Erschienen in: Medical Decision Making ; 34 (2014), 2. - S. 206-215

ORIGINAL ARTICLE

# How Do Physicians Provide Statistical Information about Antidepressants to Hypothetical Patients?

*Wolfgang Gaissmaier, PhD, Britta L. Anderson, PhD, Jay Schulkin, PhD*

***Background.*** *Little is known about how physicians provide statistical information to patients, which is important for informed consent.* ***Methods.*** *In a survey, obstetricians and gynecologists (*N = *142) received statistical information about the benefit and side effects of an antidepressant. They received information in various formats, including event rates (antidepressant v. placebo), absolute risks, and relative risks. Participants had to imagine 2 hypothetical patients, 1 for whom they believed the drug to be safe and effective and 1 for whom they did not, and select the information they would give those patients. We assessed whether the information they selected for each patient was complete, transparent, interpretable, or persuasive (i.e., to nudge patients toward a particular option) and compared physicians who gave both patients the same information with those who gave both patients different information.* ***Results.*** *A similar proportion of physicians (roughly 25% each) selected information that was 1) complete and transparent, 2) complete but not transparent, 3) not interpretable for the patient because necessary comparative information was missing, or 4) suited for nudging. Physicians who gave both patients the same information (61% of physicians) more often selected at least complete information, even if it was often not transparent. Physicians who gave both patients different information (39% of physicians), in contrast, more often selected information that was suited for nudging in line with the belief they were asked to imagine. A limitation is that scenarios were hypothetical.* ***Conclusions.*** *Most physicians did not provide complete and transparent information. Clinicians who presented consistent information to different patients tended to present complete information, whereas those who varied what information they chose to present appeared more prone to nudging.* ***Key words:*** *informed consent; risk communication; patient-physician communication; statistical literacy. (****Med Decis Making 2014;34:206–215****)*

Informed consent requires that patients understand the benefits and side effects of treatments. Because physicians are a primary source of information for patients during informed consent, it is important that physicians inform their patients both accurately and transparently. This study focuses on one aspect of physician-patient communication—how physicians communicate statistical information to patients.

It has been suggested that patients have more trust in numeric information compared with qualitative, verbal information. It has also been shown that patients are more comfortable and satisfied with numeric information compared with only verbal.[1,2] However, previous studies have documented that physicians do not present medical information in transparent formats[3] and that at times they even communicate incorrect information.[4,5] One study found that 60% of 160 gynecologists believed that the positive predictive value of a positive mammogram in asymptomatic women who participated in breast cancer screening was 80% or even 90% (even after receiving all relevant statistical information to calculate the positive predictive value), although the true answer is about 10%.[6]

Physicians' ability to provide accurate and transparent information is important because many

Received 3 August 2012 from Harding Center for Risk Literacy, Max Planck Institute for Human Development, Berlin, Germany (WG); and the American College of Obstetricians and Gynecologists, Washington, DC, USA (BLA, JS). This study is funded in part by grant UA6MC19010 through the U.S. Department of Health and Human Services, Health Resources and Services Administration, Maternal and Child Health Research Program. Revision accepted for publication 8 July 2013.

Address correspondence to Wolfgang Gaissmaier, Harding Center for Risk Literacy, Max Planck Institute for Human Development, Lentzeallee 94, 14195 Berlin, Germany; telephone: ++49-30-82406691; fax: ++49-30-82406394; e-mail: gaissmaier@mpib-berlin.mpg.de.


DOI: 10.1177/0272989X13501720

**Table 1** Information Provided to Physicians about the Antidepressant

| | No. out of 1000 Who Took Placebo | No. out of 1000 Who Took Antidepressant | Absolute Risk Change | Relative Risk Change |
|---|---|---|---|---|
| Benefit | | | | |
| Chance of improving | 261/1000 (26%) | 454/1000 (45%) | 19.3 percentage points | 74% |
| Side effects | | | | |
| Dry mouth | 121/1000 (12%) | 224/1000 (22%) | 10.3 percentage points | 85% |
| Sexual problems | 55/1000 (5%) | 114/1000 (11%) | 5.9 percentage points | 107% |

people struggle to understand quantitative information[6–8] and some patients prefer physicians who provide quantitative, numerical information.[9,10] Providing quantitative information to patients is also necessary because patients often have different preferences than their physicians—for instance, regarding the acceptance of the risk of side effects.[11] Given that it is therefore absolutely crucial that statistical information be provided to patients, and that it be presented in a way that facilitates its understanding, it is important to examine how physicians actually present statistical information.

