ELSEVIER

Patient Education and Counseling 66 (2007) 100–107

Patient Education and Counseling

www.elsevier.com/locate/pateducou

Mortality versus survival graphs: Improving temporal consistency in perceptions of treatment effectiveness☆

Brian J. Zikmund-Fisher [a,b,c,*], Angela Fagerlin [a,b,c], Peter A. Ubel [a,b,c,d]

[a] *VA Health Services Research & Development Center for Practice Management and Outcomes Research, VA Ann Arbor Healthcare System, Ann Arbor, MI, United States*
[b] *Division of General Internal Medicine, University of Michigan, Ann Arbor, MI, United States*
[c] *Center for Behavioral and Decision Sciences in Medicine, Ann Arbor, MI, United States*
[d] *Department of Psychology, University of Michigan, Ann Arbor, MI, United States*

Received 3 March 2006; received in revised form 25 October 2006; accepted 30 October 2006

Abstract

Objective: Previous research has demonstrated that people perceive treatments as less effective when survival graphs show fewer years of data versus more data. We tested whether using mortality graphs would reduce this temporal inconsistency bias.
Methods: A demographically diverse sample of 1461 Internet users read about a hypothetical disease and then were randomized to view either survival or mortality graphs that showed either 5 years of data or 15 years of treatment outcomes data. Participants identified the most effective treatment, provided ratings comparing the effectiveness of two treatments, and answered comprehension questions.
Results: Treatment effectiveness ratings varied significantly between respondents seeing the 5 year and 15 year survival graphs even though the relative risk reduction was the same in both cases. This variation was significantly reduced in the mortality graph conditions. Responses on comprehension measures were mixed: viewers of mortality graphs were less able to identify which treatment was more effective but better able to correctly report individual data points.
Conclusions: Perceptions of treatment effectiveness appear more temporally consistent with mortality graphs than with survival graphs.
Practice implications: All line-based risk graphics (whether framed in survival or mortality terms) should highlight duration information to facilitate improved comprehension of treatment effectiveness.

Keywords: Risk communication; Visual displays; Cognitive biases

1. Introduction

Survival graphs are commonly used to visually display treatment effectiveness information in reports of randomized clinical trials [1,2] and have been shown to be useful in communicating complex treatment problems to patients [3–9]. In such graphs, differential treatment effectiveness is visually illustrated by the area between a treatment curve and a baseline or placebo curve. Recent research, however, has shown that perceptions of survival curves are not unbiased [5,10,11]. In particular, there is a visual illusion caused by the fact that the area between two curves varies with the length of time displayed [12]. When less time is displayed, fewer people die regardless of treatment choice, the survival curves are flatter, and hence the area between two curves is visually smaller even if the relative effectiveness of treatments is unchanged. This illusion leads to a temporal inconsistency bias: people fail to adjust their perception of the visual differences between curves to account for the how many years of data are displayed [13]. As a result, when directly asked to rate the effectiveness of different possible treatments, people perceive larger differences in effectiveness when survival graphs show fewer years of data versus more data, even

☆ This work was presented at the annual meeting of the Society for Medical Decision Making, October 16, 2006.

* Corresponding author at: Center for Behavioral and Decision Sciences in Medicine, 300 North Ingalls Building, Rm. 7C27, Ann Arbor, MI 48109-0429, United States. Tel.: +1 734 936 9179; fax: +1 734 936 8944.
E-mail address: bzikmund@umich.edu (B.J. Zikmund-Fisher).

0738-3991/$ – see front matter © 2006 Published by Elsevier Ireland Ltd.
doi:10.1016/j.pec.2006.10.013

when annual mortality risks and relative risk reductions are constant over time.

One explanation for this bias is that people adjust their perceptions of treatment effectiveness to account for the statistical reliability of the observed differences. In graphs showing only a few years of data, fewer people have died and hence the differences in survival rates are likely to be less significant than at later points in time. Even though the effect size may be constant over time, viewers of survival graphs may be discounting their perceptions of treatment effectiveness to account for difference in their confidence about this result.

