = 0.34$	3.40, $p = 0.001$	2.81, $p = 0.005$	0.36, $p = 0.72$
Accuracy: relative frequency	2.07, $p = 0.04$	2.93, $p = 0.003$	3.11, $p = 0.002$	1.70, $p = 0.09$
Accuracy: placebo	0.91, $p = 0.36$	2.85, $p = 0.004$	4.23, $p < 0.001$	2.38, $p = 0.02$

Note. Accuracy (absolute frequency, percentage, relative frequency, and placebo) was measured as the percentage who reported the correct quantitative benefit information. Bonferroni-adjusted significance was defined as $p < 0.005$. All *t*-tests had 1308 degrees of freedom.

Table N. Visual Print Conditions: Test Statistics (*t*-tests and *p*-values) from Analyses Comparing Accuracy Measures Among Visual Formats

	Pie Chart vs.				Bar Chart vs.			Table vs.
	Bar Chart	Table	Pictograph		Table	Pictograph		Pictograph
Accuracy: absolute frequency	-2.10, $p = 0.04$	-1.84, $p = 0.07$	0.94, $p = 0.35$		0.15, $p = 0.88$	2.99, $p = 0.003$		2.69, $p = 0.007$
Accuracy: percent	-2.36, $p = 0.02$	-1.73, $p = 0.08$	0.67, $p = 0.51$		0.50, $p = 0.62$	2.95, $p = 0.003$		2.30, $p = 0.02$
Accuracy: relative frequency	-0.91, $p = 0.36$	-1.04, $p = 0.30$	0.44, $p = 0.66$		-0.22, $p = 0.83$	1.31, $p = 0.19$		1.40, $p = 0.16$
Accuracy: placebo	-2.03, $p = 0.04$	-3.55, $p < 0.001$	-1.54, $p = 0.12$		-1.54, $p = 0.12$	0.41, $p = 0.68$		1.91, $p = 0.06$

Note. Accuracy (absolute frequency, percentage, relative frequency, and placebo) was measured as the percent who reported the correct quantitative benefit information. Bonferroni-adjusted significance was defined as $p < 0.005$. All *t*-tests had 1304 degrees of freedom.

Table O. Visual Television Conditions: Weighted Percents of Accuracy Variables, by Visual Format

	Visual Format				
	Pie Chart	Bar Chart	Table	Pictograph	No Visual
N	241	237	238	241	238
Accuracy: Absolute frequency	56.20%	68.56%	52.15%	47.89%	27.85%
Accuracy: percent	52.37%	65.58%	50.96%	44.93%	26.63%
Accuracy: relative frequency	28.22%	31.90%	19.83%	19.22%	19.78%
Accuracy: placebo	61.72%	72.01%	66.69%	61.87%	31.42%

Note. Accuracy (absolute frequency, percentage, relative frequency, and placebo) was measured as the percentage who reported the correct quantitative benefit information.

Table P. Visual Television Conditions: Test Statistics (*t*-tests and *p*-values) from Analyses Comparing Accuracy Measures for Each Visual Format to the No Visual Condition

	Visual Format vs. No Visual			
	Pie Chart	Bar Chart	Table	Pictograph
Accuracy: absolute frequency	5.45, $p < 0.001$	7.57, $p < 0.001$	4.68, $p < 0.001$	3.99, $p < 0.001$
Accuracy: percent	5.01, $p < 0.001$	7.26, $p < 0.001$	4.70, $p < 0.001$	3.69, $p < 0.001$
Accuracy: Relative frequency	1.91, $p = 0.06$	2.81, $p = 0.005$	-0.01, $p = 0.99$	-0.13, $p = 0.90$
Accuracy: placebo	5.66, $p < 0.001$	7.26, $p < 0.001$	6.57, $p < 0.001$	5.73, $p < 0.001$

Note. Accuracy (absolute frequency, percent, relative frequency, and placebo) was measured as the percent who reported the correct quantitative benefit information. Bonferroni-adjusted significance was defined as $p < 0.005$. All *t*-tests had 1194 degrees of freedom.

Table Q. Visual Television Conditions: Test Statistics (*t*-tests and *p*-values) from Analyses Comparing Accuracy Measures Among Visual Formats

	Pie Chart vs.				Bar Chart vs.			Table vs.
	Bar Chart	Table	Pictograph		Table	Pictograph		Pictograph
Accuracy: absolute frequency	-2.39, $p = 0.02$	0.76, $p = 0.44$	1.58, $p = 0.12$		3.15, $p = 0.002$	3.99, $p < 0.001$		0.80, $p = 0.42$
Accuracy: percent	-2.50, $p = 0.01$	0.26, $p = 0.79$	1.41, $p = 0.16$		2.75, $p = 0.006$	3.92, $p < 0.001$		1.14, $p = 0.26$
Accuracy: relative frequency	-0.75, $p = 0.46$	1.86, $p = 0.06$	2.08, $p = 0.04$		2.56, $p = 0.01$	2.79, $p = 0.005$		0.15, $p = 0.88$
Accuracy: placebo	-2.00, $p = 0.04$	-0.97, $p = 0.33$	-0.03, $p = 0.98$		1.06, $p = 0.29$	1.99, $p = 0.05$		0.95, $p = 0.34$