There are many situations in which patients evaluate quantitative information during their medical decision making. We investigate patient-physician communication regarding the quantitative information associated with the benefits and sides effects of a medication for depression. It is important to examine physician-patient communication regarding prescription drugs because the use of prescription drugs is on the rise; a 2010 report of the National Health and Nutrition Examination Survey (NHANES) stated that in 2007–2008, 48% of the people in the United States took at least 1 prescription drug in the past month compared with 44% in 1999–2000. The most common type of prescription drug used by 20- to 59-year-olds was antidepressants.[12]

We examined how one physician specialty group, obstetricians and gynecologists (ob-gyns), would communicate statistical information about the benefits and side effects of antidepressants to hypothetical patients. Ob-gyns are primary health care providers for women. Many ob-gyns provide primary care for their patients such as depression care screening and treatment, including prescribing antidepressants.[13,14] In our study, we provided statistical information about an antidepressant[15] to the physicians in a variety of numeric formats (summarized in Table 1) and asked which pieces physicians would use to explain the benefits and side effects to their patients.

We assessed whether physicians chose statistical information that was complete, transparent, and interpretable.[6] *Complete* means that both benefits and side effects were communicated. A *transparent* representation makes clear which proportion of people are affected by the antidepressant. An example of a transparent representation would be reporting the event rates for both the placebo and treatment groups (i.e., that the proportion of patients with sexual problems increases from 5.5% to 11.4%). Alternatively, one could report the absolute risk change (i.e., that it increases by 5.9 percentage points) for both groups. The use of relative risk changes, such as that the antidepressant increases the risk of sexual problems by 107% [100% $\times$ (11.4 − 5.5)/5.5], is an example of a *nontransparent* representation. It is well documented that relative risks lead people to overestimate the effects of drugs, which holds true for even physicians and health professionals.[16] One problem with relative risks is that they do not include baseline risk information. But even including baseline risks together with relative risks does not completely amend the problem.[17] Finally, *interpretability* means that the information is actually meaningful without additional information. This is not the case when only 1 event rate (either under treatment or placebo) is presented in isolation without comparative information about the other event rate or a measure of risk change. For instance, knowing that 45% of patients who took the antidepressant got better is not interpretable without knowing that 26% who took placebo also got better. Past research on consumer choice[18] and medical decision making[19] has shown that choices can become inconsistent when comparative information that is needed to evaluate an attribute is missing, which underlines the importance of interpretability (labeled ''evaluability'' in that research).

Importantly, we assume that not all physicians will want to foster informed choice but that some could aim to ''nudge'' patients toward their own opinion about the antidepressant. Physicians who wanted to foster informed choice would inform each patient independently of their own belief about whether a treatment is good for a particular patient,

whereas physicians who wanted to nudge patients would adapt the numeric information they provide according to their beliefs. We also assume that their goal of informing or nudging should be reflected in the selection of pieces of information. A physician with the goal of informing would make sure that the information is indeed complete, transparent, and interpretable. A physician with the goal of nudging patients, in contrast, would present the information in a way that makes the drug look either more or less favorable than it would look if information were presented transparently. For instance, it would look more favorable if the benefit were presented as relative risk change (big number) but the side effects as absolute risk change (smaller number)—a technique called ''mismatched framing.''[6] The drug can also be made to look more or less favorably by completely omitting one side (i.e., either the benefit or the side effects).

The primary goal of our study was to assess which representations of statistical information physicians would use to explain the benefit and side effects of an antidepressant to hypothetical patients. In this regard, we also assessed whether physicians adapted their information provision according to beliefs about the treatment they were asked to imagine and how the selected representations depended on whether the physicians adapted the information provision. A secondary goal was to investigate whether physicians' self-reported numerical skills (referred to as subjective numeracy) would predict which kind of information they would select.

## METHODS

### Participants

Two-hundred and twelve ob-gyns who had responded to a survey about antidepressants in the past year were surveyed. The participants were all part of the Collaborative Ambulatory Research Network (CARN), which is a group of practicing ob-gyns who are members of the American College of Obstetricians and Gynecologists (ACOG). CARN members volunteer to participate in 2–4 survey research studies each year (no compensation is provided). The CARN group is representative of the ACOG membership with regard to age, gender, and geographic location.

A total of 152 responded for a 72% response rate. Of those, 142 provided completed questionnaires and were included in the final sample. Physicians in the final sample ($N$ = 142) were born, on average, in the year 1956 ($s$ = 27 years; data missing from 2 physicians) and had 20 years of practice ($s$ = 9 years; data missing from 14 physicians). Fifty-nine percent were female (data missing from 2 participants).

### Design and Measures

The primary outcome was the selection of pieces of information in the information choice task. To assess whether physicians adapted their information provision according to their beliefs about the patient, they were asked to imagine 2 hypothetical patients, 1 for whom they believed the antidepressant to be safe and effective, and 1 for whom they believed this not to be the case. More particularly, we looked at the relation between the two: Would physicians who gave both patients the same information more often select complete and transparent statistical information, whereas physicians who gave both patients different information would more often select statistical information that is suited for nudging by framing the statistics in a way to support their view in each case? Note that we explicitly chose a relatively indirect assessment of this preference for informing versus nudging rather than include a direct question in this regard. In pilot testing, physicians informed us that a more direct question called attention to the concept of informed choice versus nudging and they felt compelled to respond in line with informed choice because they felt it was more socially desirable. The indirect assessment that we used avoids this potential problem. The secondary outcome was subjective numeracy and how it relates to the primary outcomes.

