

Why not? – Communicating stochastic information by use of unsorted frequency pictograms – a randomised controlled trial

Warum eigentlich nicht? – Kommunikation stochastischer Informationen mit unsortiert arrangierten Häufigkeits-Piktogrammen – eine randomisierte kontrollierte Studie

Abstract

Objective: Statistical health risk information has been proven confusing and difficult to understand. While existing research indicates that presenting risk information in frequency formats is superior to relative risk and probability formats, the optimal design of frequency formats is still unclear. The aim of this study was to compare presentation of multi-figure pictographs in consecutive and random arrangements regarding accuracy in perception and vulnerability for cognitive bias.

Methods: A total of 111 patients with multiple sclerosis were randomly assigned to two experimental conditions: patient information using 100 figure pictographs in 1) unsorted (UP group) or 2) consecutive arrangement (CP group). The study experiment was framed as patient information on how risks and benefit could be explained. The information comprised two scenarios of a treatment decision with varying levels of emotional relevance. Primary outcome measure was accuracy of information recall (errors made when recalling previously presented frequencies of benefits and side effects). Cognitive bias was measured as additional error appearing with higher emotional involvement. The uncertainty tolerance scale and a set of items to assess risk attribution were surveyed.

Results: The study groups did not differ in their accuracy of recalling benefits, but recall of side effects was more accurate in the CP-group. Cognitive bias when recalling benefits was higher in the UP-group than in the CP-group and equal for side effects in both groups. Results were similar in subgroup analyses of patients 1) with highly irrational risk attribution 2) with experience regarding the hypothetical contents or 3) with experience regarding pictograph presentation of frequencies. Overall, benefit was overestimated by more than 100% and variance of recall was extremely high.

Conclusions: Consecutive arrangement as commonly used seems not clearly superior to unsorted arrangement which is more close to reality. General poor performance and the corresponding high variance of recall might have clouded existing effects of the arrangement types. More research is needed with varying proportions and other samples.

Keywords: risk communication, risk perception, medical decision-making, graphical format, quantitative information

Zusammenfassung

Ziele: Statistische Gesundheitsinformationen (z.B. Krankheitsrisiken) können für Patienten schwierig zu verstehen und verwirrend sein. Während die Forschung gezeigt hat, dass absolute Risikoangaben (z.B. natürliche Häufigkeiten) gegenüber relativen Risikoangaben oder Wahrscheinlichkeitsangaben zur Risikokommunikation überlegen sind,

Jürgen Kasper^{1,2}

Christoph Heesen¹

Sascha Köpke^{2,3}

Ingrid Mühlhauser²

Matthias Lenz²

1 Institute of Neuroimmunology and Clinical MS Research (inims), University Medical Center Hamburg, Germany

2 Faculty of Mathematics, Informatics, and Natural Sciences (MIN-Faculty), Unit of Health Sciences and Education, University of Hamburg, Germany

3 Nursing Research Group, Institute of Social Medicine, University of Lübeck, Lübeck, Germany

ist bislang unklar, wie Häufigkeiten optimalerweise dargestellt werden sollten.

Ziel dieser Studie ist ein Vergleich zwischen Häufigkeitspiktogrammen mit sortierter Anordnung der einzelnen Figuren und Häufigkeitspiktogrammen mit zufälliger Anordnung hinsichtlich der Wahrnehmungsgenauigkeit und einer Anfälligkeit für irrationale Wahrnehmungsverzerrung.

Methoden: Per Randomisierung wurden 111 Patienten mit Multipler Sklerose zwei unterschiedlichen experimentellen Bedingungen zugeordnet: Patienteninformation mit 100er Piktogrammen in 1) zufälliger Anordnung (unsorted arrangement = UP-Gruppe) oder 2) sortierter Anordnung (consecutive arrangement = CP-Gruppe). Das Experiment bestand in einer Patienteninformation über verschiedene Darstellungsmöglichkeiten von Krankheitsrisiken und Nutzen und Nebenwirkungen medizinischer Interventionen. Die Information enthielt zwei Szenarien zu hypothetischen Behandlungsentscheidungen mit unterschiedlichen Ausprägungen emotionaler Relevanz. Primärer Endpunkt war die Genauigkeit der Informationswiedergabe (als Wiedergabefehler). Irrationale Wahrnehmungsverzerrung wurde als zusätzlicher Wiedergabefehler bei stärkerer emotionaler Involviertheit definiert. Ungewissheitstoleranz und Risikoattribution wurden ebenso evaluiert.

Ergebnisse: UP-Gruppe und CP-Gruppe unterschieden sich nicht in der Wiedergabegenauigkeit von Interventionsnutzen. Nebenwirkungshäufigkeiten wurden in der CP-Gruppe genauer geschätzt. Bei der Wiedergabe des Interventionsnutzen war die Wahrnehmungsverzerrung in der UP-Gruppe größer als in der CP-Gruppe und der Nebenwirkungen in beiden Gruppen gleich. Subgruppenanalysen der Patienten mit 1) stärker irrationaler Risikoattribution 2) ohne Vorerfahrungen mit der betreffenden Entscheidung oder 3) ohne Vorerfahrungen mit der Darstellungsform ergaben ähnliche Ergebnisse. Insgesamt wurde der Interventionsnutzen um mehr als 100% überschätzt. Die Varianz in der Wiedergabegenauigkeit war extrem hoch.

