Erschienen in: Psychological Science ; 23 (2012), 6. - S. 547-553

General Article

Psychological Science
23(6) 547–553

DOI: 10.1177/0956797612437428
http://pss.sagepub.com

Psychological Research and the Prostate-Cancer Screening Controversy

Hal R. Arkes[1] and Wolfgang Gaissmaier[2]
[1]Department of Psychology, Ohio State University, and [2]Harding Center for Risk Literacy, Max Planck Institute for Human Development, Berlin, Germany

Abstract

In October of 2011, the U.S. Preventive Services Task Force released a draft report in which they recommended against using the prostate-specific antigen (PSA) test to screen for prostate cancer. We attempt to show that four factors documented by psychological research can help explain the furor that followed the release of the task force's report. These factors are the persuasive power of anecdotal (as opposed to statistical) evidence, the influence of personal experience, the improper evaluation of data, and the influence of low base rates on the efficacy of screening tests. We suggest that augmenting statistics with facts boxes or pictographs might help such committees communicate more effectively with the public and with the U.S. Congress.

Keywords

decision making, health, judgment, policy making, scientific communication

Received 11/12/11; Revision accepted 1/9/12

In early October of 2011, the U.S. Preventive Services Task Force (USPSTF) released a draft report in which they recommended against using the prostate-specific antigen (PSA) test to screen for prostate cancer. The resulting furor, fueled by presidential candidates, spokespersons for advocacy organizations, and prostate-cancer survivors, involved a number of serious misunderstandings. This episode closely resembled the situation in 1997, when the National Institutes of Health (NIH) Consensus Development Conference concluded that women ages 40 to 49 should decide for themselves whether to get mammograms (NIH Consensus Development Panel, 1997). This recommendation contradicted the prevailing view of many organizations—that mammograms should be routinely given to women in this age group. Both the PSA and the mammogram controversies engendered rancorous "discussion" punctuated by denigrating personal attacks on the panel members by politicians and other individuals.

Our goal here is to examine the PSA testing controversy from the perspective of psychologists who do research in the area of judgment and decision making. We hope to resolve this paradox: How can the personal experience of some people be so contrary to the scientific evidence that motivated the panel's recommendations? A secondary goal is to provide some guidance to future panels so that they might be able to communicate their science-based recommendations more successfully.

PSA Screening: Some Background and Clinical Evidence

We begin by very briefly outlining the goals of PSA screening and the bases for the USPSTF report.

First, it is important to understand that the goal of all cancer screening methods is not simply the early detection of disease. Rather, screening is aimed at reducing mortality or improving quality of life. Screening is targeted at people without symptoms, to test for hidden disease. To be useful, early detection needs to enable earlier treatment that is either more effective or safer than later treatment.

Second, it is important to understand that medical treatments usually both have benefits and cause harm, and this is also true for cancer screening. Although the principal alleged benefit of PSA screening—the reduction of prostate-cancer mortality—is relatively obvious, the harms associated with treatment are more subtle. The most important harm is overdiagnosis, which can eventually lead to overtreatment. Overdiagnosis is defined as the detection of an abnormality that

Corresponding Author:
Hal R. Arkes, Department of Psychology, Ohio State University, Columbus, OH 43210-1234
E-mail: arkes.1@osu.edu

would never progress to cause problems in a patient's lifetime, such as a nonprogressive prostate cancer. Treating a nonprogressive prostate cancer is obviously not beneficial, and it can even be harmful and cause impotence or incontinence as a side effect. In addition to conveying the risk of overdiagnosis and overtreatment, participation in screening bears the risk of yielding false-positive test results that can lead to psychological distress and unnecessary biopsies.

The USPSTF based its conclusion on the best available scientific evidence about the benefits and harms of PSA screening. We highlight the central facts here. In a recent large clinical trial, it was demonstrated that over a period of 9 years, out of every 1,410 men who regularly participated in prostate-cancer screening, there was 1 fewer death due to prostate cancer compared with an equally large group of men who did not participate in screening (Schröder et al., 2009). Additionally, the comparison with men who did not participate in screening revealed that 48 out of the 1,410 men who regularly participated in prostate-cancer screening were unnecessarily treated (i.e., without receiving a benefit such as increased longevity) and hence subjected to risks such as incontinence and impotence. Furthermore, in this study, PSA screening did not reduce overall mortality at all. That is, the slight reduction in mortality due to prostate cancer was balanced out by a slight increase in mortality due to other causes. Also, a second large clinical trial, reported in the same issue of the same medical journal, found no reduction in prostate-cancer mortality through screening (Andriole et al., 2009). More recently, two meta-analyses summarizing the best available evidence on PSA screening concluded that there is currently no indication that it is effective in reducing mortality (Djulbegovic et al., 2010; Ilic, O'Connor, Green, & Wilt, 2011). Also, a 20-year follow-up did not find a benefit of PSA screening, whereas the harms, such as unnecessary treatments, are unquestioned (Sandblom, Varenhorst, Rosell, Löfman, & Carlsson, 2011).