Note. Accuracy (absolute frequency, percentage, relative frequency, and placebo) was measured as the percentage who reported the correct quantitative benefit information. Bonferroni-adjusted significance was defined as $p < 0.005$. All *t*-tests had 1190 degrees of freedom.

Table R. Statistical Print and Television Conditions: Weighted Means (Standard Errors) and Percents of Dependent Variables, by Drug Efficacy Level

	Print Conditions				Television Conditions		
	Drug Efficacy Level				Drug Efficacy Level		
	High	Low	Control		High	Low	Control
N	646	680	136		600	621	124
Intention: Talk to doctor	1.99 (0.05)	1.95 (0.04)	1.93 (0.11)		1.88 (0.05)	1.77 (0.04)	1.79 (0.10)
Intention: Search for information	2.08 (0.05)	2.00 (0.05)	2.03 (0.12)		2.06 (0.05)	1.91 (0.05)	1.98 (0.11)
Intention: Take drug	2.41 (0.05)	2.36 (0.05)	2.30 (0.11)		2.39 (0.05)	2.24 (0.05)	2.47 (0.11)
Attitude toward the drug	3.10 (0.04)	3.04 (0.04)	2.89 (0.10)		3.09 (0.04)	2.91 (0.04)	2.88 (0.07)
Risk/benefit tradeoff	4.12 (0.07)	3.89 (0.06)	3.60 (0.17)		3.98 (0.07)	3.69 (0.07)	3.83 (0.14)
Perceived benefit	4.42 (0.08)	4.24 (0.08)	4.03 (0.22)		4.46 (0.08)	4.24 (0.08)	4.31 (0.16)
Perceived risk	4.11 (0.08)	4.11 (0.08)	4.56 (0.20)		4.13 (0.07)	4.32 (0.07)	4.23 (0.05)
Risk recall (aided)	5.86 (0.07)	5.80 (0.07)	6.12 (0.15)		6.46 (0.07)	6.44 (0.07)	6.36 (0.14)
Risk recall (unaided)	83.67%	80.41%	83.18%		85.32%	86.69%	87.83%

Note. Intention measures were assessed on a scale (1 = not at all likely, 4 = extremely likely). Attitude toward the drug was assessed on a scale (1 = very bad, 5 = very good). Risk-benefit tradeoff was assessed on a scale (1= more risks than benefits, 7 = more benefits than risks). Perceived benefit was assessed on a scale (1 = not at all effective/well, 7 = very effective/well). Perceived risk was assessed on a scale (1 = very safe/not at all risky, 7 = not at all safe/very risky). Risk recall was the sum of correct responses (0 to 8).

Table S. Statistical Print and Television Conditions: Test Statistics (F-tests and p-values) from Analyses Comparing Dependent Measures Among Drug Efficacy Levels

	Print Conditions				Television Conditions		
	Drug Efficacy Level				Drug Efficacy Level		
	High vs. Low	High vs. Control	Low vs. Control		High vs. Low	High vs. Control	Low vs. Control
Intention: Talk to doctor	*ns*	*ns*	*ns*		*ns*	*ns*	*ns*
Intention: Search for information	*ns*	*ns*	*ns*		F (1, 1213) = 4.36, p = 0.04 [a]	*ns*	*ns*
Intention: Take drug	*ns*	*ns*	*ns*		F (1, 1207) = 4.74, p = 0.03	*ns*	*ns*
Attitude toward the drug	*ns*	*ns*	*ns*		F (1, 1219) = 9.56, p = 0.002	F (1, 1342) = 6.51, p = 0.01	*ns*
Risk/benefit tradeoff	F (1, 1321) = 6.00, p = 0.01	--	--		F (1, 1216) = 9.31, p = 0.002	--	--
Perceived benefit	*ns*	*ns*	*ns*		F (1, 1220) = 3.70, p = 0.055	*ns*	*ns*
Perceived risk	*ns*	--	--		F (1, 1218) = 3.85, p = 0.05	--	--
Risk recall (aided)	*ns*	*ns*	*ns*		*ns*	*ns*	*ns*
Risk recall (unaided)	*ns*	*ns*	*ns*		*ns*	*ns*	*ns*

Note. *ns* = not significant. [a] This effect was qualified by an interaction with statistical format, F (1, 1213) = 3.12, p = 0.01.