Schlussfolgerung: Die üblicherweise verwendete sortierte Anordnung von Figuren in Häufigkeitspiktogrammen ist einer zufälligen Anordnung nicht klar überlegen. Dabei spiegelt die zufällige Anordnung die Verteilung von Fällen in der Realität besser wider. Die schlechte Wiedergabegenauigkeit und damit einhergehend hohe Varianz könnten allerdings tatsächliche Unterschiede zwischen den Darstellungsformen verschleiern haben. Weiterer Forschungsbedarf besteht insbesondere bezüglich verschiedener Risikogrößen und anderer Zielgruppen.

Introduction

To make informed health decisions, patients need the relevant information presented in an adequate manner. Criteria for high quality patient information have been established based on ethical and legal and scientific considerations [1]. According to these criteria patient information has to be understandable, evidence based, complete, relevant for patients, and balanced. Developers are guided in selecting content and appropriate presentation methods. This has been referred to as “evidence-based patient information (EBPI)” [2]. However, successful communication of probabilistic information, such as health risks, frequencies of benefit and harm of therapies, prognostic or diagnostic tests still seems the most difficult task in patient information and even physician information [3], [4]. In particular, single event probabilities, conditional probabilities, and relative risks have been found to be confusing, largely due to poor presentation of information

[3]. Erroneous communication of risks may cause serious consequences such as medical decisions based on inappropriate expectations. Various graphical formats exist to visualise empirical information to patients or consumers including bar graphs, survival curves, and pie charts [1], [5], [6]. In particular, diagrams comprising different forms of multiple figures (e.g. stick-figures or smileys) are frequently used within EBPI or patient decision aids [7] to visualize disease risks, and benefits and harms of medical interventions. These illustrations often show multi-figure pictographs (MFPs), i.e. an arrangement of figures (N=100 or N=1000) representing the distribution of patients with effects/no effects or side effects/no side effects of an intervention. Supplemental graphical illustration of proportions of events of benefit or harm has been shown to support understanding when compared to text only risk information [8] and to positively influence physician patient communication [4]. Within the various graphical formats MFPs have been proven

advantageous [9], [10], [11]. Qualitative data suggest that people prefer MFPs to bar charts since it “is clearer that you’re talking about human beings and not statistics” [12]. However, it is not clear, how such risk formats are best composed to support understanding and to improve decision making quality. Evaluation of pictographs’ ability to appropriately mediate risk information needs to consider two aspects of such information: Firstly, the information aims at accurately transmitting a specific proportion e.g. the percentage of people who benefit from a certain therapy and of people who do not benefit. Secondly, the information has to indicate its probabilistic character, i.e. that uncertainty remains about the prediction whether an individual belongs to the group with or without a benefit.

The process of information is an interaction of the graphical presentation with the perceptual, emotional, and cognitive structure of its recipients. For instance, as a result from emotional involvement with the decision, patients might tend to irrationally process given quantities. This may lead to over- or underestimation or over-salient perception of events that objectively are not prominently presented. Also, depending on individual degrees of tolerance towards complex and ambiguous situations – called uncertainty tolerance [13] – patients’ information processing might vary widely. As a cognitive style, uncertainty tolerance has been proven an important factor in learning and information processing not only in education but also in health communication [13]. Graphical formats to communicate likelihoods of health related events should as far as possible be robust against such biases. Therefore, understanding of internal mechanisms of processing risk information might be helpful to minimize cognitive bias.

To adequately account for the state of knowledge in this area we systematically searched Medline, EMBASE, PsycINFO, PSYINDEXplus, and CINAHL (last search June 2010) for original studies investigating the MFP-method used to visualize probabilistic information in the context of shared decision making. The search strategy was based on the latest Cochrane Review on decision aids [14]. Searches were limited to English or German language publications but not limited according to type of study design or publication year. Only original studies were included that compared MFPs with other graphical formats to communicate likelihoods in the context of shared decision making or that compared different design features of MFPs. Reference lists of retrieved publications were screened for additional publications. Retrieved material was analysed independently by two investigators (JK and ML).

After duplicates were removed a total of 1129 potentially relevant titles were identified. 1111 of which could be excluded with certainty. Of the remaining 18 references assessed for eligibility as full-text, four were found relevant for the study question [12], [15], [16], [17].

Two studies compared MFP with other presentation formats [16], [17].

In a randomised trial [16] response time of a hypothetical medical choice and accuracy of recall of previously presented risks were compared among six different presentation formats (pie charts, vertical bars, horizontal bars, numbers, consecutively arranged MFPs, and randomly arranged MFPs).

Pie charts and random ovals caused the slowest and less accurate performances [16].

In another study [17] focus groups were used to assess the ease of interpretation, simplicity and perceived vulnerability against bias of perceived risk magnitude in bar graphs compared to MFPs. MFPs turned out to be appreciated for higher salience compared to bar graphs, but also to be associated with potential bias regarding the accurate perception of magnitude of a risk [17].

Four studies evaluated differential effects of a variety of features of MFPs [12], [15], [16], [17]. In a controlled study, differences in accuracy and speed of perception of quantitative risk information were surveyed for specific types of MFPs, in particular two single pictographs compared to one compound pictograph, pictographs arranged in horizontal compared to vertical direction, and shading compared to no shading of the highlighted groups of figures in the pictographs [15]. Horizontal pictographs were perceived faster and more accurately than vertically arranged ones and two single pictographs showed faster perception compared to compound formats particularly with shaded background. Although shading increased response time in compound pictograms, this feature was preferred to the more effective forms [15].