*Information choice task.* The information choice task was designed for this survey to provide physicians with statistical information from a review on antidepressants. Table 1 was provided to the participants. It included information about the most important benefit (chance of improving) and side effects (dry mouth and sexual problems) of the antidepressant. The information was provided in 4 different formats, namely as event rate under placebo, event rate under treatment, absolute risk change, and relative risk change.

In the initial ''Safe'' scenario, participants were told: ''Imagine that you have a depressed nonpregnant patient who is considering taking an antidepressant. You believe that an antidepressant will be safe and effective for this particular patient. You decide to use summarized information from a Cochrane Review to educate the patient on the advantages and disadvantages of taking antidepressants.'' They were then asked, ''Which pieces of information

from the Cochrane Review (displayed above) would you most likely use to explain the benefits or disadvantages of antidepressants to this patient?'' They were told, ''You may choose more than 1, but please do not choose more than 4.''

In the subsequent ''Not Safe'' scenario, participants were then told, ''You now have another patient. For this patient, you believe that an antidepressant will not be safe and effective. Again, you decide to use summarized information from a Cochrane Review to educate the patient on the advantages and disadvantages of taking antidepressants. Would you show this patient different information than the patient in the question above?'' Respondents were asked to check ''yes'' or ''no.'' If they checked ''yes,'' they were asked to indicate which pieces of information they would show this patient.

Each piece of information in the table was marked with a letter from A to L, and participants were asked to mark those letters on a list below the table. The restriction to select only a maximum of 4 pieces of information was included to ensure that participants would not simply indicate all pieces of information. Also, if they wanted to provide information about the benefit and each of the 2 side effects, the restriction to 4 pieces of information would force them to choose between absolute and relative risk changes, because they simply could not adhere to the limit of 4 pieces of information when choosing to present event rates under treatment and placebo for the benefit and both side effects.

*Subjective numeracy.* The subjective numeracy scale (SNS)[20,21] consists of 8 questions to which participants responded on scales from 1 to 6. Four questions measured subjective numerical ability (SNS-Ability subscale), and 4 questions measured preferences for numerical information (SNS-Preference subscale). The SNS has been found to have good internal reliability ($\alpha = 0.82$)[20] and to be well correlated with objective numeracy on probabilistic national samples in the United States and Germany.[7] The SNS rather than an objective measure was used for this survey because pilot samples were too intimidated by the objective numeracy questions.

### Data Analysis

Data were analyzed with a personal computer–based version of SPSS 16.0 (SPSS Inc., Chicago, IL). Descriptive and frequency data were computed for primary analysis. $\chi^2$ analyses were conducted for categorical variables, Mann-Whitney *U* tests were conducted for ordinal variables, and analysis of variance and a *t* test were used for comparing subjective numeracy scores. Physicians were grouped according to whether they used identical or different pieces of information to explain the benefit and side effects to the 2 hypothetical patients.

First, we then looked at how frequently each particular piece of information was selected by the overall sample and each of the groups, which we calculated as proportion of physicians who selected a particular piece of information. These proportions were compared within subjects with a sign test for 2 related samples. Next, we looked at patterns of pieces of information provided. Five patterns were of particular interest: (I) Information that is complete and transparent; (II) information that is complete but not transparent, although not suited for nudging; (III) information that is not interpretable for the patient, because 1 event rate is presented in isolation without necessary comparative information about a reference standard or measure of risk change; (IV) information that is suited for nudging by making the benefit appear larger than the side effects; and (V) information that is suited for nudging by making the side effects appear larger than the benefit. Precise definitions for each of those patterns can be found in Table 2.

## RESULTS

### Information Choice Task

Of the 142 physicians, 87 (61.3%) provided identical information independent of whether they were asked to imagine that the drug was safe and effective (*identically informing physicians* subsequently), while 55 (38.7%) provided different information across both scenarios (*differently informing physicians* subsequently). There was a higher proportion of women among identically informing physicians (67% female) compared with differently informing physicians (45%, $\chi^2 = 6.2$, $P = 0.014$).

The proportion of identically informing physicians who provided 1, 2, 3, 4, and >4 pieces of information was 2.3%, 11.5%, 54.0%, 28.7%, and 3.3%, respectively (irrespective of scenario). In the scenario Safe, the proportion of differently informing physicians who provided 1, 2, 3, 4, and >4 pieces of information was 21.8%, 7.3%, 43.6%, 23.6%, and 3.6%, respectively. In the scenario Not Safe, the proportion of differently informing physicians who provided 1, 2, 3, 4, and >4 pieces of information was 10.9%, 23.6%, 34.5%, 25.5%, and 5.5%. In both scenarios, there was a slight tendency that identically informing