Table T. Visual Print and Television Conditions: Weighted Means (Standard Errors) and Percents of Dependent Variables, by Drug Efficacy Level

	Print Conditions				Television Conditions		
	Drug Efficacy Level				Drug Efficacy Level		
	High	Low	Control		High	Low	Control
N	641	668	136		599	596	124
Intention: Talk to doctor	2.05 (0.05)	1.96 (0.05)	1.93 (0.11)		1.98 (0.04)	1.84 (0.05)	1.79 (0.10)
Intention: Search for information	2.06 (0.05)	2.00 (0.05)	2.03 (0.12)		2.05 (0.05)	1.94 (0.05)	1.98 (0.11)
Intention: Take drug	2.42 (0.05)	2.36 (0.05)	2.30 (0.11)		2.52 (0.05)	2.23 (0.05)	2.47 (0.11)
Attitude toward the drug	3.22 (0.04)	2.90 (0.05)	2.89 (0.10)		3.16 (0.04)	2.88 (0.04)	2.88 (0.07)
Risk/benefit tradeoff	4.13 (0.07)	3.71 (0.07)	3.60 (0.17)		4.25 (0.07)	3.50 (0.07)	3.83 (0.14)
Perceived benefit	4.56 (0.09)	4.03 (0.08)	4.03 (0.22)		4.75 (0.08)	4.03 (0.09)	4.31 (0.16)
Perceived risk	4.09 (0.08)	4.35 (0.07)	4.56 (0.20)		3.91 (0.07)	4.31 (0.07)	4.07 (0.16)
Risk recall (aided)	5.98 (0.07)	5.88 (0.08)	6.12 (0.15)		6.51 (0.07)	6.51 (0.07)	6.36 (0.14)
Risk recall (unaided)	85.41%	82.48%	83.18%		84.25%	88.28%	87.83%

Note. Intention measures were assessed on a scale (1 = not at all likely, 4 = extremely likely). Attitude toward the drug was assessed on a scale (1 = very bad, 5 = very good). Risk-benefit tradeoff was assessed on a scale (1= more risks than benefits, 7 = more benefits than risks). Perceived benefit was assessed on a scale (1 = not at all effective/well, 7 = very effective/well). Perceived risk was assessed on a scale (1 = very safe/not at all risky, 7 = not at all safe/very risky). Risk recall was the sum of correct responses (0 to 8).

Table U. Visual Print and Television Conditions: Test Statistics (*F*-tests and *p*-values) from Analyses Comparing Dependent Measures Among Drug Efficacy Levels

	Print Conditions				Television Conditions		
	Drug Efficacy Level				Drug Efficacy Level		
	High vs. Low	High vs. Control	Low vs. Control		High vs. Low	High vs. Control	Low vs. Control
Intention: Talk to doctor	*ns*	*ns*	*ns*		F (1,1191) = 5.07, p = 0.02	*ns*	*ns*
Intention: Search for information	*ns*	*ns*	*ns*		*ns*	*ns*	*ns*
Intention: Take drug	*ns*	*ns*	*ns*		F (1, 1184) = 16.45, p < 0.001	*ns*	*ns*
Attitude toward the drug	F (1, 1306) = 26.69, p < 0.001	F (1, 1441) = 9.31, p = 0.002	*ns*		F (1, 1194) = 24.77, p < 0.001	F (1, 1317) = 12.31, p < 0.001	*ns*
Risk/benefit tradeoff	F (1, 1304) = 17.42, p < 0.001	--	--		F (1, 1193) = 56.87, p < 0.001	--	--
Perceived benefit	F (1, 1307) = 17.43, p < 0.001	*ns*	*ns*		F (1, 1194) = 37.47, p < 0.001	F (1, 1318) = 6.25, p = 0.01	*ns*
Perceived risk	F (1, 1306) = 6.45, p = 0.01	--	--		F (1, 1192) = 18.62, p < 0.001	--	--
Risk recall (aided)	*ns*	*ns*	*ns*		*ns*	*ns*	*ns*
Risk recall (unaided)	*ns*	*ns*	*ns*		*ns*	*ns*	*ns*

Note. *ns* = not significant.

Table V. Statistical Print and Television Conditions: Weighted Percents of Accuracy Variables, by Drug Efficacy Level

	Print Conditions				Television Conditions		
	Drug Efficacy Level				Drug Efficacy Level		
	High	Low	Control		High	Low	Control
N	646	680	136		600	621	124
Accuracy: absolute frequency	25.72%	28.63%	3.36%		24.21%	27.70%	3.52%
Accuracy: percent	24.53%	27.32%	2.50%		22.37%	27.00%	4.09%
Accuracy: Relative frequency	10.64%	24.23%	9.45%		15.44%	29.00%	8.51%
Accuracy: placebo	27.68%	25.35%	0.00%		24.83%	25.70%	0.56%

Note. Accuracy (absolute frequency, percent, relative frequency, and placebo) was measured as the percentage who reported the correct quantitative benefit information.