In another study [12] qualitative methods were used to investigate the effects of 1) an interactive application of MFPs in comparison to a static application and 2) random pictograms compared to consecutive arrangements. Animation and interactive use of MFPs evoked more emotional response, both positive and negative than viewing static graphs. Random arrangements of figures (indicating persons affected by the disease throughout a group of figures) made it more difficult to judge the proportions but were considered more realistic than consecutive arrangements [12].

Two variants of pictographs were part of an aforementioned trial [15] that compared formats regarding perception speed and accuracy. Consecutively arranged pictographs were superior to randomly arranged formats regarding both accuracy and speed of perception.

One study employed focus groups with women who were asked to comment different formats illustrating breast cancer risk [17]. The study investigated the size of the denominator MFPs are based on (amount of individuals in a reference group). MFPs using a lower denominator were preferred to those with higher denominator. In addition, the order (arrangement) of pictograms was assessed. Random order was difficult to process for some participants, however, also recognized as part of concept of conveying random occurrence.

In summary, it seems that formats subjectively perceived best for making a choice to some extent differ from those best for estimating the size of an amount.

The actual state of evidence on MFPs does not provide much guidance whether or how to use this format to support understanding of probabilistic risk information. Whereas evidence suggests that consecutive arrangements and numerical data are perceived easier and more accurately, potential advantages of the random order as found in some qualitative studies are hitherto not sufficiently understood. No studies were identified, evaluating different types of arrangements regarding potential differences in accuracy of information recall and the understanding of randomness.

Having in mind the importance of patients' accurate comprehension of numerical risk data and the issue of randomness in the concept of risk, the need becomes obvious to understand the effects of pictogram arrangement on cognitive and emotional perception of risk.

Objectives

This study aimed at investigating advantages of the use of 100 figure pictographs presented in unsorted compared to those in consecutive order. Although the common way to present pictographs in consecutive order seems more likely to yield precise comprehension of proportions, unsorted pictograms are recognized closer to reality. Perception of this information might therefore be less vulnerable for irrational biases evoked by individual cognitive attribution styles or uncertainty tolerance. We assessed the effects of the order within a pictograph on accuracy of comprehension and on vulnerability for irrational biases.

Methods

A randomised controlled trial was conducted in 2007 testing two experimental conditions: patient information using 100 figure pictographs 1) in unsorted (UP group) and 2) in consecutive arrangement (CP group). Within each condition two hypothetical scenarios representing different levels of relevance were provided in order to investigate potential differences in information processing caused by varying degrees of importance of the content.

Participants

The study population consisted of a convenience sample of adults suffering from multiple sclerosis (MS). Participants were recruited at the MS Outpatient Unit of University Medical Centre Hamburg-Eppendorf, the Centre of Neurology Hamburg-Bramfeld, and the Hamburg MS Society. Patients with confirmed MS were eligible if they were between 18 and 70 years old, German speaking and not obviously cognitively or, according to own statement, heavily visually impaired.

Experimental intervention

The study experiment was framed as a booklet for patients providing information on methods used in EBPI to communicate risks and benefits. Study participants were asked to read the booklet thoroughly and then to recall specific parts of it. The booklet consisted of 17 charts. (The experiment booklets are attached online to this paper: consecutive order in English language (Attachment 1), consecutive order in German language (Attachment 2), random order in English language (Attachment 3), random order in German language (Attachment 4)). The first five charts introduced the issue of absolute and relative risk reduction. Two examples were given by using pictures of glasses each containing 1000 blue and yellow pearls in different quantities referring to proportions of persons with benefit or no-benefit. This introductory part was intended to motivate participants and to equalize potentially varying degrees of understanding about the concept of risk. The remaining 12 charts contained the actual experiment, risk information, embedded in two hypothetical scenarios presenting treatment decisions. Each consisted of two "therapeutic" options for which probabilities of benefit and side effects were displayed. The first scenario – 'garden pond algae pollution' – illustrated a non medical problem and was assumed to be of low relevance and is subsequently referred to as "neutral scenario (NSc)". As 'therapeutic' option we chose the 'utilization of water slugs' compared to 'no action'. The second scenario – 'hypothetical MS-treatment' – was assumed to be perceived as highly relevant for participants ('relevant scenario' (RSc)). In both scenarios, fictitious probabilities of benefits and side effects were given. Proportions were based on existing MS-treatment effects. To avoid recognition, probabilities were slightly varied between both scenarios. A previously evaluated five step proceeding was used to structure the two scenarios. This approach was shown to be feasible to provide risk information about specific treatment effects comprehensibly [18]: 1) the four possible outcomes were explained theoretically: intervention & no benefit, intervention & benefit, no intervention/placebo & no benefit, no intervention/placebo & benefit (2 charts). 2) Quantities of ponds/patients showing no benefit and those showing benefit were presented in two 100 figure pictographs, one for treatment and one for placebo/no treatment (NSc: without treatment 31, with treatment 45 ponds/patients showing improvement out of 100; RSc: without treatment ponds/patients showing no worsening 31, with treatment 43 ponds/patients showing no worsening) (1 chart). 3) One chart displayed the absolute risk reduction using one single compound pictograph (NSc: 14 out of 100; RSc: 12 out of 100) (1 chart). 4) The potential side effects were explained theoretically: having (or not having) side effects with (or without) intervention (2 charts). 5) An additional chart visualized the quantity of side effects using another 100 figure pictograph. Pictographs were presented without any numerical or textual explanation (Figure 1), so that participants could



Figure 1: Different formats of 100-figure pictogram illustrations

gather the quantitative information solely from the chart. This proceeding intended to maximize potential differences between the two formats.