We turn now to the factors that we hypothesize made the expression of these facts in the USPSTF report so controversial.

Factor 1: The Power of the Anecdote

Several studies have shown that an anecdote or two can have a more powerful effect on decision making than a compendium of more reliable statistical data (de Wit, Das, & Vet, 2008; Fagerlin, Wang, & Ubel, 2005; Hamill, Wilson, & Nisbett, 1980; Ubel, Jepson, & Baron, 2001). For example, Fagerlin et al. (2005) presented people with statistics indicating that bypass surgery had a 75% cure rate for angina, and balloon angioplasty had a 50% cure rate. All participants also read anecdotes purportedly written by persons who had successful or unsuccessful outcomes of one of the treatments. For half of the participants, the proportion of successful-treatment anecdotes was representative of the actual base rates of success (i.e., 75% for bypass surgery and 50% for angioplasty). The other half of the participants read anecdotes that contained equal numbers of successful and unsuccessful outcomes for each treatment. Thus, for these participants, the frequency of the successful and unsuccessful anecdotal outcomes was not representative of the statistics for surgery. When participants were asked to choose which treatment they would prefer, the percentage of the representative-anecdote group who chose bypass surgery was more than twice the percentage of the unrepresentative-anecdote group who chose that surgery (41% vs. 20%). The anecdotes thus had a substantial influence on treatment choice, even though identical statistical data were presented to the two groups. Public comments on the USPSTF Web site are replete with just such anecdotes. These members of the public were not persuaded by the statistical data but were convinced by the case of an individual they knew. This same type of anecdotal analysis was used by several U.S. politicians, such as former New York City Mayor Rudy Giuliani (Campanile, 2011) and U.S. Congress member Donna Christensen (2011).

Closely related to the power of the anecdote is the *identifiable-victim effect*, which is defined as the willingness "to expend greater resources to save the lives of identified victims than to save equal numbers of unidentified or statistical victims" (Jenni & Loewenstein, 1997, p. 235). Charities often utilize this principle by highlighting an individual "poster child" or by pairing a potential donor with one potential recipient of the donor's largesse. Statistical lives tend to elicit weaker reactions than specific, individual lives. In fact, Slovic (2007) has shown that people who are informed about a large number of statistical victims tend to experience psychological numbness, rather than enhanced concern. Surprisingly, this identifiable-victim effect can be observed even when the comparison is between one victim and a group of as few as eight victims (Kogut & Ritov, 2005).

When a reader of the USPSTF report tries to digest the information about statistical lives, this information does not have the same impact as information, say, about the reader's mail carrier's older brother who had a positive PSA test, a biopsy, and a radical prostatectomy, and is now still alive. The information that "no trial has shown a decrease in overall mortality with the use of PSA-based screening through 11 years of followup" (USPSTF, 2011, first sentence in the section "Estimate of Magnitude of Net Benefit") will not have the same probative value as awareness of a putative identified beneficiary of the PSA test. Thus, such anecdotes may contribute to the widespread gross overestimation of the benefits of PSA screening (Gigerenzer, Mata, & Frank, 2009).

Factor 2: Epidemiology Versus Personal Experience

Consider two auditoriums, each of which contains 1,000 men age 50 or older. Auditorium "Screened" contains 1,000 men who have had a PSA screening test. Auditorium "Not Screened" contains 1,000 men who have not had such a test. About 8 men from each auditorium will die from prostate cancer in the next 10 years (Djulbegovic et al., 2010). A very

important conclusion to be drawn from these numbers is that screening does not decrease prostate-cancer mortality. How can this be, given that so many men claim to have been saved by a PSA test?

