

Assessing the Impact of Visualizations of Quantitative and Qualitative Risk Information on the Comprehension of Hypothetical Risk-Based Breast Cancer Screening Results

Inge S. van Strien-Knippenberg, Danielle R.M. Timmermans, Maaïke Weber, Yasmina Okan, Carla H. van Gils, Mireille J.M. Broeders & Olga C. Damman

To cite this article: Inge S. van Strien-Knippenberg, Danielle R.M. Timmermans, Maaïke Weber, Yasmina Okan, Carla H. van Gils, Mireille J.M. Broeders & Olga C. Damman (2026) Assessing the Impact of Visualizations of Quantitative and Qualitative Risk Information on the Comprehension of Hypothetical Risk-Based Breast Cancer Screening Results, Health Communication, 41:9, 1668-1683, DOI: [10.1080/10410236.2025.2582779](https://doi.org/10.1080/10410236.2025.2582779)

To link to this article: <https://doi.org/10.1080/10410236.2025.2582779>



© 2025 The Author(s). Published with license by Taylor & Francis Group, LLC.



[View supplementary material](#)



Published online: 10 Nov 2025.



[Submit your article to this journal](#)



Article views: 1397



[View related articles](#)



[View Crossmark data](#)

Assessing the Impact of Visualizations of Quantitative and Qualitative Risk Information on the Comprehension of Hypothetical Risk-Based Breast Cancer Screening Results

Inge S. van Strien-Knippenberg ^a, Danielle R.M. Timmermans ^a, Maaïke Weber^a, Yasmina Okan ^b,
Carla H. van Gils ^c, Mireille J.M. Broeders ^d, and Olga C. Damman ^a

^aDepartment of Public and Occupational Health, Amsterdam Public Health Research Institute, Amsterdam University Medical Center; ^bDepartment of Communication, Pompeu Fabra University; ^cJulius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht University; ^dDepartment IQ Health, Radboud University Medical Center & Dutch Expert Centre for Screening

ABSTRACT

This study aimed to assess whether, in the context of communicating risk-based breast cancer screening results, visualizations of risk information consisting of an icon array depicting quantified breast cancer risk and simplified drawings for “qualitative” risk information (e.g. risk factors, screening frequency) offer added value compared to a textual format presenting numerical information. An online experiment was conducted with a 3 (information format: text with numbers – text with risk visualizations – interactive) × 4 (breast cancer risk category: slightly decreased, average, slightly increased, increased) between-subjects design. We recruited $n = 1,047$ women aged 40–74 years from the general Dutch population, including 413 (39.4%) with lower education and 269 (25.7%) with a family history of breast cancer. Participants viewed hypothetical screening results. Primary outcomes were verbatim and gist comprehension of numerical probability information and verbatim comprehension of qualitative risk information. Secondary outcomes (e.g. affective reactions, perceived likelihood) and health literacy, numeracy, and graph literacy were also assessed. Results showed no significant differences in comprehension between “text with risk visualizations” and “text with numbers” formats. Verbatim comprehension of numerical probability and qualitative risk information was worse in the “interactive” format compared to “text with numbers” format. Those assigned to a non-average risk category exhibited lower verbatim and gist comprehension. Overall, the visualizations tested in this study did not enhance comprehension of risk-based breast cancer screening results compared to a text format with numerical information. Understanding risk-based screening results may be challenging regardless of format, particularly for people assigned to a non-average risk category.

Risk-based breast cancer screening

In the Netherlands, all women between the ages of 50–74 are invited biennially for breast cancer screening with mammography. If no breast abnormality is found, a woman receives a letter stating that no abnormality was found and that she will receive a new invitation after two years until she is 75. Although still a subject of research, a risk-based approach may be implemented for breast cancer screening in the future (Broeders & Paci, 2015; Fitzgerald et al., 2022; RIVM, 2023). Risk-based screening may include considering risk factors such as age of first menstrual period, family history, and breast density (Cliff et al., 2022). Based on a mammography and such risk factors, women can be classified into a breast cancer risk category which can be used to determine, for example, the screening frequency and method. Although there are currently no risk-based screening programs for breast cancer, effective risk communication will be key for the successful implementation of such programs in the future.