Another potential explanation, however, derives from the fact that survival graphs violate one of the basic standards of graphic design. Interpreting a risk graphic involves identifying the data point of interest and comparing its visual position in the graph to a comparison standard. In virtually every other type of line graph, comparisons are made to the ''L-shaped'' framework of the *X*- and *Y*-axes, focused at the origin [14]. Thus, the height of a data point is compared to bottom of the graph. With survival graphs, however, the height of a data point should be compared to the *top* of the graph, i.e., the level of complete survival. This means that, as shown in left-hand graph in Fig. 1, evaluating the risk reduction achieved by treatment requires visually measuring the vertical distance between two curves (e.g., the solid vertical line labeled ''*Y*'') and comparing those distances to the distance to 100% survival (e.g., the dotted vertical line labeled ''*Z*''). As this task is dissimilar from most other graph interpretation tasks people are familiar with, often confusion and/or misinterpretation results.

If the bias is caused by confusion about the visual standards of comparison in survival graphs, however, a potential solution to this problem involves the use of mortality graphs, which display the identical information as survival curves but use the *X*-axis (0% mortality) as the natural standard of comparison. As shown in right-hand graph in Fig. 1, interpretation of treatment effectiveness as displayed in mortality graphs still involves identifying the same vertical distances between curves (e.g., Y'), but that distance is now compared to the vertical distance to the bottom of the graph (Z'). As comparisons of graph data points to the *X*-axis are more familiar to most readers, we hypothesized that people would be better able to use the total mortality (as displayed in height Z') to calibrate their expectations regarding how the visual space between curves (Y') relates to relative treatment effectiveness. As a result, we expected respondents to be more temporally consistent in their interpretations of treatment effectiveness when using mortality graphs instead of survival graphs. Note, however, that the use of mortality graphs should make no difference if the temporal inconsistency is caused by people adjusting for the statistical reliability of the displayed results.

Nevertheless, there are two concerns about the use of mortality graphs instead of survival graphs. First, research by Armstrong et al. raises the possible concern that members of the general public may have worse accuracy when interpreting the same information in mortality graph form instead of as a survival graph [10]. Such comprehension difficulties could more than offset any consistency gains achieved by framing the data in terms of mortality. Second, mortality graphs are less commonly used in medical applications, and, as a result, people may have to expend greater cognitive effort when viewing or interpreting such graphics.

We sought to replicate the temporal inconsistency bias previously demonstrated in research of the general public's perceptions of survival graphs [13] and test whether presenting the equivalent data in a mortality graph format would reduce or eliminate the bias. We hypothesized that people viewing survival graphs would perceive different levels of treatment effectiveness between 15 year graphs and 5 year graphs, while people viewing mortality graphs would (correctly) perceive equal degrees of treatment effectiveness regardless of time period displayed.

2. Methods

2.1. Overview of study design

In an Internet-administered survey, participants read a description of the hypothetical health condition used in prior

Fig. 2. Experimental conditions.

research [13]. We randomized participants to receive either a survival curve graph showing 15 years' worth of data, an abbreviated survival graph showing only 5 years' worth of data, or 1 of 2 analogous mortality graphs and then assessed individuals' comprehension of the graphs and their perceptions of treatment effectiveness and disease seriousness. The name of the condition was also randomly varied, with half of subjects reading about a ''disease'' and the remainder reading about a ''cancer.''

2.2. Participants

Study participants were drawn from a panel of Internet users administered by Survey Sampling International (SSI) who voluntarily agreed to participate in research surveys. To ensure demographic diversity (but not perfect representativeness), e-mail invitations were sent to a stratified random sample weighted to offset expected differences in response rates (especially for African-Americans and Hispanic-Americans), with the goal of approximating the U.S. census in the final subject pool. To ensure a broad age distribution, we drew our sample from three distinct age groups: one-third each ages 21–40, 41–59, and 60 and older. Upon completion, participants were entered into a draw administered by SSI for cash prizes totaling US$ 10,000.