Let us take a look at three possible subgroups of men in Auditorium Screened. These three subgroups have in common that they received a positive PSA test result. The first group consists of those men whose positive PSA test actually did detect a progressive prostate cancer and who are still alive because their cancer was detected early and subsequently treated successfully. This group's members rightfully believe that they were saved by a PSA test. However, given that the USPSTF report and some of the meta-analyses cited earlier raise doubt that PSA screening does in fact reduce prostate-cancer mortality (not to speak of overall mortality), it is unclear whether there actually are such men in the auditorium. Even in light of the more favorable trials, such as the European Randomized Study of Screening for Prostate Cancer (Schröder et al., 2009) and its recent update (Schröder et al., 2012), it seems safe to assume that PSA screening saved at most 1 man in the auditorium from dying from prostate cancer; and in light of the less favorable evidence (Andriole et al., 2009; Djulbegovic et al., 2010; Ilic et al., 2011; Sandblom et al., 2011), quite possibly none have been saved.

The second group in Auditorium Screened consists of about 20 men who were unnecessarily diagnosed and treated, because their cancers would have never caused them harm had the cancers not been detected by screening (Djulbegovic et al., 2010). Quite a few of these men will have serious side effects, such as impotence or incontinence, as a consequence of the treatment. About a quarter of the men who have a radical prostatectomy will have such symptoms (USPSTF, 2011). Elevated risks of impotence and bowel dysfunction follow radiation treatment, too. Thus, in this subgroup of 20 men, there will be about 5 who will have to live with such consequences, which they would not have experienced had they not undergone treatment. Moreover, these men faced a small risk of dying from the surgery: About 5 out of every 1,000 men (which equals 0.1 in this subgroup of 20 men) will die within 30 days of undergoing a radical prostatectomy. Yet these 20 men do not know that the screening did not improve their eventual outcome, so they mistakenly believe that they were saved by the PSA test. Even those with serious side effects will probably believe that the side effects are worth it, even though the screening, with subsequent treatment, only caused them harm with no benefit. Thus, they are angry about the task force's recommendation to forgo the PSA test. Their families also think that the prostate cancer is the reason why they had to undergo treatment, and do not know that the treatments were actually unnecessary and only harmful. These families might be very unhappy with the PSA screening test if they did know.

Auditorium Screened also contains a third group of about 180 men who received a false-positive test result and had an unnecessary biopsy. Even if a man has only a biopsy with no further medical treatment, there is the risk of harm. Seven percent of the men who have prostate biopsies (i.e., 12.6 of these 180) have to be hospitalized because of infections and other complications (Loeb, Carter, Berndt, Ricker, & Schaeffer, 2011). These 180 men will probably not think that they have been saved, but will be relieved that their test result was only a false alarm. Probably only few of them, if any, will blame the test, however, for having produced a false alarm in the first place.

So, in Auditorium Screened, there are a lot of men who think that the PSA screening provided more benefit than harm, when in fact the opposite is true. We know that it is true because (a) the number of prostate-cancer deaths is the same in the two auditoriums, but (b) Auditorium Screened contains many men with serious side effects that could have been avoided had the men not been screened with a test that has a high false-positive rate.

We cannot blame the men in Auditorium Screened for their beliefs that they are healthier than the men in Auditorium Not Screened and that death from prostate cancer has reduced their numbers to a lesser degree. In the real world, there are no such adjacent auditoriums whose proximity allows an easy comparison of the health status of the two groups. Of course, epidemiological studies allow for such comparisons. When the men in Auditorium Screened rely on their personal experience, they are not aware of the counterfactuals. They do not know about the fate of the men who were not screened or of the men who made different treatment choices than they did. An individual man might only know that he was screened, had a biopsy, was treated, and now does not have prostate cancer. He thinks the PSA test saved his life, and he places an unkind comment about the task force members' professional competence on USPSTF's Web site.

What if there were adjacent auditoriums, one containing men who had been screened and one containing men who had not been screened? Even under such ideal circumstances allowing for straightforward comparison of outcomes, the research we describe next (in discussing Factor 3) suggests that men in Auditorium Screened still might evaluate such data improperly.