To make informed decisions about cancer screening participation, people need to understand the key benefits and

potential harms. This has proven difficult, as such information often involves complex screening statistics which can be counterintuitive for lay people, such as the risk of overdiagnosis (Hersch et al., 2018; Petrova et al., 2015; Schwartz et al., 1997). Although it is not yet clear whether risk-based screening will be implemented and what it will look like, the results of such screening will in any case contain a more multifaceted risk message. For the current study, in consultation with experts, we assumed a risk-based screening approach in which a woman is assigned to one of four risk categories (i.e., slightly decreased, average, slightly increased, or increased risk of developing breast cancer), each with a corresponding screening interval and method.

When communicating the results of risk-based screening, in addition to information about the absence of abnormalities, information about the classification into a risk category and the associated recommended follow-up screening interval and method need to be communicated. The risk communication literature suggests that providing comparative risk information can help people put risk information into context (Waters

CONTACT Inge S. van Strien-Knippenberg  i.vanstrien@amsterdamumc.nl  Department of Public and Occupational Health, Amsterdam University Medical Center, Vrije Universiteit Amsterdam, Amsterdam Public Health Research Institute, Boelelaan 1117, Amsterdam 1081 HV, The Netherlands
 Supplemental data for this article can be accessed online at <https://doi.org/10.1080/10410236.2025.2582779>

© 2025 The Author(s). Published with license by Taylor & Francis Group, LLC.
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

et al., 2021) but that it also has an impact on people's risk perception, independently of the absolute magnitude of the probability (Fagerlin et al., 2007; Windschitl et al., 2002). For example, in an experiment in which women received hypothetical comparative risk information, participants perceived themselves as more vulnerable when their risk was framed as higher compared to a group whose risk was framed as lower, even though the absolute risk was the same (Windschitl et al., 2002). In a similar experiment in which absolute risk was also kept the same, women were more likely to take medication when a risk was framed as higher than average compared to lower than average (Fagerlin et al., 2007). There is also evidence that people who have more difficulty with information use, such as those with lower health literacy or numeracy skills, generally have more difficulty understanding health risk information (Berkman et al., 2011; Lipkus & Peters, 2009; Reyna et al., 2009). This raises the question of how the results of risk-based screening can best be communicated, including to those with lower health literacy/numeracy, especially when no healthcare provider is involved and people interpret the information themselves. The current study investigated the added value of visualizations of risk information, both an icon array¹ for quantified breast cancer risk and visualizations for "qualitative" risk information (e.g., risk factors and the recommended screening plan), on comprehension compared to textual information with numerical probability information, with all formats designed according to evidence-based principles.

Communicating quantitative health risk information using visualizations

General recommendations for communicating numerical probability information include using numbers instead of verbal terms only and specifying the denominator (Bonner et al., 2021). However, there is no simple answer to the question of how best to communicate quantified risk. This depends on the communication goal and information characteristics, for example, a single number or multiple numbers when comparing different treatment options (Ancker et al., 2022). Quantified risk information in risk-based breast cancer screening indicates how many people in the relevant risk group are expected to develop breast cancer in the coming period and how this compares to the average breast cancer risk. Key recommendations in this context include specifying the reference class, using a consistent reference class, and ensuring that readers do not need to do the math themselves, for example by explicitly stating differences in the probabilities between risk categories (Bonner et al., 2021).

Additionally, numerical probability information can be displayed visually, for example through icon arrays or bar charts. Visual formats are generally recommended to facilitate interpretation of quantitative information (Trevena et al., 2021). Visual formats can reduce the cognitive effort required to process abstract information by making calculations and comparisons visually accessible (Ancker et al., 2006; Franconeri et al., 2021). For example, graphs depicting both the number of people affected by a risk and the total number of people at risk help to make part-to-whole relationships visible, which in turn