2.3. Intervention

Respondents read a brief description of a hypothetical condition, which was randomized to be named either Crawford's Disease or Crawford's Cancer to test whether the cancer label affected perceptions. (For simplicity, we only refer to Crawford's Disease from here on.) Respondents were not asked to imagine that they had the condition, only to evaluate information about the disease in question. They were informed that patients with this disease have a constant risk of dying each year but can lower that risk by taking one of two medications. Survey participants then were randomly assigned to view one of four graphs. Two graphs showed survival curves indicating the proportion of Crawford's Disease patients who took either Pill A, Pill B, or No Pills still surviving as time since diagnosis increases. The remaining two graphs showed the equivalent mortality data for each option (see Appendix A for the full text of the scenario and a sample graphic).

Fig. 2 summarizes the four graph conditions in this study, which comprise a 2×2 factorial design ($2 \times 2 \times 2$ including the disease label manipulation). We started with a survival graph showing 15 years of data as our base condition (Survival Graph A). The annual survival rates for the three survival curves were 96% (Pill A), 94.5% (Pill B), and 90% (No Pills), and risks were compounded annually, yielding exponential curve shapes. We then created an alternate, abbreviated graph (Survival Graph B) that displays only 5 years, with the *X*-axis rescaled accordingly. As noted in Table 1, doing so does not affect the annual risk levels or the relative risk reductions but does create a graph with flatter sloped curves and hence a smaller area between the treatment and No Pills curves. Furthermore, since survey participants were explicitly told that ''the chance of dying from Crawford's Disease is the same every year no matter how long you live,'' viewers of these two graphs received equivalent risk information, merely displayed for different periods of time. Next, we took both of these graphs and

Table 1
Characteristics of the graphs used

Characteristic	Type A graphs	Abbreviated Type B graphs
Amount of data displayed	15 Years	5 Years
Annual risk levels	Constant over time	Constant over time
Slope of survival/mortality curves	Steeper than Type B	Flatter than Type A
Relative risk reductions (RRR)	Constant over time	Constant over time

created corresponding mortality graphs by simply calculating 100 minus the percent surviving at each time point.

2.4. Outcome measures

After completing an initial assessment of disease seriousness for Crawford's Disease (not analyzed here), respondents answered a screening question which asked them to identify which pill was more effective (Pill A or Pill B). The next question was our primary dependent variable: a judgment as to how much more effective the respondent believed their selected treatment was, as compared to the alternate treatment. Thus, if a respondent stated (correctly) that Pill A was more effective, they next saw a question which asked them, ''How much *more effective* do you think Pill A is compared to Pill B for treating Crawford's Disease?'' Responses were made on an 11-point (0–10) scale, with endpoints labeled as ''Just as effective'' and ''Much more effective.''

After completing their assessment of effectiveness, respondents also answered a sequence of comprehension questions about the information presented in the graph. For example, respondents were asked to indicate how many people who took Pill A, out of 100, were dead after 12 years (4 years if respondents saw the abbreviated 5-year ''B'' graphs). To account for the difficulty of this task, responses were coded as correctly answered if they were within ± 4 people of the true answer.

Participants also completed the subjective numeracy scale (SNS) [15,16] to control for variations in ability with and preferences for numerical information, as well as standard demographics questions. The research design and measures received Institutional Review Board exempt status approval.

2.5. Hypotheses

Because, in principle, reading survival and mortality graphs requires exactly the same skill set, we did not hypothesize any differences in respondents' ability to correctly identify the most effective treatment or to report particular numbers from the graphs viewed.

Consistent with previous research on temporal inconsistency in interpretation of survival graphs, however, we hypothesized that people's perceptions of treatment effectiveness would be primarily driven by the size of the visual space between the Pill A and Pill B curves rather than by the actual risk levels shown. This ''separation'' between curves is then compared to the baseline (either 100% survival or 0% mortality) to determine whether the magnitude of this difference is practically significant. We thus hypothesized that respondents would see a larger difference in effectiveness between Pill A and Pill B with Survival Graph A than with the abbreviated Survival Graph B. Because mortality graphs conform to the usual expectations that the standard of comparison is the *X*-axis, however, we hypothesized that we would observe a significant reduction or elimination of temporal inconsistency. Specifically, we anticipated that perceptions of the difference in treatment effectiveness between Pill A and Pill B would not significantly differ between the two mortality graph conditions.