Before turning to Factor 3, we briefly add two caveats about our example of the two auditoriums. First, in the United States, the majority of men above the age of 50 have had a PSA test (Ross, Berkowitz, & Ekwueme, 2008). Indeed, 95% of male urologists and 78% of male primary-care physicians age 50 and older have had a PSA test (Chan, Barry, Vernon, & Ahn, 2006). Contrary to our example, the number of men in U.S. Auditorium Screened is in fact much larger than the number in U.S. Auditorium Not Screened. Thus, the relative number of men whose PSA screening has led to more harm than benefit is actually larger than our example suggests. Second, we have used averages in our calculations, so the numbers might vary depending on such factors as whether any subgroups of men have particular risk factors, such as a family history of prostate cancer.

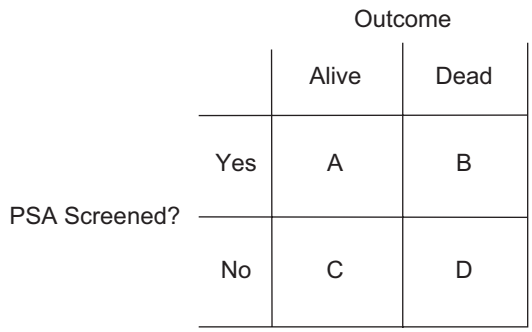

Fig. 1. Matrix showing the categories of data needed to determine whether prostate-specific antigen (PSA) screening is effective in reducing mortality.

Factor 3: Improper Data Evaluation

Consider the 2 × 2 matrix depicted in Figure 1. In evaluating the relation between the PSA test and mortality, laypersons generally pay the most attention to the data in cell A, and they pay more attention to the data in the top row than to the data in the bottom row (Arkes & Harkness, 1983; Shaklee & Tucker, 1980; Ward & Jenkins, 1965). This implies that people's evaluation of the PSA test would be most influenced by the large number of men they know who have gotten a PSA test and who are still alive. These men are in cell A, whose data contribute to the conclusion that the test is beneficial. This positive view of the PSA test is exacerbated by the fact that most men who receive PSA screening do not realize that their outcome would have been the same had they not been screened. In other words, the proportion of the data in the top row that belongs in cell A is nearly identical to the proportion of the data in the bottom row that belongs in cell C; the overall mortality rates for screened and not-screened men is the same (Schröder et al., 2009, 2012).

All four cells of Figure 1 are needed to ascertain whether screening is more effective than no screening with regard to mortality. Although research results suggest that many people do not think the evidence in cells C and D is important in evaluating the effectiveness of the PSA test (e.g., Arkes & Harkness, 1983), the people on the USPSTF think that this evidence is highly relevant. The people on the task force know that cells C and D comprise the control group's data. Many laypersons do not understand that a control group is needed in order to evaluate the effectiveness of a medical intervention or test. By attending primarily to cell A and by giving insufficient attention to cells C and D, many members of the general public will not comprehend a central basis for the task force's negative opinion of the PSA test.

Factor 4: Screening Tests for Low-Base-Rate Events

Meehl and Rosen (1955) pointed out long ago that screening tests that are highly informative for a select subgroup are not necessarily well suited for screening the general population.

Table 1. Results of a Prostate-Specific Antigen (PSA) Test in a Population of 1,000 Men With a 50% Base Rate of Cancer

PSA test result	Men with cancer	Men without cancer
Positive	105	30
Negative	395	470
Total	500	500

Note: The table shows the number of men in each indicated category. The positive predictive value of the PSA test in this population is 78% (i.e., 105/(105 + 30) × 100).

With a cutoff point of 4 ng/ml, the PSA test is reported to have a sensitivity of approximately 21% and a specificity of approximately 94% (Thompson et al., 2005). That means the PSA test will correctly classify 21% of the men with prostate cancer and 94% of the men who do not have prostate cancer. Conversely, the test will miss about 79% of the men who actually have prostate cancer, and raise a false alarm in 6% of the men who actually do not have prostate cancer. Suppose that this test is given to 1,000 patients at a urology clinic who have symptoms diagnostic of prostate cancer. Perhaps 50% of these men truly have prostate cancer. Table 1 depicts this situation. Of the 135 men who test positive, 105 actually have prostate cancer. Thus, the positive predictive value of the PSA test in this situation is approximately 78% (i.e., 105/135 × 100).