**Table 2** Definitions of Patterns of Information

| Pattern | Definition |
|---|---|
| I Complete and transparent | Benefit and at least 1 side effect are reported transparently; that is, either both event rates or absolute risk changes are reported in each case. |
| II Complete but not transparent | Benefit and at least 1 side effect are reported, but the information is not transparent or is only partially transparent. For instance, relative risk changes are reported in each case. While difficult to understand, this is not obviously suited for nudging. |
| III Not interpretable for the patient due to missing comparative information | Only the event rates under treatment are reported for benefit and at least 1 side effect, which is not interpretable without additional information such as event rate under placebo or a measure of risk change. |
| IV Suited for nudging: benefit appears larger | Benefit is reported as relative risk change, while side effects are reported as event rates or absolute risk changes. This makes the benefit look large in comparison. Or, only benefit is reported, while side effects are omitted. |
| V Suited for nudging: harms appear larger | Side effects are reported as relative risk changes, while benefit is reported as event rates or absolute risk change. This makes the side effects look large in comparison. Or, only side effects are reported, while benefit is omitted. |

physicians provided more pieces of information, $z = 1.82$, $P = 0.068$ and $z = 1.63$, $P = 0.104$, respectively.

Table 3 depicts the proportion of physicians—overall and within each group—who would provide each particular piece of information, dependent on the scenario (Safe v. Not Safe). Overall, differently informing physicians provided similar pieces of information as identically informing physicians in scenario Safe but quite different ones in scenario Not Safe. Moreover, differently informing physicians made the antidepressant look more favorably in scenario Safe and less favorably in scenario Not Safe.

In more detail: First, physicians provided more information on the benefit than on each of the harms (all $P$s $< 0.001$), with the exception of differently informing physicians in scenario Not Safe ($P$s $> 0.851$). Second, raw event rates were more often provided than risk changes, be they absolute or relative (all $P$s $\leq 0.020$), again with the exception of differently informing physicians in scenario Not Safe ($P = 0.771$). Third, very often those event rates were only provided for the treatment, but less often for the placebo condition (all $P$s $< 0.001$; except differently informing physicians, scenario Not Safe: $P = 0.093$). And fourth, when we compared the scenarios Safe and Not Safe within differently informing physicians, there seemed to be a flip in the use of absolute risk changes and relative risk changes: In the scenario Safe, the benefit was more often presented as relative than as absolute risk change, and the side effects were more often presented as absolute than as relative risk change ($P = 0.012$ for the benefit and $P = 0.013$ for side effects); in the scenario Not Safe, in contrast, the benefit was more often presented as absolute than relative risk change, and the side effects were more often presented as relative than as absolute risk change ( $P = 0.019$ for the benefit and $P < 0.001$ for side effects).

Patterns of information on the whole sample indicated that 22.5% and 19% of physicians provided pattern I ''Complete and transparent'' in scenarios Safe and Not Safe, respectively, and an additional 23.2% and 23.9%, respectively, provided pattern II ''Complete but not transparent'' (Figure 1). As can be expected, there were more physicians who provided pattern IV ''Benefit appears larger'' in the scenario Safe than in scenario Not Safe (23.9% v. 13.4%, respectively), whereas the reverse held true for pattern V ''Harms appear larger'' (4.9% v. 25.4%, respectively). Pattern III ''Not interpretable'' was shown by 25.4% and 18.3% of the physicians in scenarios Safe and Not Safe, respectively, predominantly by only providing event rates under treatment without event rates under placebo or a measure of risk change.

Identically informing and differently informing physicians differed in the frequencies of choosing various information patterns, as expected (Figure 1). In the scenario Safe, identically informing physicians were more likely to show pattern II ''Complete but not transparent'' than were differently informing physicians (29.9% v. 12.7%) but were less likely to show pattern IV ''Benefit appears larger'' (16.1% v. 36.4%), overall $\chi^2$ across categories = 13.904 ($df = 4$, $N = 142$), $P = 0.007$. Similarly, in the scenario Not Safe, identically informing physicians were more likely to show pattern I ''Complete and transparent'' than were differently informing physicians (24.1% v. 10.9%) and pattern II ''Complete but not transparent'' (29.9% v. 14.5%) but were much less likely to

**Table 3** Proportion of Physicians Providing a Particular Piece of Information

| | Identically Informing Physicians ($n$ = 87) | Differently Informing Physicians ($n$ = 55) | | Total ($N$ = 142) | |
|---|---|---|---|---|---|
| | Scenarios Safe and Not Safe | Scenario Safe | Scenario Not Safe | Scenario Safe | Scenario Not Safe |
| *Benefit* | | | | | |
| Event rate placebo | 26% | 16% | 38% | 23% | 31% |
| Event rate treatment | 70% | 60% | 29% | 66% | 54% |
| Absolute risk change | 21% | 20% | 29% | 20% | 24% |
| Relative risk change | 26% | 42% | 9% | 32% | 20% |
| *Side effect 1* | | | | | |
| Event rate placebo | 7% | 9% | 7% | 8% | 7% |
| Event rate treatment | 43% | 40% | 33% | 42% | 39% |
| Absolute risk change | 18% | 16% | 5% | 18% | 13% |
| Relative risk change | 13% | 4% | 44% | 9% | 25% |
| *Side effect 2* | | | | | |
| Event rate placebo | 13% | 9% | 11% | 11% | 12% |
| Event rate treatment | 56% | 40% | 33% | 50% | 47% |
| Absolute risk change | 20% | 22% | 9% | 20% | 15% |
| Relative risk change | 13% | 2% | 49% | 8% | 27% |