Table W. Statistical Print and Television Conditions: Test Statistics (*F*-tests, *t*-tests, and *p*-values) from Analyses Comparing Accuracy Measures Among Drug Efficacy Levels

	Print Conditions				Television Conditions		
	Drug Efficacy Level				Drug Efficacy Level		
	High vs. Low	High vs. Control	Low vs. Control		High vs. Low	High vs. Control	Low vs. Control
	F (1, 1325)	*t* (1461)	*t* (1461)		*F* (1, 1220)	*t* (1344)	*t* (1344)
Accuracy: absolute frequency	0.78, $p = 0.38$ [b]	4.68, $p < 0.001$	4.99, $p < 0.001$		1.80, $p = 0.18$	4.26, $p < 0.001$	4.62, $p < 0.001$
Accuracy: percent	0.78, $p = .38$	4.71, $p < 0.001$	4.99, $p < 0.001$		3.16, $p = 0.08$	4.10, $p < 0.001$	4.64, $p < 0.001$
Accuracy: relative frequency	33.88, $p < 0.001$ [c]	0.35, $p = 0.73$	3.08, $p = 0.002$		23.44, $p < 0.001$ [d]	1.79, $p = 0.07$	4.02, $p < 0.001$
Accuracy: placebo	0.86, $p = 0.36$	--[a]	--[a]		0.16, $p = 0.69$ [e]	4.03, $p < 0.001$	4.08, $p < 0.001$

Note. Accuracy (absolute frequency, percent, relative frequency, and placebo) was measured as the percentage who reported the correct quantitative benefit information.

[a] Test statistics could not be calculated because the control condition percent equaled 0.
[b] This was qualified by an interaction between drug efficacy level and statistical format, F (4, 1325) = 2.46, $p = 0.04$.
[c] This was qualified by an interaction between drug efficacy level and statistical format, F (4, 1325) = 2.49, $p = 0.04$.
[d] This was qualified by an interaction between drug efficacy level and statistical format, F (4, 1220) = 4.11, $p = 0.003$.
[e] This was qualified by an interaction between drug efficacy and statistical format, F (4, 1220) = 3.37, $p = 0.01$.

Table X. Visual Print and Television Conditions: Weighted Percents of Accuracy Variables, by Drug Efficacy Level

	Print Conditions				Television Conditions		
	Drug Efficacy Level				Drug Efficacy Level		
	High	Low	Control		High	Low	Control
N	641	668	136		599	596	124
Accuracy: absolute frequency	36.18%	51.49%	3.36%		45.54%	55.52%	3.52%
Accuracy: percent	31.61%	50.54%	2.50%		42.11%	54.09%	4.09%
Accuracy: Relative frequency	8.92%	25.51%	9.45%		12.19%	35.43%	8.51%
Accuracy: placebo	44.34%	45.83%	0.00%		58.61%	58.89%	0.56%

Note. Accuracy (absolute frequency, percentage, relative frequency, and placebo) was measured as the percentage who reported the correct quantitative benefit information.

Table Y. Visual Print and Television Conditions: Test Statistics (*F*-tests, *t*-tests, and *p*-values) from Analyses Comparing Accuracy Measures Among Drug Efficacy Levels

	Print Conditions				Television Conditions		
	Drug Efficacy Level				Drug Efficacy Level		
	High vs. Low	High vs. Control	Low vs. Control		High vs. Low	High vs. Control	Low vs. Control
	F (1, 1308)	*t* (1444)	*t* (1444)		*F* (1, 1194)	*t* (1318)	*t* (1318)
Accuracy: absolute frequency	20.90, $p < 0.001$	5.70, $p < 0.001$	6.98, $p < 0.001$		9.78, $p = 0.002$	6.19, $p < 0.001$	6.98, $p < 0.001$
Accuracy: percent	32.91, $p < 0.001$	5.37, $p < 0.001$	6.86, $p < 0.001$		13.84, $p < 0.001$	6.12, $p < 0.001$	7.17, $p < 0.001$
Accuracy: relative frequency	46.32, $p < 0.001$	-0.17, $p = 0.87$	3.26, $p = 0.001$		61.52, $p < 0.001$	1.05, $p = 0.30$	4.84, $p < 0.001$
Accuracy: placebo	0.34, $p = 0.56$	--[a]	--[a]		0.02, $p = 0.90$	5.48, $p < 0.001$	5.49, $p < 0.001$

Note. Accuracy (absolute frequency, percentage, relative frequency, and placebo) was measured as the percentage who reported the correct quantitative benefit information.

[a] Test statistics could not be calculated because the control condition percent equaled 0.