Now let us consider using the same screening test when prostate cancer is a low-base-rate event. If the test is used to screen the general population of males, the base rate will be much lower than the 50% depicted in Table 1. In one of the largest studies of prostate-cancer screening (Schröder et al., 2009), the base rate of prostate cancer was only 6.3%. Table 2 depicts this low-base-rate situation. Now the positive predictive value is down to 19%. This means that 81% of the positive test results are false positives! Many of the men with these positive test results will have biopsies with associated morbidity (Loeb et al., 2011). Some of these men will have prostatectomies, radiation therapy, or hormone therapy with their associated morbidities (Chou et al., 2011). The problem is that a test with modest sensitivity or specificity is not appropriate for screening the general population unless the cost of false positives or false negatives is very low.

Table 2. Results of a Prostate-Specific Antigen (PSA) Test in a Population of 1,000 Men With a 6.3% Base Rate of Cancer

PSA test result	Men with cancer	Men without cancer
Positive	13	56
Negative	50	881
Total	63	937

Note: The table shows the number of men in each indicated category. The positive predictive value of the PSA test in this population is 19% (i.e., 13/(13 + 56) × 100).

Prophylactic Measures

One of us has served on government committees whose reports are not welcomed (Arkes, 2003). In part on the basis of that experience, we suggest using what has been learned from psychological research to help such committees communicate more successfully with the public and the U.S. Congress.

One way of effectively communicating clinical evidence to the public is to use facts boxes, which are simple tabular representations of the benefits and harms of particular treatments. Schwartz, Woloshin, and Welch (2007, 2009) developed the concept of drug facts boxes and successfully tested their use with laypeople. Inspired by Woloshin and Schwartz (2009), the Harding Center for Risk Literacy at the Max Planck Institute for Human Development has designed a facts box about the benefits and harms of PSA screening (Fig. 2).

Another very effective means of informing the public about clinical evidence is to use visual displays. Let us consider again the study by Fagerlin et al. (2005). In addition to the groups we have already described, other groups of subjects saw the base rates of success of bypass surgery and angioplasty depicted with pictographs. Thus, for angioplasty, a matrix of 100 small figures was presented, with half of them colored to represent those cured of angina. For bypass surgery, the matrix had 75 figures colored to represent those cured of angina. For these two groups of subjects, the differing proportions of anecdotal successful and unsuccessful outcomes did not have a significant influence on treatment choices. Several other studies suggest that such pictorial depictions significantly increase understanding of statistical data and counteract the influence of anecdotal data (Fischhoff, Brewer, & Downs, 2011; Hawley et al., 2008).[1] We suggest that the USPSTF report would have been more successfully understood had such pictorial representations been used. In its current form, the report contains a very large number of statistics, which will overwhelm the general public. Research has shown that numeracy among the American public is shockingly low (Galesic & Garcia-Retamero, 2010; Lipkus, Samsa, & Rimer, 2001; Schwartz, Woloshin, Black, & Welch, 1997; Woloshin, Schwartz, & Welch, 2005). Pictorial representations help, and Figure 3 illustrates what an icon array for PSA screening could look like.

It is instructive that the public reaction to the USPSTF report can be understood in part by referring to psychological research. It would be particularly helpful if such research could be used to educate the public and elevate the level of civic discussion.

Prostate Cancer Early Detection

HARDING CENTER FOR RISK LITERACY

by PSA screening and digital-rectal examination.
Numbers are for men aged 50 years or older, not participating vs. participating in screening for 10 years.

	1,000 men without screening	1,000 men with screening
Benefits		
How many men died from prostate cancer?	8*	8
How many men died from any cause?	200	200
Harms		
How many men were diagnosed and treated** for prostate cancer unnecessarily?	—	20
How many men without cancer got a false alarm and a biopsy?	—	180

* This means that about 8 out of 1,000 men (50+ years of age) without screening died from prostate cancer within 10 years.

** With prostate removal or radiation therapy, which can lead to incontinence or impotence.

Fig. 2. Facts box illustrating the benefits (or lack thereof) and harms of prostate-specific antigen (PSA) screening for men age 50 and older. The underlying epidemiological data are taken from Djulbegovic et al. (2010). Note that the numbers are not meant to be the final verdict on PSA screening, but rather serve to illustrate the order of magnitude of the effects. Copyright 2012 by the Harding Center for Risk Literacy.