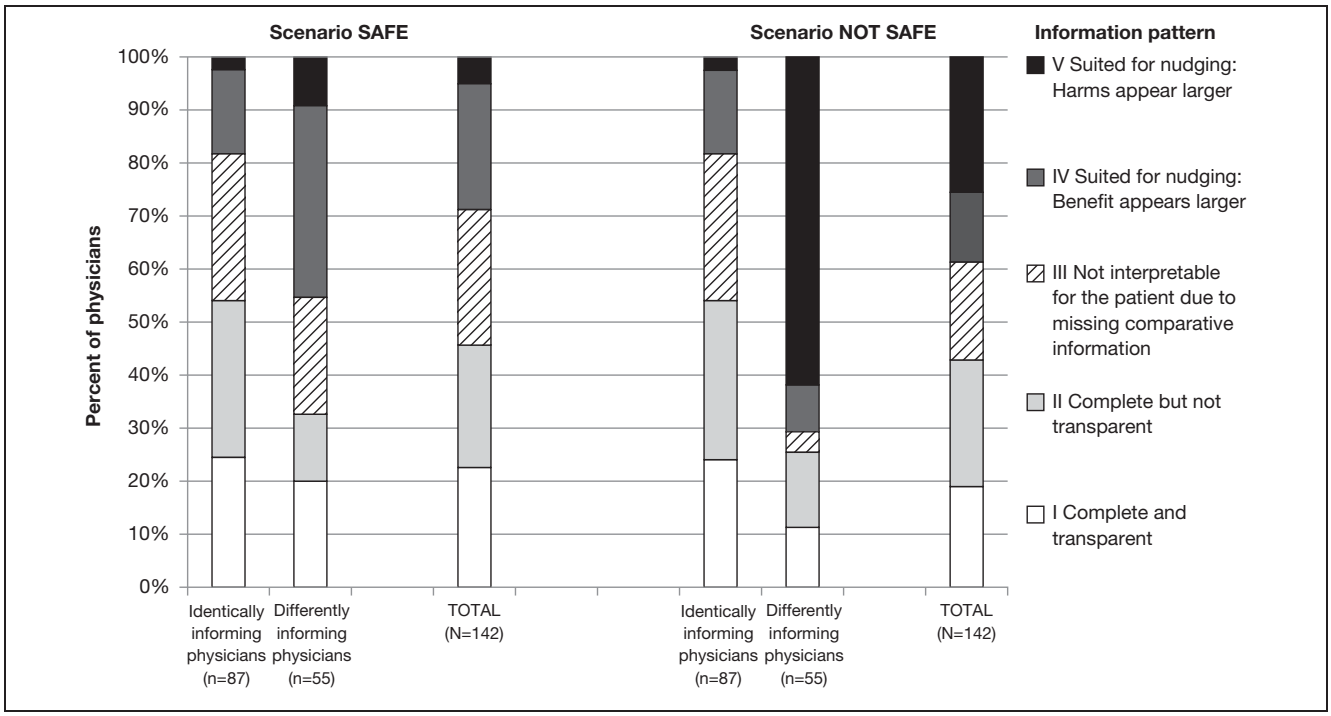


*Figure 1 Proportions of physicians who showed a particular pattern of information provision in the scenarios Safe and Not Safe. Data are depicted separately for the 87 physicians who informed patients identically independent of whether they believed the drug to be safe and effective for them (identically informing physicians), for the 55 physicians who informed patients differently dependent on whether they believed the drug to be safe and effective for them (differently informing physicians), and for the overall sample of 142 physicians (total).*

show pattern V ''Harms appear larger'' (2.3 % v. 61.8%), overall $\chi^2$ across categories = 65.290 ($df$ = 4, $N$ = 142), $P < 0.001$. Identically informing physicians were generally more likely to show pattern III ''Not interpretable'' than were differently informing physicians, and this difference was small in the scenario

Safe (27.6% v. 21.8%) but quite pronounced in the scenario Not Safe (27.6% v. 3.6%).

### Subjective Numeracy

The average score on the subjective numeracy scale for the overall sample was 4.4 on a 6-point scale ($s$ = 0.5; data missing from 2 participants), which is well within range of other studies that found mean subjective numeracy scores between approximately 4 and 5.[20–22] Subjective numeracy scores did not differ between identically informing physicians and differently informing physicians, $t(138) = -0.83$, $P = 0.93$. Subjective numeracy scores were not related to the choice of pattern within both groups of physicians; for identically informing physicians, both scenarios: $F(4,80) = 0.32$, $P = 0.86$; for differently informing physicians, scenario Safe: $F(4,50) = 1.08$, $P = 0.38$; for differently informing physicians, scenario Not Safe: $F(4,50) = 1.94$, $P = 0.12$.