supports understanding of absolute risk magnitudes (Garcia-Retamero & Cokely, 2017; Stone et al., 2018). When processing numerical information, according to Fuzzy Trace Theory, two different types of mental representations are formed: verbatim and gist representations (Reyna, 2008; Reyna et al., 2022). Verbatim representations concern the exact numbers and words, that is, the literal content of the message, for example 9% risk of developing a particular disease (Reyna, 2008; Reyna et al., 2022). These verbatim representations can also be used for calculations, such as calculating the difference in risk between 9% and 4% (Brust-Renck & Reyna, 2023). On the other hand, gist representations concern the bottom-line meaning of the information, that is, the essence of the information (Reyna, 2008; Reyna et al., 2022). For example, gist representations can include general representations, such as the likelihood of a risk occurring, e.g., likely or unlikely, or representations of the order of magnitude, such as more or less (Brust-Renck & Reyna, 2023; Reyna & Brust-Renck, 2020). It has been argued that visualizations of numerical information can be particularly effective in supporting gist comprehension by making patterns more immediately visible (Trevena et al., 2021). However, there is also evidence that some visualizations can help to improve verbatim comprehension (Garcia-Retamero & Cokely, 2017), for example, icon arrays used to communicate the harms and benefits of two medications (Tait et al., 2010).

There is no single visual format that is best in all situations or for all communication goals (Ancker et al., 2024a). For example, visualization features effective at increasing comprehension are often not the same as features that help to change behavior (Ancker et al., 2006; Okan et al., 2018). When only the numerator is visually displayed, it can lead to higher perceived anxiety, worry, and risk aversion which may in turn lead to behavior change. In contrast, when visualizations direct attention equally to both the numerator and denominator of a probability, comprehension is often enhanced, but there is generally less effect on behavior change (Ancker et al., 2006; Stone et al., 2017). Different visualization formats can also elicit different affective responses. For example, one study found that health risk information presented in an icon array elicited a stronger affective response than the same information presented as frequencies or percentages (Timmermans et al., 2008). Moreover, visualizations that are not well-designed can lead to confusion or misinterpretation (Ancker et al., 2006; Bresciani & Eppler, 2015), highlighting the importance of design based on evidence-based principles. Such principles include avoiding inverted or truncated scales, keeping visualizations simple, and respecting common associations, such as that a higher bar graph means "more" (Franconeri et al., 2021; Garcia-Retamero & Cokely, 2017; Trevena et al., 2021).

To communicate quantified breast cancer risk in the context of risk-based breast cancer screening, an icon array is a potentially suitable visual format. Icon arrays, together with bar graphs, have been widely studied and they generally seem to support comprehension (Trevena et al., 2021; Zipkin et al., 2014). One of the advantages of icon arrays is that the reference class, i.e., the denominator, is clearly visible (Garcia-Retamero et al., 2010). The question is whether an icon array

supports understanding of risk-based breast cancer screening results compared to a text format with numerical probability information when both formats are designed using evidence-based principles.

Communicating qualitative health risk information using visualizations

Communications about the results of risk-based breast cancer screening involve not only quantitative information but also qualitative risk information, such as whether no abnormality was found, which risk factors contribute to breast cancer, and the recommended screening plan. This information can also be visualized. Visualizations of qualitative information include, for example, illustrations, photographs, pictures, cartoons, and pictograms² (Houts et al., 2006; Schubbe et al., 2020). General advantages of such visualizations include their ability to draw attention, structure and order information, and enhance information comprehension and recall (Houts et al., 2006). However, as with visualizations of numerical information, visualizations of qualitative information can cause confusion or misinterpretation (Bresciani & Eppler, 2015; Newman & Schwarz, 2024). Design principles are therefore important, such as the application of Gestalt principles (Koffka, 1922), the use of simple pictograms, minimizing the use of abstract or ambiguous symbols, and taking into account cultural differences in symbol interpretation (Bresciani & Eppler, 2015; Houts et al., 2006). Simplified drawings appear to facilitate comprehension better than photographs, potentially because photographs often contain more details that may be distracting (Houts et al., 2006).