2.6. Statistical analysis

We used Chi-squared tests to compare whether more people correctly identified the most effective treatment and correctly answered our comprehension questions when they viewed survival graphs versus mortality graphs. We then utilized *t*-tests to determine whether respondents' ratings of both ''B'' graph conditions differed from ratings of the corresponding base case graphs. However, to test whether the use of mortality graphs mitigated any observed temporal inconsistency, we also conducted an ANOVA analysis which included dummy variables for graph frame (survival or mortality), disease label (Crawford's Disease or Cancer), and graph type (A or B), as well as a frame $\times$ type interaction term (to test if the effect of graph type differed between survival and mortality graphs) and respondent's average SNS score and respondent age as continuous covariates. All analyses were performed using STATA Version 8.

3. Results

This study was part of a larger Internet survey effort in which people were randomly assigned to different tasks. Overall, 6.6% of people receiving e-mail invitations clicked the embedded link to see the survey cover page. This rate varied significantly by age; response rates were 9.4% for invitees age 40 and above but only 3.9% for those under age 40. A total of 1704 participants were randomized to the conditions reported here. While some dropout occurred as people progressed through other parts of the survey, 1461 people (86%) provided useable data. Sample mean age was 50 (range 21–89), 49% were males, and, of the 1339 who

reported racial and ethnic background information, 82% described themselves as Caucasian, 13% Hispanic (any race), 11% African-American, 2% Asian-American, and 5% multi-ethnic or other race. We observed a wide range of educational achievement, with 37% having completed a Bachelor's or higher degree but also 20% with only a High School education or less. Compared to non-respondents (using information from SSI), participants were older (mean age 51 versus 41) and more educated (37% with Bachelor's or higher versus 30%), due to the substantial over-sampling of younger adults needed to offset their lower response rates. Most importantly, there were no significant variations in sample characteristics across the randomized experimental groups.

Contrary to expectations, we observed a significant difference in people's ability to correctly identify the most effective treatment. While 94% of participants who viewed survival graphs accurately reported that Pill A was most effective, this percentage drops to 85% among those who viewed mortality curves.

Since the following question regarding perceptions of treatment was tailored based on the response to the screening question, respondents who failed to correctly identify Pill A as more effective were excluded from all subsequent analyses. However, to check whether this selection process influenced our results, we also repeated all analyses using all subjects. (Doing so is appropriate since, even if participants misidentified the most effective treatment, they all were asked to evaluate the magnitude of the difference between the Pill A and Pill B curves.) All results were qualitatively similar, so we only report the restricted analyses here.

Among those correctly interpreting survival graphs, we again saw a distinct pattern of temporal inconsistency. Participants who viewed an abbreviated graph showing only 5 years of data (Survival Graph B) perceived a significantly smaller difference in treatment effectiveness than those who viewed the base 15 year graph (Survival Graph A) (5.63 versus 6.73, $t = -6.07$, $p < 0.001$). This difference occurred despite the fact that these graphs both illustrated identical risk levels and hence identical relative risk reductions due to treatment. Consistent with our expectations, however, subjects who had correctly interpreted the corresponding mortality graphs showed a much reduced pattern of temporal inconsistency (6.02 versus 6.42), although the difference is still significant ($t = -2.13$, $p = 0.034$).

Table 2
ANOVA results

Model parameter	F	p-Value
Abbreviated Type B graph (vs. 15 year Type A)	34.66	<0.001
Mortality graph (vs. survival)	0.03	0.869
"Crawford's Cancer" (vs. "Disease")	1.04	0.307
Subjective numeracy score (continuous variable)	0.24	0.624
Participant age (continuous variable)	11.75	<0.001
Graph type × mortality frame interaction	7.44	0.006

Note: Sample restricted to only those subjects correctly answering the screening question ($N = 1135$). Dependent variable: how much more effective Pill A is vs. Pill B.