## DISCUSSION

A similar proportion of physicians (roughly 25% each) provided information that was 1) complete and transparent, 2) complete but not transparent, 3) not interpretable for the patient, and 4) suited for nudging. As expected, identically informing physicians (61% of physicians) were more likely to present information that was suited to foster informed choice. Differently informing physicians (39% of physicians), in contrast, were more likely to present information that was suited for nudging in line with the belief they were asked to imagine. To do so, they made either the benefit or harms appear larger, including mismatched framing (e.g., relative risks for benefit and absolute risks for harms)[6] and omission of information. These results suggest that assessing whether physicians adapted the provision of information to beliefs they were asked to imagine distinguished those who more often informed understandably from those who more often provided information suited for nudging (see Figure 1).

However, the results also showed that even among the majority of identically informing physicians, most did not present information that is fully understandable or interpretable; for instance, they provided event rates for treatment groups in isolation, without providing the event rate for control (placebo) groups or a measure of risk change. This practice of focusing only on the treatment group and ignoring the control group (despite that group's importance for evaluating the efficacy of a treatment) is commonly observed in basic research on judgments of contingency between 2 dichotomous variables.[23] This suggests that clinicians may have a lack of awareness for the importance of control groups, similar to what can often be observed in public debates about clinical evidence. For instance, in the recent debate that followed the US Preventive Services Task Force's recommendation not to routinely screen for prostate cancer with the prostate specific antigen (PSA) test, many argued in favor of the test by referring to the large proportion of people who took the test and were still alive 10 years later. To actually evaluate the efficacy of the test, however, the proportion of people who took the test and were still alive 10 years later needs to be compared with the proportion of people who did *not* take the test and were also still alive 10 years later, which is about equally high.[24]

One underlying reason for why even identically informing physicians provided information that was not fully understandable could be a lack of understanding statistical information.[25–29] However, we did not find any relation between the way physicians reported that they would inform patients and their subjective numeracy. It could be that subjective numeracy is not sensitive enough in this regard, although it is correlated with objective numeracy.[7]

However, both the subjective and the objective numeracy scales are concerned with relatively basic operations such as computing fractions. Given that most physicians are likely to have such basic skills,[6] these skills can probably not discriminate well among physicians with regard to interpreting and communicating clinical evidence. Thus, future research should use more advanced measures to assess physicians' abilities to understand statistical evidence[28,30] and how they are related to the communication of quantitative information to patients. While previous research showed that subjective, but not objective, numeracy was related to how likely physicians were to provide quantitative information to their patients at all,[22] it still needs to be determined which skills they require to provide this information in a clearly understandable fashion.

Furthermore, for the question of how physicians communicate quantitative information, it is important to consider not only their quantitative skills but also how they adapt this communication to the (presumed) numerical skills of patients. Given that prior research suggests that the numerical skills of the general population are relatively poor overall[7,8] and that physicians have limited ability to identify patients with low literacy in medicine more generally,[31,32]

physician should always strive to make quantitative information as easily accessible and understandable as possible. One promising way to summarize clinical evidence is the drug facts box, which is a simple tabular representation of medical information. Information is understood well when presented in the drug facts box format.[33,34] Additionally, facts boxes can be accompanied by visual aids to help patients who have difficulties in understanding numbers.[35,36]

### Limitations and Future Directions

The results of this investigation should be considered within the context of the study. First, it only asked physicians to make hypothetical choices for hypothetical patients who were described to them by very brief vignettes. Vignettes are widely used and have been shown to be valid and effective tools.[37,38] Given that the majority of physicians informed identically, and given that those who informed differently did so in a predictable fashion, we believe that physicians followed our thought experiment of 2 different patients in the way we intended.

Second, because this study was conducted via paper survey and not with actual patient encounters, it may lack ecological validity. The paper format of the survey limited our ability to include other numeracy measures (such as objective measures) and it also limited the length and the number of options of the choice task. For instance, physicians could not opt to describe the probabilities verbally as ''high risk'' or ''low risk'' instead of giving numbers.

Third, although we assume that their choices reflect, at least to some degree, each physician's preferences for informing or nudging, we did not directly assess this preference. It could be the case that identically and differently informing physicians differed with regard to whether they interpreted the treatment choice as preference-sensitive. For instance, differently informing physicians could have associated the description of scenario Not Safe with much more serious side effects than those listed in Table 1, and they could have reverted to nudging as the survey did not include the opportunity to express those more serious concerns. Future studies should therefore investigate such interpretations and also include different treatments that vary with regard to how preference-sensitive they are to see whether the proportion of physicians opting for informing or nudging depends on this property.

Fourth, we did not assess whether physicians would be able to intentionally select information that was complete and transparent or to intentionally nudge patients. Our results represent a candid response about what they would suggest to a patient and provide a good starting point for future research.

Fifth, our sample only included gynecologists and obstetricians. Different medical disciplines differ systematically with regard to aspects that could be important for the question of how quantitative information is communicated to patients, such as degree of contact with patients. It would therefore be important to include different groups of physicians in future research to find out whether and how the communication of statistical information differs between medical specialties.