In the context of health information in general, a previous systematic review showed that using pictures, e.g., pictograms and cartoons, can improve patients' knowledge and recall across a wide range of medical topics (Schubbe et al., 2020). In the context of cancer screening, schematic illustrations showing how polyps grow and how sigmoidoscopy can prevent cancer were found to improve understanding of the preventive aim of bowel cancer screening (Brotherstone et al., 2006).

In the present study, we examined whether visualizations of risk information improve understanding of hypothetical results of risk-based breast cancer screening, relative to a text format presenting numerical probability information, with both formats developed in accordance to evidence-based principles. The visualization format combined an icon array to communicate quantified breast cancer risk with simplified drawings representing the more “qualitative” risk information, co-created with women in a previous study. We also aimed to investigate the influence of assigning participants to a breast cancer risk category on information comprehension. We hypothesized that:

H1a and H1b: Women exposed to text accompanied by an icon array – as part of an information leaflet in which other risk information is also visualized- will show better verbatim (H1a) and gist (H1b) comprehension of the numerical probability information (i.e., risk of developing cancer in the next 5 years)

compared to women exposed to a text format with numerical probability information.

H2: Women exposed to text with visualizations of the “qualitative” risk information – as part of an information leaflet in which other risk information is also visualized- will show a higher verbatim comprehension of this information (e.g., no abnormalities found, assigned risk category, advice for follow-up screening interval and screening method) compared to women exposed to a text format.

Additionally, we aimed to explore whether the effects of visualizations of risk information on information comprehension differ for women assigned to different risk categories (RQ1).

Individual differences

People with lower health literacy (HL) and/or numeracy skills generally have more difficulty understanding numerical probability information (Davis et al., 2002; Fransen et al., 2017; Schwartz et al., 1997) and may therefore particularly benefit from visualizations (Garcia-Retamero & Cokely, 2017; Oudhoff & Timmermans, 2015). A recent review found some evidence that visualizations such as icon arrays can support comprehension of numerical information for those with lower HL (Richter et al., 2023). For people with lower numeracy, visualizations of numerical information may help to provide meaning to information that would otherwise remain abstract (Lipkus & Peters, 2009). In addition to HL and numeracy, graph literacy (GL), defined as “the ability to understand graphically displayed information,” has also been shown to influence people's understanding of visualizations. People with lower GL may benefit less from visualizations of quantitative information than people with higher GL (Galesic & Garcia-Retamero, 2011; Okan et al., 2019; Richter et al., 2023).

Visualizations of qualitative information have also been shown to increase comprehension for people with lower HL. Examples include illustrations in wound care instructions and pictograms in medication-taking instructions (Houts et al., 2006; Schubbe et al., 2020). However, a review of visualizations of qualitative information showed that positive effects of visualizations among the less literate were only found when this target group was involved in designing the visualizations (Mbanda et al., 2021). Therefore, it is important to consider the different skill levels of people in both the development and evaluation of visualizations.

RQ2: Are the effects of visualizations of risk information on information comprehension different for individuals with different skill levels in HL, numeracy, and GL?

Static visualizations and interactive features

In health communication and decision support, tools sometimes use interactive elements, such as the option to click on different treatments and outcomes to compare treatments (Trevena et al., 2021). Another example is interactivity in icon arrays (Stellamanns et al., 2017), which may help direct

attention to key information and promote more active information processing (Okan et al., 2015). Interactive icon arrays may therefore offer a useful approach for communicating the results of risk-based breast cancer screening. However, icon array interactivity has been applied differently in different studies. For example, in one study, participants who had to click on the correct number of icons within icon arrays exhibited worse risk understanding than those who viewed a static icon array (Zikmund-Fisher et al., 2011). In another study, different elements of icon arrays were presented sequentially, with explanatory labels also appearing in some conditions. The results showed that the sequential presentation of elements did not influence comprehension, but adding labels improved comprehension for participants with lower GL (Okan et al., 2015). The underlying assumption is that presenting information sequentially may help reduce information overload, yet most of the available studies indicate that respondents do not perform better with interactive icon arrays compared to static icon arrays (Ancker et al., 2024b).