To test whether graph frame (survival or mortality) interacted with graph type (A or B), we conducted an ANOVA analysis (Table 2) that controlled individual differences in numerical ability in addition to our three experimental factors. Consistent with the bivariate analyses above, we find a significant main effect of graph type: respondents perceived smaller differences in treatment effectiveness between Pill A and Pill B with the abbreviated Type B graphs than with the original 15 year Type A graphs ($F = 34.66$, $p < 0.001$). We also found a significant main effect of age: older participants perceived greater differences in treatment effectiveness ($F = 11.75$, $p < 0.001$). There was no main effect of framing the graph in terms of survival versus mortality ($F = 0.03$, $p = 0.87$), and we found no significant effect of either disease label or numerical ability. However, we did find a clear and significant graph frame × type interaction to support our hypothesis that temporal inconsistency biases would be reduced when subjects viewed mortality graphs instead of survival graphs ($F = 7.44$, $p = 0.006$). An expanded ANOVA model (not reported) showed no other significant two-way or three-way interactions.

As noted above, fewer people were able to correctly identify the most effective treatment when viewing mortality

Table 3
Graph comprehension results

	Survival graphs ($n = 710$)	Mortality graphs ($n = 751$)	Significance (p)
Screening question			
% Correctly identifying Pill A as more effective	93.9	84.8	<0.001
Detailed comprehension questions[a]			
% Correctly identifying ...			
No. who took Pill A dead after 4/12 years?	61.5	80.1	<0.001
No. who took No Pills dead after 3/9 years?	54.4	72.7	<0.001
No. who took Pill B alive after 5/15 years?	69.0	72.7	0.163
No. who took Pill B who died between 1 and 5 years/between 5 and 15 years?	37.5	38.1	0.835
If someone survived 3/9 years, which pill gives the best chance of surviving another year	95.7	91.5	0.004

Note: [a] Among people correctly identifying Pill A as more effective than Pill B.

graphs versus survival graphs. However, among those subjects who correctly answered this screening question, we found that comprehension of mortality graphs was generally at least as good, and often better, then comprehension of survival graphs. Table 3 reports the percentage of subjects answering each comprehension question, broken down by graph frame. As you can see, respondents who viewed mortality graphs scored better on two questions, roughly equal on two additional questions, and worse on only one question, another question regarding which treatment reduced risk the most.

4. Discussion and conclusion

4.1. Discussion

Our results illustrate that survival graphs and mortality graphs each have strengths and weaknesses. On the one hand, survival curves appear to be most efficacious at helping people to identify which treatment was most effective. On the other hand, mortality graphs yielded comprehension levels at least as good as and often better than survival graphs on all other types of questions. More importantly, we also found that mortality graphs reduce temporal inconsistency bias: while people viewing 15 year survival graphs perceived significantly greater differences in treatment effectiveness between Pill A and Pill B than did people viewing abbreviated 5 year curves (which had flatter slopes and smaller areas between curves), the time-period effect was significantly smaller for the respondents viewing each type of mortality graph.

Why did mortality graphs reduce temporal inconsistencies yet lead to greater difficulty in identifying the most effective treatment? We believe that both of these results stem from the features of mortality graphs that are similar to and dissimilar to other types of graphs. Like most other types of line graphs, mortality graphs have the *X*-axis as the standard of comparison for all data points (as opposed to 100% survival, which is the comparison standard for survival graphs). As a result, viewers of mortality graphs can use the height of each curve (compared to the *X*-axis) to calibrate their perceptions of how significant the ''white space'' between two curves might be and adjust for differences caused by the number of years of data displayed. However, because mortality graphs show increasing levels of a *negative* outcome (death), the optimal treatment is identified by the *lowest* line on the graph, not the highest line. The idea that ''less is better'' is somewhat counter-intuitive, which may have led to higher observed error rates when subjects were asked to identify the most effective treatment.

Several factors limit the generalizability of our results. First, our Internet sample is not representative of the general population. However, the sample was demographically diverse, and our previous research using this panel has shown that Internet survey responses closely match those of representative samples [17]. Furthermore, our randomized experimental design ensured that any sample peculiarities occurred evenly across experimental groups. Second, the graphs studied here displayed risks which were constant over time, while many real-world risks either increase or decrease over time. People may be more or less able to interpret the data and develop accurate perceptions of treatment effectiveness when viewing curves showing risks which vary over time. Third, we do not know the exact size of the images viewed by participants. While all graphs were of the same size in terms of screen pixels, respondents may have had larger screens (creating larger image sizes), higher screen resolutions (resulting in smaller image sizes), or both. Secondary analyses of participants' screen resolutions (recorded directly from their Internet browsers) found that our primary result, the interaction of graph frame $\times$ type, appears much weaker among the minority of participants with screen resolutions of 800×600 or smaller. It is difficult to interpret this result, however, because lower screen resolution may indicate a smaller screen size, an older computer, or simply worse eyesight, among other reasons.