Measurement

Recall of quantities was assessed for benefit and side effects for NSc and RSc. Accuracy of recall of benefit regarding NSc was defined as the primary outcome. After participants had thoroughly studied the charts, they were interviewed by a trained researcher. The interview was structured using 10 additional charts.

Participants had to identify the correct information and to translate the graphically displayed proportions into frequency values (“how many are affected out of 100?”).

To prevent participants from perceiving the situation as a performance test, the experiment was framed as an information session. The participants had to complete both scenarios (NSc & RSc) under the same conditions before they were interviewed. Recall performance was assumed not only being a matter of (sensory) perception but also a matter of cognitive processing such as memory and subjective appraisal. The latter is sensitive to any irrational assumptions, attitudes or styles of evaluation. As potentially relevant to this kind of process, self-serving bias might impact on the result of cognitive information processing [19]. The self-serving bias refers to a tendency to evaluate ambiguous information in a way that is beneficial for one’s interests. Here, self-serving bias would lead to overestimation of the own chance of benefit and

underestimation of the risk to be affected by side effects accordingly. However, we assumed such kind of biases only in the processing of emotionally important contents i.e. the RSc. Therefore, intra-individual differences in recall performance between NSc and RSc were assumed to indicate a presentation form's vulnerability to evoke such cognitive bias. This was defined as the secondary outcome. As another indicator of patients' tendency to irrationally bias their cognitive processing either as an optimism bias or as a pessimism bias [20], disease specific risk attribution was assessed by asking participants to estimate their personal risk of becoming wheelchair bound (according to [21]). Using visual analogue scales (VAS), patients had to rate their risk higher or lower compared to a fictive person at equal age and state of disease, which was represented in the middle of the VAS. This approach had been used before in a study evaluating risk perception of MS patients [21]. Estimating one's own risk higher than that of a matching person indicates a problem-sensitive and pessimistic coping style. Downsizing one's own risk indicates a more repressive coping style. Different VAS were presented separately referring to the 2-year, 10-year and lifetime risk. In addition, by using the 8 item uncertainty tolerance-questionnaire [13], degrees of participants' cognitive affinity to uncertain and complex situations were assessed.

To determine whether presentation forms had an impact on choices, a hypothetical choice was assessed. To avoid responses according to patients' own situation, participants were asked "Would you recommend this treatment to another person affected by MS". Participants could respond on a 4 point Likert scale ranging from 0 (absolutely yes), 1 (tend to 'yes'), 2 (tend to 'no'), and 3 (absolutely no) for each scenario. Furthermore, we qualitatively assessed reasons for reported intentions by an open format question. To control for possible confounding with recall accuracy, familiarity with the specific scenario was assessed. Using three point Likert scales, participants were asked for instance whether they were personally concerned by 'algae pollution of garden ponds', whether they were experienced in making MS-treatment decisions, and whether they were familiar with MFPs similar to those used in this study. To describe the study sample demographically and regarding disease related aspects, age, gender, highest level of education, duration of MS in months, and the Cambridge Multiple Sclerosis Basic Score (CAMBS) [22] were assessed. The CAMBS considers four aspects of MS: disability and impairment, relapse, progression, and handicap. The whole study experiment was piloted with ten MS patients to warrant readability and to achieve indicators for sample size calculation.

Study conduct

Allocation to presentation of either UP or CP was conducted randomly and concealed by drawing pre-coded questionnaires from a batch after a participant had given informed consent. Subsequently, baseline data including

uncertainty tolerance and risk attribution were surveyed. In face-to-face interviews participants were asked to step-wise work through all pages of the booklet. If required, the interviewer gave assistance. To avoid misunderstanding of the concepts of relative risk reduction and absolute risk reduction, the investigator checked comprehension when participants reached the chart explaining those concepts, using a standardized question. If necessary, additional explanation was given.

After participants had read the experimental information recall of presented frequencies and the hypothetical choice were assessed. Finally, in order to assess the participants' preferences regarding presentation formats, alternative formats were presented together with the question on preference.

Hypotheses

1. We expected frequencies from the CP to be recalled more accurately.
2. On the other hand however, we expected the CP to promote a cognitive-motivational bias of the information process. We assumed an easier identification of figures representing individuals with a positive outcome to facilitate self-serving bias, whereas the random arrangement might strengthen patients' awareness of the probabilistic nature of risk. Driven by individual attribution styles, this motivational bias was presumed to particularly affect the processing of the RSc, i.e. we expected (hypothetical) choices to follow the perception bias.
3. Participants tending to overestimate benefits and to underestimate harms were assumed to make more positive choices and vice versa.
4. Accordingly, we expected risk attribution (RA) to influence processing of risk information. Thus, optimistic RA was assumed to be associated with overestimation of treatment effects and underestimation of side effects.
5. We expected participants to prefer CPs compared to UPs. However, this preference was expected to decrease with increasing uncertainty tolerance.

Methods of analysis

Accuracy of recall was drawn from the NSc data by calculating differences between raw values (frequency of benefit and side effects) and the corresponding true value. If the given value in the booklet was 15% risk, a corresponding recall risk of 25% meant a 10% error. Cognitive bias was defined on an intra-individual level based on the additional recall error occurring in the RSc in relation to the NSc. Cognitive bias was operationalized by calculating the difference of errors made when recalling the frequencies of the RSc and of the NSc, expressed as absolute value on a 0 to 100 scale. Experimental conditions were compared regarding accuracy of recall and cognitive bias using unpaired t-tests.