Interactive features can also relate to qualitative information. In a review and meta-analysis of interactive features in computer-based decision aids, one of the interactive features was content control, i.e., the ability of users to navigate through information and decide for themselves whether to read additional information (Syrowatka et al., 2016). The meta-analysis showed that decision aids with content control positively affected knowledge compared to decision aids without content control. However, the different components of content control had different effects on the outcomes. For example, being able to navigate, such as with a menu bar, hindered knowledge, while having the option for additional information and external sources had a positive effect on knowledge (Syrowatka et al., 2016). In the context of risk-based breast cancer screening, interactive elements could allow presenting the core information (e.g., the risk factors) and allowing users to decide whether they want to read further explanations (e.g., an explanation of what the risk factors exactly entail). This could contribute to reducing information overload and increasing user engagement, which could potentially improve comprehension (Syrowatka et al., 2016). The current study therefore explored the effects of interactive elements in communicating the results of risk-based breast cancer screening on women's comprehension. Due to the mixed findings regarding interactive elements in previous research, we did not formulate hypotheses.

RQ3: What are the effects of an interactive format on the verbatim and gist comprehension of the numerical probability information and the verbatim comprehension of the qualitative risk information compared to a text format with numerical probability information?

Additionally, because the use of risk visualizations and assignment to a risk category can also affect outcomes beyond comprehension, such as affective reactions (Fagerlin et al., 2007; Houts et al., 2006; Timmermans et al., 2008; Windschitl et al., 2002), we also examined several secondary outcomes, namely: affective reaction, perceived likelihood, affective risk assessment, and evaluation of information materials (for details see the Methods section).

RQ4: What are the effects of information formats and risk categories on affective reactions, perceived likelihood, affective risk assessment, and evaluation of information materials?

Methods

Design

In an online randomized experiment, women from the general Dutch population were asked to imagine themselves receiving results from risk-based breast cancer screening with fictitious but realistic information. We used a 3 (information format: text with numbers – text with risk visualizations – interactive) $\times$ 4 (breast cancer risk category: 1 – 2 (average) – 3 – 4) between-subjects experimental design. Respondents were stratified by educational level (low, versus middle/high) and randomly assigned to one of the 12 conditions.

This study was exempted from extensive review by the medical research ethics committee of Amsterdam UMC, location VUmc (FWA00017598) following local regulatory guidelines/standards for human subjects' protection in the Netherlands (Medical Research Involving Human Subjects Act). The study was pre-registered before data collection through the Open Science Framework on December 16 2021 at https://osf.io/r723n/?view_only=96b98b45ca8449a3878fcdc2c652a80c. Deviations from the pre-registration are indicated. Not all research questions from the pre-registration are addressed in the current manuscript. The research questions that relate to the acceptance of risk-based screening and the intention to participate will be addressed in another manuscript, as well as possible differences between women assigned to different risk categories.

Participants

Participants ($N=1,047$) were women from the general Dutch population aged 40–74 without a history of breast cancer. Participants were recruited by the online panel Motivaction (ISO-20252, ISO-27001 certified) in December 2021. A flow diagram of participant inclusion and randomization is presented in Figure 1.

An a priori power analysis was performed for sample size estimation. Effect sizes in studies focusing on presentations of numerical probability information in graphs tend to be small (Trevena et al., 2021). For the power analysis, we used a Cohen's f of .115 with an alpha of .05 based on previous research on the effects of different visualizations of breast and prostate cancer survival rates on comprehension (Recchia et al., 2020). For a 4×3 factorial ANOVA with interaction effects and one covariate, the required sample size was 1,037 for a power of .80.³

Materials

All study materials and survey items are available on the Open Science Framework https://osf.io/x3hzt/?view_only=5b8e7f732ae74992aa4eefc32f18781f. Screening results were