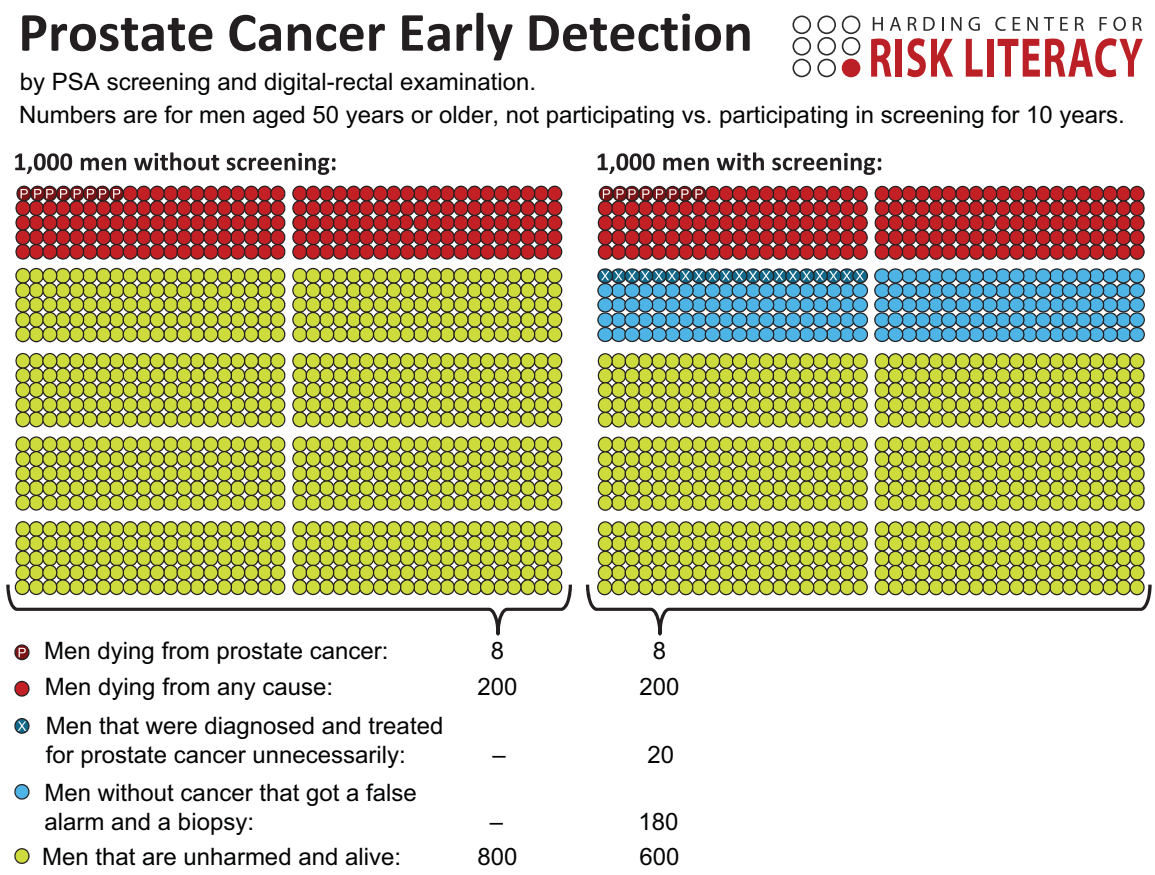

Fig. 3. Icon array illustrating the benefits (or lack thereof) and harms of prostate-specific antigen (PSA) screening for men age 50 and older. The underlying epidemiological data are taken from Djulbegovic et al. (2010). Note that the numbers are not meant to be the final verdict on PSA screening, but rather serve to illustrate the order of magnitude of the effects. Copyright 2012 by the Harding Center for Risk Literacy.

Acknowledgments

We thank Birgit Silberhorn and the members of the Harding Center for Risk Literacy for their help with the facts box and the icon array.

Declaration of Conflicting Interests

The authors declared that they had no conflicts of interest with respect to their authorship or the publication of this article.

Note

1. Pictographs additionally reduce several other biases, such as denominator neglect (Garcia-Retamero, Galesic, & Gigerenzer, 2010), duration neglect (Liersch & McKenzie, 2009), and framing effects (Garcia-Retamero & Cokely, 2011; Garcia-Retamero & Galesic, 2010).

References

Andriole, G. L., Crawford, E. D., Grubb, R. L., III, Buys, S. S., Chia, D., Church, T. R., . . . Berg, C. D. (2009). Mortality results from a randomized prostate-cancer screening trial. *New England Journal of Medicine*, *360*, 1310–1319.