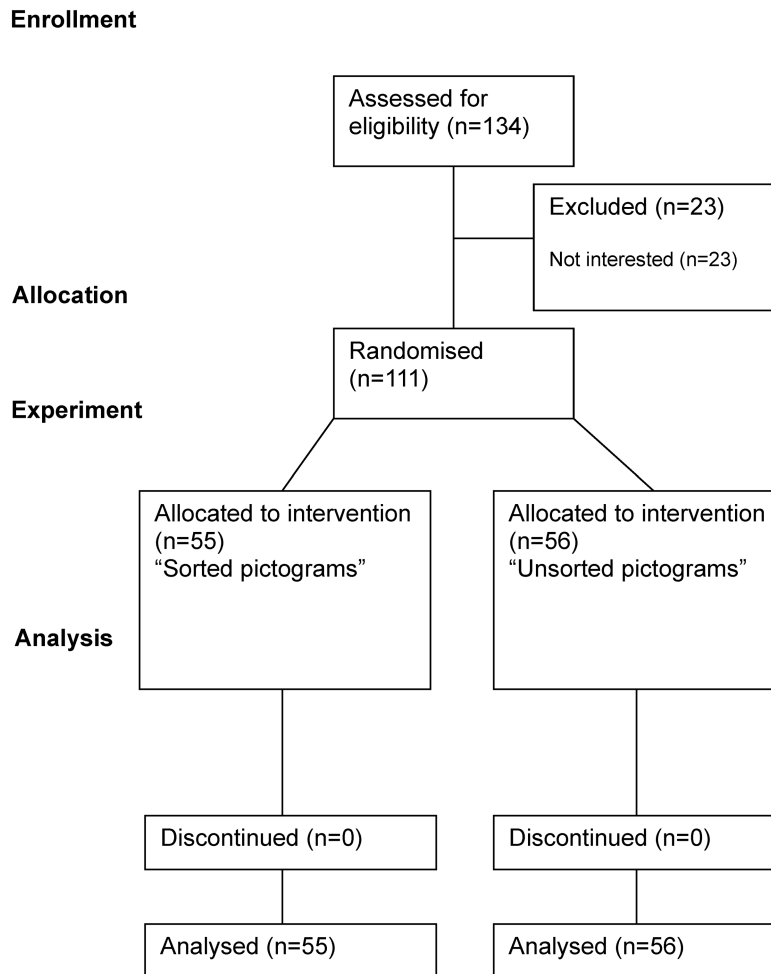


Figure 2: Flow of participants

Participants rating their personal risk of becoming wheelchair bound at least 10% higher or lower compared to their matches were defined as vulnerable towards cognitive bias. In the subgroups of these participants showing either a pronounced optimistic or pessimistic attitude, the secondary endpoint (cognitive bias) was tested between the experimental conditions. Using Pearson correlation coefficients, risk attribution and uncertainty tolerance were correlated to accuracy regarding each domain of recall: frequencies of benefit and side effects for NSc and RSc. Transformed into a scale ranging from minus 1.5 to plus 1.5, data on choice were correlated to accuracy as well.

Sample size

The primary endpoint was the accuracy of recall of benefit in the NSc, where effects of order of pictographs were assumed to be independent of cognitive biases. Sample size calculation was based on data from the pilot study (N=10) (G-power 3.0 [23]). Assuming 90% power, 0.05 α -level (two sided) and a 16.5 standard deviation, a minimal total sample size of N=96 was calculated, required to identify a mean difference of 10% in the primary endpoint, which we considered important.

Results

A total of 111 patients were assigned to the two study conditions (Figure 2). Demographic and disease-related variables were equally distributed with respect to gender, level of education, disease duration, and disability. In the UP-group participants were older (Table 1).

Accuracy of recall of frequencies was low in both groups (mean of absolute error values for benefit_{NSc}=16.5%, side effects_{NSc}=11.4%, benefit_{RSc}=18.5%, side effect_{RSc}=13.2%). Patients overestimated benefits more than 100% (mean error for the NSc=+15.0, SD=18.7, for RSc=+17.7, SD=19.0). Recall of frequencies related to side effects were more precise, however with pronounced variance (mean error for NSc=-4.9, SD=13.1; RSc=+2.7, SD=17.2). While study groups did not differ in their accuracy of recalling benefits, the recall of side effects was more accurate in the CP-group (Table 2). Contradicting our hypothesis, the cognitive bias when recalling benefits was higher in the UP-group than in the CP-group and equal for side effects in both groups (Table 2). Results were similar in subgroup analyses (Table 2) that included only participants who showed a considerable tendency to make irrational personal risk attributions. Accordingly, risk attribution was not correlated to accuracy in three of