### Implications

If the goal is to increase the likelihood that physicians present statistical information in a way that enables their patients to make informed choices, 2 conditions need to be met. First, physicians need to have a preference to present the information adequately rather than nudging patients toward their own beliefs. The marked and predictable differences between identically and differently informing physicians suggest that this cannot necessarily be taken for granted for each physician.

Second, to improve physicians' communication of statistical information to patients, it will be crucial to improve physicians' understanding of health statistics. This requires making statistical thinking a more important part of the medical curriculum.[6,39–42] In this regard, it would be helpful to demonstrate to physicians how much they themselves are often misled by confusing representations of clinical evidence. To grasp the relevance of this problem, they need to learn that nontransparent statistics or mismatched framing are ubiquitous and can even be found in the leading medical journals.[43,44] As well, one could show them how much insight can be created by good, transparent representations, including using absolute instead of relative risks, natural frequencies instead of conditional probabilities, and so forth.[6] Teaching physicians how to translate misleading representations into transparent ones, and vice versa, will allow them to communicate statistics effectively and to be vaccinated against undue attempts to persuade them of particular treatments.[45]

Some hesitation to inform patients adequately may stem from the concern that patients will not understand statistical information, so that some physicians may prefer to—benevolently—nudge patients instead. Convincing physicians that patient understanding can actually be achieved with adequate representations could help reduce this

hesitation. An improvement of the medical curriculum with regard to statistical thinking will help physicians understand statistical information and increase the likelihood that they will effectively communicate statistical information to patients.

## REFERENCES


1. Berry D, Raynor T, Knapp P, Bersellini E. Over the counter medicines and the need for immediate action: a further evaluation of European Commission recommended wordings for communicating risk. Patient Educ Couns. 2004;53:129–34.

2. Gurmankin AD, Baron J, Armstrong K. The effect of numerical statements of risk on trust and comfort with hypothetical physician risk communication. Med Decis Making. 2004;24:265–71.

3. Wegwarth O, Gigerenzer G. ''There is nothing to worry about'': gynecologists' counseling on mammography. Patient Educ Couns. 2011;84:251–6.

4. Gigerenzer G, Hoffrage U, Ebert A. AIDS counselling for low-risk clients. AIDS Care. 1998;10:197–211.

5. Stiftung Warentest. Urologen im Test: Welchen Nutzen hat der PSA-Test? [Testing urologists: what are the benefits of the PSA test?] Stiftung Warentest. 2004;February:86–9.

6. Gigerenzer G, Gaissmaier W, Kurz-Milcke E, et al. Helping doctors and patients make sense of health statistics. Psychol Sci Public Interest. 2007;8:53–96.

7. Galesic M, Garcia-Retamero R. Statistical numeracy for health: a cross-cultural comparison with probabilistic national samples. Arch Intern Med. 2010;170:462–8.

8. Reyna VF, Nelson WL, Han PK, Dieckmann NF. How numeracy influences risk comprehension and medical decision making. Psychol Bull. 2009;135:943–73.

9. Mazur DJ, Hickham DH, Mazur MD. How patients' preferences for risk information influence treatment choice in a case of high risk and high therapeutic uncertainty: asymptomatic localized prostate cancer. Med Decis Making. 1999;19:394–8.

10. Shaw NJ, Dear PRF. How do parents of babies interpret qualitative expressions of probability? Arch Dis Child. 1990;65:520–3.

11. Heesen C, Kleiter I, Nguyen F, et al. Risk perception in natalizumab-treated multiple sclerosis patients and their neurologists. Mult Scler. 2010;16:1507–12.

12. Gu Q, Dillon CF, Burt VL. Prescription drug use continues to increase: U.S. prescription drug data for 2007–2008. NCHS Data Brief. 2010;42:1–8.

13. Leddy M, Haaga D, Gray J, Schulkin J. Postpartum mental health screening and diagnosis by obstetrician-gynecologists. J Psychosom Obstet Gynecol. 2011;32:27-34.

14. Schmidt LA, Greenberg BD, Holzman GB, et al. Treatment of depression by obstetrician–gynecologists: a survey study. Obstet Gynecol. 1997;90:296–300.

15. Rayner L, Price A, Evans A, et al. Antidepressants for depression in physically ill people. Cochrane Database Syst Rev. 2010; (3):CD007503.

16. Covey J. A meta-analysis of the effects of presenting treatment benefits in different formats. Med Dec Making. 2007; 27:638–54.

17. Covey J. The effects of absolute risks, relative risks, frequencies, and probabilities on decision quality. J Health Commun. 2011;16:788–801.

18. Hsee CK. The evaluability hypothesis: an explanation for preference reversals between joint and separate evaluations of alternatives. Organ Behav Hum Decis Process. 1996;67:247–57.

19. Zikmund-Fisher BJ, Fagerlin A, Ubel PA. ''Is 28% good or bad?'' Evaluability and preference reversals in health care decisions. Med Decis Making. 2004;24:142–8.