Arkes, H. R. (2003). The nonuse of psychological research at two federal agencies. *Psychological Science*, *14*, 1–6.

Arkes, H. R., & Harkness, A. R. (1983). Estimates of contingency between two dichotomous variables. *Journal of Experimental Psychology: General*, *112*, 117–135.

Campanile, C. (2011, October 8). Rudy rips "prostate panel" a new one. *New York Post*. Retrieved from http://www.nypost.com/p/news/local/rudy_rips_prostate_test_panel_new_2wq1v7rrw1IYI8IRaJzTXI

Chan, E. C., Barry, M. J., Vernon, S. W., & Ahn, C. (2006). Physicians and their personal prostate cancer-screening practices with prostate-specific antigen: A national survey. *Journal of General Internal Medicine*, *21*, 257–259.

Chou, R., Croswell, J. M., Dana, T., Bougatsos, C., Blazina, I., Fu, R., . . . Lin, K. (2011). Screening for prostate cancer: A review of the evidence for the U.S. Preventive Services Task Force. *Annals of Internal Medicine*, *155*, 762–771.

Christensen, D. M. (2011). [Letter to Kathleen Sebelius, Secretary of U.S. Department of Health and Human Services]. Retrieved from http://donnachristensen.house.gov/sites/donnachristensen.house.gov/files/documents/Prostate%20Cancer%20Screening%20Recs.Nov.2011.PDF

de Wit, J. B. F., Das, E., & Vet, R. (2008). What works best: Objective statistics or a personal testimonial? An assessment of the persuasive

effects of different types of message evidence on risk perception. *Health Psychology*, *27*, 110–115.

Djulbegovic, M., Beyth, R. J., Neuberger, M. M., Stoffs, T. L., Vieweg, J., Djulbegovic, B., & Dahm, P. (2010). Screening for prostate cancer: Systematic review and meta-analysis of randomised controlled trials. *British Medical Journal*, *341*, c4543. Retrieved from http://www.bmj.com/content/341/bmj.c4543

Fagerlin, A., Wang, C., & Ubel, P. A. (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.

Fischhoff, B., Brewer, N. T., & Downs, J. T. (Eds.). (2011). *Communicating risks and benefits: An evidence-based user's guide*. Silver Spring, MD: U.S. Food and Drug Administration.

Galesic, M., & Garcia-Retamero, R. (2010). Statistical numeracy for health: A cross-cultural comparison with probabilistic national samples. *Archives of Internal Medicine*, *170*, 462–468.

Garcia-Retamero, R., & Cokely, E. T. (2011). Effective communication of risks to young adults: Using message framing and visual aids to increase condom use and STD screening. *Journal of Experimental Psychology: Applied*, *17*, 270–287. doi:10.1037/a0023677

Garcia-Retamero, R., & Galesic, M. (2010). How to reduce the effect of framing on messages about health. *Journal of General Internal Medicine*, *25*, 1323–1329.

Garcia-Retamero, R., Galesic, M., & Gigerenzer, G. (2010). Do icon arrays help reduce denominator neglect? *Medical Decision Making*, *30*, 672–684. doi:10.1177/0272989X10369000

Gigerenzer, G., Mata, J., & Frank, R. (2009). Public knowledge of benefits of breast and prostate cancer screening in Europe. *Journal of the National Cancer Institute*, *101*, 1216–1220.

Hamill, R., Wilson, T., & Nisbett, R. E. (1980). Insensitivity to sample size: Generalizing from atypical cases. *Journal of Personality and Social Psychology*, *39*, 578–589.

Hawley, S. T., Zikmund-Fisher, B., Ubel, P., Jankovic, A., Lucas, T., & Fagerlin, A. (2008). The impact of the format of graphical presentation on health-related knowledge and treatment choices. *Patient Education and Counseling*, *73*, 448–455.

Ilic, D., O'Connor, D., Green, S., & Wilt, T. J. (2011). Screening for prostate cancer: An updated Cochrane systematic review. *British Journal of Urology International*, *107*, 882–891.

Jenni, K. E., & Loewenstein, G. (1997). Explaining the identifiable victim effect. *Journal of Risk and Uncertainty*, *14*, 235–257.

Kogut, T., & Ritov, I. (2005). The "identified victim" effect: An identified group, or just a single individual? *Journal of Behavioral Decision Making*, *18*, 157–167.