Table 1: Demographic and disease-related data

	UP-group (n=56)	CP-group (n=55)	Total (n=111)	p-value
Female	42 (75%)	43 (78%)	85 (76.6%)	.43
Male	14 (25%)	12 (22%)	26 (23.4%)	
Age (years)	44.9 (±8.9)	40.5 (±10.4)	42.7 (±9.9)	.02
Education (years)				.44
>12	20 (36%)	22 (39%)	42 (37.8%)	
10–12	30 (54%)	25 (45%)	55 (49.5%)	
0–9	5 (9.1%)	9 (16%)	14 (12.6%)	
Duration of MS (months)	88 (±73)	91 (±83)	90 (±78)	.81
Disability (CAMBS)	1.1 (±0.5)	1.0 (±0.5)	1.1 (±0.5)	.11
Maximal walking distance				.63
>500 m	34 (61%)	39 (70.9%)	73 (66.4%)	
<500 m	10 (18.2%)	7 (12.7%)	17 (15.5%)	
<300 m	2 (3.6%)	1 (1.8%)	3 (2.7%)	
<200 m	4 (7.3%)	1 (1.8%)	5 (4.5%)	
<100 m	4 (7.3%)	5 (9.1%)	9 (8.1%)	
<20 m	1 (1.8%)	2 (3.6%)	3 (2.7%)	

CAMBS=Cambridge Multiple Sclerosis Basic Score, assessing physical disability status

Table 2: Results

	UP-group means (SD)	CP-group means (SD)	p-value
Analyses of primary endpoint accuracy (regarding recall of neutral scenario, NSc)			
Accuracy benefit	14.7 (15.5)	18.2 (19)	.30
Accuracy side effects	13.5 (8.84)	9.3 (6.8)	.007
Accuracy (regarding relevant scenario referring to RSc)			
Accuracy benefit	20.2 (20.4)	16.8 (16.1)	.337
Accuracy side effects	15.7 (12.4)	10.8 (9.6)	.024
Secondary endpoint: cognitive bias			
Net mistake (RSc minus NSc)			
Benefit	5.18 (18.0)	-1.4 (16.5)	0.05
Side effects	2.1 (12.0)	1.3 (10.1)	.70
1. Subgroup (participants vulnerable for cognitive bias) analyses of secondary endpoint			
Groups drawn from two year risk attribution			
Cognitive bias benefit (N=28/32)	2.9 (14.7)	-5.5 (19)	.06
Cognitive bias side effects (N=29/30)	.4 (12.5)	1.2 (7.2)	.76
Groups drawn from ten year risk attribution			
Cognitive bias benefit (N=31/33)	4.1 (21.3)	-5.2 (19)	.07
Cognitive bias side effects (N=32/31)	.1 (11.7)	2.2 (10.2)	.44
Groups drawn from life time risk attribution			
Cognitive bias benefit (N=29/26)	5.2 (21.6)	-3.42 (16.9)	.12
Cognitive bias side effects (N=30/24)	-2.0 (9.6)	.1 (6.9)	.36
2. Subgroup (participants not concerned with garden ponds) analyses of primary endpoint			
Accuracy benefit (N=40/42)	15.1 (16.4)	16.7 (19.6)	.69
Accuracy side effects (N=41/41)	14.2 (8.8)	9.1 (6.5)	.004
3. Subgroup (participants not previously involved in immunotherapy decisions) analyses of secondary endpoint			
Cognitive bias benefit (N=10/13)	-1.4(17.2)	1.4 (22.1)	.75
Cognitive bias side effects (N=11/12)	-1.0(14.7)	.0 (7.6)	.84
4. Subgroup (participants not familiar with pictographs) analyses of primary & of secondary endpoints			
Accuracy benefit (N=41/34)	14.4 (14.5)	18.4 (19.5)	.31
Accuracy side effects (N=41/41)	13.6 (7.9)	9.0 (7.3)	.01
Cognitive bias benefit (N=41/34)	4.7 (14.4)	-.3 (17.0)	.18
Cognitive bias side effects (N=41/41)	3.1 (11.9)	1.3 (10.4)	.48

All mean values are estimation errors referring to an absolute scale of 0 to 100%.

the four domains of recall ($NSc_{benefit} r = -0.027$, $NSc_{side\ effects} r = .147$, $RSc_{benefit} r = 0.118$, $RSc_{side\ effects} r = 0.204$, Pearson correlations with comparative two year risk of becoming wheelchair bound) and slightly associated with $RC_{side\ effect}$. Participants with a more pessimistic attribution style tended to overestimate side effects and vice versa ($r = 0.21$, $p = .035$). Uncertainty tolerance was not correlated with the accuracy in all four recall domains. However, a low negative correlation appeared, indicating that higher uncertainty tolerance might be associated with lower irrational information processing (uncertainty tolerance and cognitive bias $r = -0.20$, $p = 0.04$; uncertainty tolerance cognitive bias $r = -0.13$, $p = 0.18$). Hypothetical choices were predominantly positive towards ‘treatment’ in both groups and in both scenarios with a tendency to more positive choices in the UP-group (NC: $p = 0.22$, RC: $p = 0.09$). Accuracy was not correlated with choice in the total sample, indicating that differences in perceiving the information had no impact on the decision. Unsorted pictographs were preferred by only two participants (2%), one out of each study group. Confounding variables were equally distributed between both conditions, except of familiarity with pictograms (UP-group: 75% = “not at all”, CP-group = 62%). Subgroup analyses 1) of participants not concerned with garden pond pollution (Table 2), 2) of participants without previous experience with MS-treatment decisions (Table 2), and 3) of participants not familiar with using pictographs did not lead to different results (Table 2).