20. Fagerlin A, Zikmund-Fisher BJ, Ubel PA, et al. Measuring numeracy without a math test: development of the subjective numeracy scale. Med Decis Making. 2007;27:672–80.

21. Zikmund-Fisher BJ, Smith DM, Ubel PA, Fagerlin A. Validation of the Subjective Numeracy Scale: effects of low numeracy on comprehension of risk communications and utility elicitations. Med Decis Making. 2007;27:663–71.

22. Anderson BL, Obrecht NA, Chapman G, Driscoll DA, Schulkin J. Physicians' communication of Down syndrome screening test results: the influence of physician numeracy. Genet Med. 2011; 13:744–9.

23. Arkes HR, Harkness AR. Estimates of contingency between two dichotomous variables. J Exp Psychol Gen. 1983;112:117–35.

24. Arkes HR, Gaissmaier W. Psychological research and the prostate-cancer screening controversy. Psychol Sci. 2012;23:547–53.

25. Berwick DM, Fineberg HV, Weinstein MC. When doctors meet numbers. Am J Med. 1981;71:991–8.

26. Rao G. Physician numeracy: essential skills for practicing evidence-based medicine. Fam Med. 2008;40:354–8.

27. Wegwarth O, Gigerenzer G. Statistical illiteracy in doctors. In: Gigerenzer G, Gray JAM, eds. Better Doctors, Better Patients, Better Decisions: Envisioning Healthcare in 2020. Struengmann Forum Report. Cambridge, MA: MIT Press; 2011. p 137–51.

28. Wegwarth O, Gaissmaier W, Gigerenzer G. Deceiving numbers: survival rates and their impact on doctors' risk communication. Med Decis Making. 2011;31:386–94.

29. Wegwarth O, Schwartz LM, Woloshin S, Gaissmaier W, Gigerenzer G. Do physicians understand cancer screening statistics? A national survey of primary care physicians in the United States. Ann Intern Med. 2012;156:340–9.

30. Anderson BL, Gigerenzer G, Parker S. Schulkin J. Statistical literacy in obstetricians and gynecologists. J Healthc Qual. 2012. doi: 10.1111/j.1945-1474.2011.00194.x.

31. Bass PF III, Wilson JF, Griffith CH, Barnett DR. Residents' ability to identify patients with poor literacy skills. Acad Med. 2002; 77:1039–41.

32. Kelly PA, Haidet P. Physician overestimation of patient literacy: a potential source of health care disparities. Patient Educ Couns. 2007;66:119–22.

33. Schwartz LM, Woloshin S, Welch HG. The drug facts box: providing consumers with simple tabular data on drug benefit and harm. Med Decis Making. 2007;27:655–62.

34. Schwartz LM, Woloshin S, Welch HG. Using a drug facts box to communicate drug benefits and harms: two randomized trials. Ann Intern Med. 2009;150:516–27.

35. Garcia-Retamero R, Galesic M. Communicating treatment risk reduction to people with low numeracy skills: a cross-cultural comparison. Am J Public Health. 2009;99:2196–202.

36. Gaissmaier W, Wegwarth O, Skopec D, et al. Numbers can be worth a thousand pictures: individual differences in understanding graphical and numerical representations of health-related information. Health Psychol. 2012;31:286–96.

37. Veloski J, Tai S, Evans AS, Nash DB. Clinical vignette-based surveys: a tool for assessing physician practice variation. Am J Med Qual. 2005;20:151–7.

38. Peabody JW, Luck J, Glassman P, et al. Comparison of vignettes, standardized patients, and chart abstraction. JAMA. 2000;283:1715–22.

39. Altman DG, Bland JM. Improving doctors' understanding of statistics. J R Stat Soc A. 1991;154:223–67.

40. Appleton DR. What statistics should we teach medical undergraduates and graduates? Stat Med. 1990;9:1013–21.

41. Gigerenzer G. Calculated Risks: How to Know When Numbers Deceive You. New York: Simon & Schuster. (UK version: Reckoning with Risk: Learning to Live with Uncertainty, London: Penguin); 2002.

42. Hertwig R, Buchan H, Davis D, et al. How will health care professionals and patients work together in 2020? A manifesto for change. In: Gigerenzer G, Gray JAM, eds. Better Doctors, Better Patients, Better Decisions: Envisioning Healthcare in 2020. Strueng-mann Forum Report. Cambridge: MIT Press; 2011. p. 317–37.

43. Gigerenzer G, Wegwarth O, Feufel M. Misleading communication of risk: editors should enforce transparent reporting in abstracts. Br Med J. 2010;341:791–2.

44. Schwartz LM, Woloshin S, Dvorin EL, Welch HG. Ratio measures in leading medical journals: structured review of accessibility of underlying absolute risks. Br Med J. 2006;333:1248–52.

45. Gigerenzer G, Gray JAM, eds. Better Doctors, Better Patients, Better Decisions: Envisioning Healthcare 2020. Cambridge, MA: MIT Press; 2011.