Fourth, our dependent variable, treatment ''effectiveness,'' could be interpreted in multiple ways. We asked respondents to consider the graphics and to make judgments about treatment effectiveness in the abstract. It is possible that some respondents were considering treatment effectiveness in the context of specific individual treatment decisions. As noted above, however, such variations among study participants should be randomly distributed across our experimental conditions. Still, the temporal inconsistency we document here could be stronger or weaker in other decision contexts. Lastly, using mortality graphs instead of survival graphs could lead patients to make different treatment decisions due to the well-documented impact of framing statistics in terms of death instead of survival [10,18,19].

4.2. Conclusion

There are many good reasons to use survival or mortality graphs to present treatment effectiveness information to health care providers and patients. Both types of graphs concisely represent not only individual risk levels but also the relative risk or benefit from different treatments over time. Yet, the very complexity and information density which makes these graphs so useful can also lead to confusion and misinterpretation, and each type of graph has different strengths and weaknesses. People appear to be better at identifying which treatment is optimal when viewing survival graphs instead of the corresponding mortality graphs. The results presented here suggest, however, that the more familiar structure of mortality graphs mitigates the temporal inconsistency bias previously demonstrated in interpretation of survival graphs, and using mortality graphs may improve people's ability to read off

specific data points from such graphs. We can only speculate whether optimal treatment identification would improve if mortality graphs became as commonly used as survival graphs currently are.

4.3. Practice implications

Primary practice clinicians must increasingly incorporate clinical research results into their practice, and the advent of the Internet increases availability of clinical research results to patients. Both of these factors imply that more and more people will be exposed to survival and mortality graphs as vehicles for conveying treatment effectiveness data. The present study reiterates the warning message of our previous research: the number of years of data provided in a survival curve can change beliefs about treatment effectiveness *by itself*. Presenting data in mortality graph format significantly reduces this unwanted effect, although special care will need to be taken to ensure that readers of mortality graphs recognize that the optimal treatment is shown by the *lowest* curve. We encourage greater consideration of mortality graphs in communications of treatment effectiveness and continue to recommend that all presentations of risk graphics (whether framed in survival or mortality terms) highlight duration information to facilitate comprehension.

Acknowledgements

Financial support for this study was provided by a grant from the National Institutes for Health (R01 CA87595). Dr. Zikmund-Fisher is supported by a career development award from the American Cancer Society (MRSG-06-130-01-CPPB), and Dr. Fagerlin is supported by an MREP award from the U.S. Department of Veterans Affairs. The funding agreements ensured the authors' independence in designing the study, interpreting the data, and publishing the report.

The authors would also like to acknowledge the web programming and research assistance provided by Todd Roberts.

Appendix A. Text of the hypothetical disease scenario

A.1. Crawford's Disease[Cancer]

Imagine there is a rare, serious disease[type of cancer] called Crawford's Disease[Cancer]. Without treatment, the chance of dying from this disease[type of cancer] is the same in any year no matter how long you live.

However, doctors agree that people who have Crawford's Disease[Cancer] have a better chance of surviving if they take one of two pills every day for the rest of their lives. These are called Pill A and Pill B, and they change your chance of dying from the disease[cancer] by a different amount. Once you have decided whether to take Pill A, Pill B, or No Pills, your chance of dying in any year remains stable for the rest of your life. The two pills cost about the same and are paid for by most insurance companies.

Pretend that 300 people were diagnosed with Crawford's Disease[Cancer] at year 0. A group of 100 people started taking Pill A, a second group of 100 people began taking Pill B, and a third group of 100 people did not take any pills. On the next page, you will see information about what happens to each of these groups.