Discussion

The study demonstrated that participants had pronounced difficulties to correctly perceive and process risk information presented by using MFPs. Recall of benefits displayed showed marked overestimation. This result was not explained by either experimental variation of pictograph format (unsorted vs. consecutive) or by indicators of irrational cognitive processing (risk attribution and uncertainty tolerance). We found no clear difference in patients’ ability to process consecutive or unsorted arrangements. Although patients showed better recall of side effects when using consecutive pictographs, no corresponding effect was seen for presentation of benefit. In contradiction to our hypothesis that consecutive arrangement could mislead vulnerable individuals to bias their perception in accordance with their risk attribution style, we found a more pronounced cognitive bias in the UP-group for benefit, while no such effect was seen for side effects. Also, no correlation was seen between accuracy of recall, risk attribution and uncertainty tolerance.

Regarding the ambiguous results, recommendations about how to present MFPs seem premature. Moreover, generalizability of the study results may be limited by the experimental character of the design. The unusual presentation of frequencies (figuratively and without supplemental numerical or textual information) might have increased uncertainty of recall and might even have

clouded actually existing effects [12]. The present study has further limitations. The presentation example used only one certain proportion which was chosen as representative for MS treatment effects. It cannot be ruled out that using different proportions would yield different results. Furthermore, given proportions did not allow for recall mistakes up to the same extent in both directions. Therefore, overestimation was more likely than underestimation. However, systematic variation of proportions would have further increased the required sample size. Furthermore, the hypothetical character of the given information may have limited the validity of reported choices. It is questionable whether and to which extent reported choices were based on the given information or based on already existing beliefs. In contrast to our study Schapira [17], Ancker [12], and Feldman-Stewart [16] found that randomly arranged pictographs were more difficult to process cognitively. Interestingly, although participants expressed a clear preference for the consecutive order, our study does not indicate that comprehension is worse in the unsorted format. This preference might be explained by the higher cognitive efforts required to process unsorted arrangements. Taking into account the indistinct results, further research seems warranted e.g. with modified design features. Also, it could be helpful to address bigger samples of healthy persons not previously contaminated with either pictograph presentation or medical decisions.

There are a number of inherent goals of presenting different frequency formats to support patients’ decision making [8]. Beyond accuracy and speed of capturing given frequencies, other aspects of the information process should be considered. On the cognitive-behavioural level usability, preference or behaviour planning might be important parameters. Although potentially interfering with some of the abovementioned goals, addressing uncertainty of given risk information can be another important goal. Accordingly, research should focus on awareness of randomness as another potential effect of presentation of frequency formats. Developers should use different existing formats according to their priority of specific communication goals [17]. For instance, some evidence exists that unsorted arrangements of figures emphasize randomness [12], [17] while consecutive arrangements seem to emphasize accuracy of perception [12], [15], [17]. Therefore, developers could purposefully combine these formats to facilitate either accurate perception or awareness of randomness or both.

Conclusions

Evidence based patient information remains challenging as tools and strategies are still far from reaching satisfactory standards. Therefore, further research about optimizing technologies to communicate probabilities is required.

Abbreviations

CP-group = study group receiving pictograms arranged consecutively

EBPI = Evidence Based Patient Information

MFP = multi-figure pictograph

MS = multiple sclerosis

NSc = neutral scenario

RSc = highly relevant scenario

UP-group = unsorted pictogram group

VAS = visual analogue scales

RA = risk attribution

Notes

Acknowledgements

We thank Petra Gutzeit and Johanna Kaiser for their support in designing the study protocol and data assessment.

Conflict of interests

Christoph Heesen has received financial support from Biogen Idec, Bayer-Health Care, Serono, Sanfor-Aventis and Teva Pharma; Sascha Köpke, Jürgen Kasper, Ingrid Mühlhauser and Matthias Lenz have nothing to declare. We are not aware of any potential conflict of interests.

Attachments

Available from

<http://www.egms.de/en/journals/psm/2011-8/psm000077.shtml>

1. ExperimentConsecutiveOrderEnglish.pdf (90 KB)
Experiment booklet consecutive order in English language
2. ExperimentConsecutiveOrdeGerman.pdf (90 KB)
Experiment booklet consecutive order in German language
3. ExperimentRandomOrderEnglish.pdf (93 KB)
Experiment booklet random order in English language
4. ExperimentRandomOrderGerman.pdf (90 KB)
Experiment booklet random order in German language