Liersch, M. J., & McKenzie, C. R. M. (2009). Duration neglect by numbers—and its elimination by graphs. *Organizational Behavior and Human Decision Processes*, *108*, 303–314.

Lipkus, I. M., Samsa, G., & Rimer, B. K. (2001). General performance on a numeracy scale among highly educated samples. *Medical Decision Making*, *21*, 37–44.

Loeb, S., Carter, H. B., Berndt, S. I., Ricker, W., & Schaeffer, E. M. (2011). Complications after prostate biopsy: Data from SEER-Medicare. *Journal of Urology*, *186*, 1830–1834.

Meehl, P. E., & Rosen, A. (1955). Antecedent probability and the efficiency of psychometric signs, patterns, or cutting scores. *Psychological Bulletin*, *52*, 194–216.

National Institutes of Health Consensus Development Panel. (1997). National Institutes of Health Consensus Development Conference Statement: Breast cancer screening for women ages 40-49, January 21-23, 1997. *Journal of the National Cancer Institute*, *89*, 1015–1026.

Ross, L. E., Berkowitz, Z., & Ekwueme, D. U. (2008). Use of the prostate specific antigen test among U.S. men: Findings from the 2005 National Health Interview Survey. *Cancer Epidemiology, Biomarkers & Prevention*, *17*, 636–644.

Sandblom, G., Varenhorst, E., Rosell, J., Löfman, O., & Carlsson, P. (2011). Randomised prostate cancer screening trial: 20 year follow-up. *British Medical Journal*, *342*, d1539. Retrieved from http://www.bmj.com/content/342/bmj.d1539

Schröder, F. H., Hugosson, J., Roobol, M. J., Tammela, T. L. J., Ciatto, S., Nelen, V., . . . Auvinen, A. (2009). Screening and prostate-cancer mortality in a randomized European study. *New England Journal of Medicine*, *360*, 1320–1328.

Schröder, F. H., Hugosson, J., Roobol, M. J., Tammela, T. L. J., Ciatto, S., Nelen, V., . . . Auvinen, A. (2012). Prostate-cancer mortality at 11 years of follow-up. *New England Journal of Medicine*, *366*, 981–990.

Schwartz, L. M., Woloshin, S., Black, W. C., & Welch, G. H. (1997). The role of numeracy in understanding the benefit of screening mammography. *Annals of Internal Medicine*, *127*, 966–971.

Schwartz, L. M., Woloshin, S., & Welch, H. G. (2007). The drug facts box: Providing consumers with simple tabular data on drug benefit and harm. *Medical Decision Making*, *27*, 655–662.

Schwartz, L. M., Woloshin, S., & Welch, H. G. (2009). Using a drug facts box to communicate drug benefits and harms. *Annals of Internal Medicine*, *150*, 516–527.

Shaklee, H., & Tucker, D. (1980). A rule analysis of judgments of covariation between events. *Memory & Cognition*, *8*, 459–467.

Slovic, P. (2007). "If I look at the mass I will never act": Psychic numbing and genocide. *Judgment and Decision Making*, *2*, 79–95.

Thompson, I. M., Ankerst, D. P., Chi, C., Lucia, M. S., Goodman, P. J., Crowley, J. J., . . . Coltman, C. A. (2005). Operating characteristics of prostate-specific antigen in men with an initial PSA level of 3.0 ng/mL or lower. *Journal of the American Medical Association*, *294*, 66–70.

Ubel, P. A., Jepson, C., & Baron, J. (2001). The inclusion of patient testimonials in decision aids: Effects on treatment choices. *Medical Decision Making*, *21*, 60–68.

U.S. Preventive Services Task Force. (2011). *Screening for prostate cancer: U.S. Preventive Services Task Force recommendation statement draft: Summary of recommendation and evidence*. Retrieved from http://www.uspreventiveservicestaskforce.org/uspstf12/prostate/draftrecprostate.htm

Ward, W. C., & Jenkins, H. M. (1965). The display of information and the judgment of contingency. *Canadian Journal of Psychology*, *19*, 231–241.

Woloshin, S., & Schwartz, L. M. (2009). Numbers needed to decide. *Journal of the National Cancer Institute*, *101*, 1163–1165.

Woloshin, S., Schwartz, L. M., & Welch, H. G. (2005). Patients and medical statistics: Interest, confidence, and ability. *Journal of General Internal Medicine*, *20*, 996–1000.