This graph below is called a *survival[mortality] curve*. It shows the number of people with Crawford's Disease[Cancer] who survive[die] each year for the first 5 [15] years after being diagnosed with the disease[cancer]. The chance of dying from Crawford's Disease[Cancer] is the same every year is the same every year for the next 5 [15] years.

The blue line with the diamonds represents the people who took Pill A. The green line with the squares represents the people who took Pill B. The orange line with the triangles represents the people who did not take any pills. Each of the diamonds (Pill A), squares (Pill B), and triangles (No Pills) stand for the number of people who are alive[dead] at each point in time.

References

[1] Landolfi R, Marchioli R, Kutti J, Gisslinger H, Tognoni G, Patrono C, Barbui T, European Collaboration on Low-Dose Aspirin in Polycythemia Vera Investigators. Efficacy and safety of low-dose Aspirin in polycythemia vera. N Engl J Med 2004;350:114–24.

[2] Fisher B, Dignam J, Bryant J, Wolmark N. Five versus more than five years of Tamoxifen for lymph node-negative breast cancer: updated findings from the national surgical adjuvant breast and bowel project B-14 randomized trial. J Natl Cancer Inst 2001;93:684–90.

[3] Armstrong K, Weber B, Ubel P, Peters N, Holmes J, Schwartz JS. Individualized survival curves improve satisfaction with cancer risk management decisions in women with BRCA1/2 mutations. J Clin Oncol December 20, 2005;23:9319–28.

[4] Armstrong K, FitzGerald G, Schwartz JS, Ubel PA. Using survival curve comparisons to inform patient decision making can a practice exercise improve understanding? J Gen Intern Med 2001;16:482–5.

[5] Mazur DJ, Hickam DH. The effect of physician's explanations on patients' treatment preferences: five-year survival data. Med Decis Making 1994;14:255–8.

[6] Mazur DJ, Hickam DH. Patients' and physicians' interpretations of graphic data displays. Med Decis Making 1993;13:59–63.
[7] Mazur DJ, Hickam DH. Interpretation of graphic data by patients in a general medicine clinic. J Gen Intern Med 1990;5:402–5.
[8] Mazur DJ, Hickam DH. Five-year survival curves: how much data are enough for patient–physician decision making in general surgery? Eur J Surg 1996;162:101–4.
[9] Mazur DJ, Merz JF. Five-year survival data in surgical decision making: what aspects of graphical data influence patients' preferences? Theor Surg 1994;9:76–81.
[10] Armstrong K, Schwartz JS, Fitzgerald G, Putt M, Ubel PA. Effect of framing as gain versus loss on understanding and hypothetical treatment choices: survival and mortality curves. Med Decis Making 2002;22:76–83.
[11] Mazur DJ, Merz JF. How the manner of presentation of data influences older patients in determining their treatment preferences. J Am Geriatr Soc 1993;41:223–8.
[12] Lau EW, Ng GA. Visual illusions created by survival curves and the need to avoid potential misinterpretation. Med Decis Making 2002;22:238–44.
[13] Zikmund-Fisher BJ, Fagerlin A, Ubel PA. Inattention to the length of time displayed biases comprehension of survival graphs. Risk Anal 2005;25:589–95.
[14] Kosslyn SM. Elements of graph design. New York: W.H. Freeman and Company; 1994.
[15] Fagerlin A, Zikmund-Fisher BJ, Ubel PA, Jankovic A, Derry H, Smith DM. Measuring numeracy without a math test: development of the subjective numeracy scale (SNS). Med Decis Making, in press.
[16] Zikmund-Fisher BJ, Smith DM, Ubel PA, Fagerlin A. Validation of the subjective numeracy scale (SNS): effects of low numeracy on comprehension of risk communications and utility elicitations. Med Decis Making, in press.
[17] Lacey HP, Smith DM, Ubel PA. Hope I die before I get old: mispredicting happiness across the lifespan. J Happiness Stud 2006;7:167–82.
[18] Tversky A, Kahneman D. The framing of decisions and the psychology of choice. Science 1981;211:453–8.
[19] Moxey A, O'Connell D, McGettigan P, Henry D. Describing treatment effects to patients: how they are expressed makes a difference. J Gen Intern Med 2003;18:948–59.