References

1. Bunge M, Mühlhauser I, Steckelberg A. What constitutes evidence-based patient information? Overview of discussed criteria. *Patient Educ Couns.* 2010;78(3):316-28. DOI: 10.1016/j.pec.2009.10.029
2. Coulter A. Evidence based patient information. is important, so there needs to be a national strategy to ensure it. *BMJ.* 1998;317(7153):225-6.
3. Gigerenzer G, Edwards A. Simple tools for understanding risks: from innumeracy to insight. *BMJ.* 2003;327(7417):741-4. DOI: 10.1136/bmj.327.7417.741
4. Paling J. Strategies to help patients understand risks. *BMJ.* 2003;327(7417):745-8. DOI: 10.1136/bmj.327.7417.745
5. Steckelberg A, Berger B, Köpke S, Heesen C, Mühlhauser I. Kriterien für evidenzbasierte Patienteninformationen [Criteria for evidence-based patient information]. *Z Arztl Fortbild Qualitätssich.* 2005;99(6):343-51.
6. Trevena LJ, Davey HM, Barratt A, Butow P, Caldwell P. A systematic review on communicating with patients about evidence. *J Eval Clin Pract.* 2006;12(1):13-23. DOI: 10.1111/j.1365-2753.2005.00596.x
7. O'Connor AM, Légaré F, Stacey D. Risk communication in practice: the contribution of decision aids. *BMJ.* 2003;327(7417):736-40. DOI: 10.1136/bmj.327.7417.736
8. Edwards A, Elwyn G, Mulley A. Explaining risks: turning numerical data into meaningful pictures. *BMJ.* 2002;324(7341):827-30. DOI: 10.1136/bmj.324.7341.827
9. Fagerlin A, Zikmund-Fisher BJ, Derry HA, Sarr BJ, Smith D, Ubel PA. What does order have to do with it? The role of risk/benefit order in shaping risk perception. 3rd International Shared Decision Making Conference; University of Ottawa, Ontario Canada, 14-16 June 2005. Program and Abstracts. p. 74. Available from: http://decisionaid.ohri.ca/ISDM2005/ISDM2005_book.pdf
10. Hawley ST, Zikmund-Fisher B, Ubel P, Jancovic A, Lucas T, Fagerlin A. The impact of the format of graphical presentation on health-related knowledge and treatment choices. *Patient Educ Couns.* 2008;73(3):448-55. DOI: 10.1016/j.pec.2008.07.023
11. Pylar J, Wills CE, Lillie J, Rovner DR, Kelly-Blake K, Holmes-Rovner M. Men's interpretations of graphical information in a videotape decision aid. *Health Expect.* 2007;10(2):184-93. DOI: 10.1111/j.1369-7625.2007.00443.x
12. Ancker JS, Chan C, Kukafka R. Interactive graphics for expressing health risks: development and qualitative evaluation. *J Health Commun.* 2009;14(5):461-75. DOI: 10.1080/10810730903032960
13. Dalbert C. Die Ungewißheitstoleranzskala: Skaleneigenschaften und Validierungsbefunde. Halle: Martin-Luther Universität Halle-Wittenberg, Institut für Pädagogik; 1999. (Hallesche Berichte zur Pädagogischen Psychologie; Nr. 1). Available from: <http://psydok.sulb.uni-saarland.de/volltexte/2004/393/>
14. O'Connor AM, Bennett CL, Stacey D, Barry M, Col NF, Eden KB, Entwistle VA, Fiset V, Holmes-Rovner M, Khangura S, Llewellyn-Thomas H, Rovner D. Decision aids for people facing health treatment or screening decisions. *Cochrane Database Syst Rev.* 2009;(3):CD001431. DOI: 10.1002/14651858.CD001431.pub2
15. Price M, Cameron R, Butow P. Communicating risk information: the influence of graphical display format on quantitative information perception-Accuracy, comprehension and preferences. *Patient Educ Couns.* 2007;69(1-3):121-8. DOI: 10.1016/j.pec.2007.08.006
16. Feldman-Stewart D, Kocovski N, McConnell BA, Brundage MD, Mackillop WJ. Perception of quantitative information for treatment decisions. *Med Decis Making.* 2000;20(2):228-38. DOI: 10.1177/0272989X0002000208
17. Schapira MM, Nattinger AB, McHorney CA. Frequency or probability? A qualitative study of risk communication formats used in health care. *Med Decis Making.* 2001;21(6):459-67. DOI: 10.1177/0272989X0102100604

18. Kasper J, Köpke S, Mühlhauser I, Heesen C. Evidence-based patient information about treatment of multiple sclerosis—a phase one study on comprehension and emotional responses. *Patient Educ Couns*. 2006;62(1):56-63. DOI: 10.1016/j.pec.2005.06.002
19. Miller DT, Ross M. Self-serving biases in the attribution of causality: fact or fiction? *Psychol Bull*. 1975;82 (2):213-25. DOI: 10.1037/h0076486
20. Kahneman D, Slovic P, Tversky A, eds. *Judgment under uncertainty: Heuristics and biases*. Cambridge, UK: Cambridge University Press; 1982.
21. Janssens C. Perception of prognostic risk in multiple sclerosis [thesis]. Rotterdam: Erasmus University; 2003. Available from: http://www-fgg.eur.nl/medbib/EUR-diss/030521_Janssens_C/Janssens_C.html
22. Mumford CJ, Compston A. Problems with rating scales for multiple sclerosis: a novel approach—the CAMBS score. *J Neurol*. 1993;240(4):209-15. DOI: 10.1007/BF00818706
23. Faul F, Erdfelder E, Lang AG, Buchner A. G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav Res Methods*. 2007;39(2):175-91. DOI: 10.3758/BF03193146

Corresponding author:

Dr. Matthias Lenz, PhD
Faculty of Mathematics, Informatics, and Natural Sciences (MIN-Faculty), Unit of Health Sciences and Education, University of Hamburg, Martin Luther King Platz 6, 20146 Hamburg, Germany, Phone: + 49 40 42838 7232, Fax: + 49 40 42838 3732
matthias.lenz@uni-hamburg.de

Please cite as

Kasper J, Heesen C, Köpke S, Mühlhauser I, Lenz M. Why not? – Communicating stochastic information by use of unsorted frequency pictograms – a randomised controlled trial. GMS Psychosoc Med. 2011;8:Doc08. DOI: 10.3205/psm000077, URN: urn:nbn:de:0183-psm0000777

This article is freely available from

<http://www.egms.de/en/journals/psm/2011-8/psm000077.shtml>

Published: 2011-11-02

Copyright

©2011 Kasper et al. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by-nc-nd/3.0/deed.en>). You are free: to Share — to copy, distribute and transmit the work, provided the original author and source are credited.