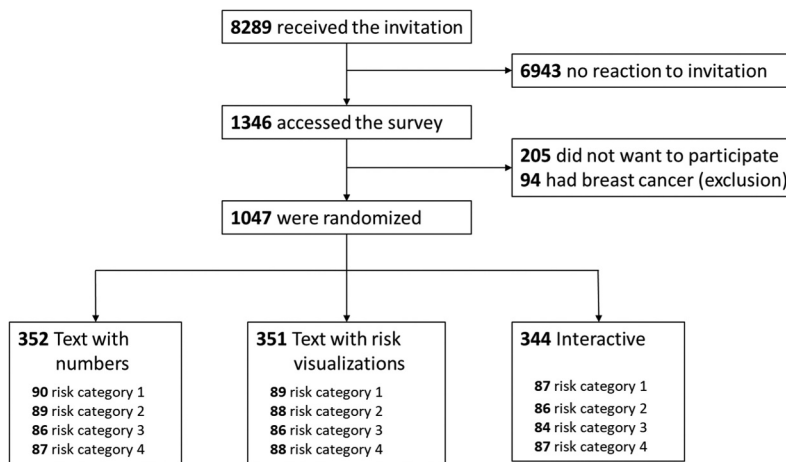


Figure 1. Flow diagram of survey participation and randomization.

presented in one of three formats: 1) text with numbers; 2) text with risk visualizations; and 3) interactive. For each format, four different breast cancer risk categories were created, reflecting the categories to which women could be hypothetically assigned (see Supplementary materials A Table A1 for an overview of the characteristics of the risk categories).

The “text with risk visualizations” format was developed in collaboration with a UX designer based on the risk communication literature, our previous research involving interviews, participatory design sessions, and user tests with the target group and expert consultation (van Strien-Knippenberg, Arjangi-Babetti, et al., 2024). After the user tests, a plain language specialist translated the information to a reading level of up to sixth grade. Prior to the current experiment, information materials were pre-tested with four women.

The numerical probability information, or quantified breast cancer risk, conveyed how many people in the relevant risk group would develop breast cancer in the coming 5 years and how this compared to the average risk of breast cancer. In the “text with numbers” format, this information was presented as numbers, communicating both the number of women expected to develop breast cancer and those not expected to develop breast cancer, as well as the difference with the average risk category. In the “text with risk visualizations” this numerical information was accompanied by an icon array using so-called “restroom icons,” i.e., icons that represent schematic male or female figures. The icon array consisted of 1,000 figures indicating how many women would develop breast cancer in the next 5 years and how many would not. The difference with the average risk in the icon array was also indicated. The “text with risk visualizations” format also included visualizations depicting qualitative elements, including that no abnormality was found, the risk factors, and the screening plan, i.e., the time period until the next invitation and which screening method will then apply. The communication goal of these visualizations was to enhance understanding and give meaning to the information. Figure 2 shows an example of the icon array and visualizations of qualitative risk information.

The “interactive” format was developed by a UX designer from our research team, using the “text with risk visualizations” as a starting point. The interactive format included

a dynamic display of the icon array whereby 1,000 icons were shown first. After pressing a button, the time frame, the number of icons affected by breast cancer, and the difference from the average breast cancer risk were displayed. Furthermore, participants only saw basic information on their screen and could decide whether they wanted to view further explanations by pressing a button. For example, for each risk factor, participants could decide whether they wanted more explanation about this risk factor. This information was visible by default in the “text with numbers” and “text with risk visualizations” format.

Procedure

Participants received a link to an online survey from the panel, with information about the study aim. Consenting participants were exposed to the stimulus materials followed by a questionnaire assessing outcome measures (see measures). The information materials were not visible to participants for the comprehension questions, but they were visible when answering the other questions. Finally, women’s GL (Okan et al., 2019), numeracy (Lipkus et al., 2001; Schwartz et al., 1997), and HL (Pander Maat et al., 2014; Woudstra et al., 2019) were assessed. Participants were thanked and rewarded according to the panel’s agreements.

Measures

Table 1 presents the questions used to assess all primary and secondary outcomes. The answers considered correct for the comprehension questions are provided in Supplementary material A, Tables A2 and A3. “Don’t know” answers were considered incorrect. The background characteristics measured were: age, education level (highest diploma obtained), medical education, awareness of BC screening, family history of BC, prior invitation to BC screening, and BC screening participation. All questions were reviewed by a plain language specialist and pre-tested among four women.

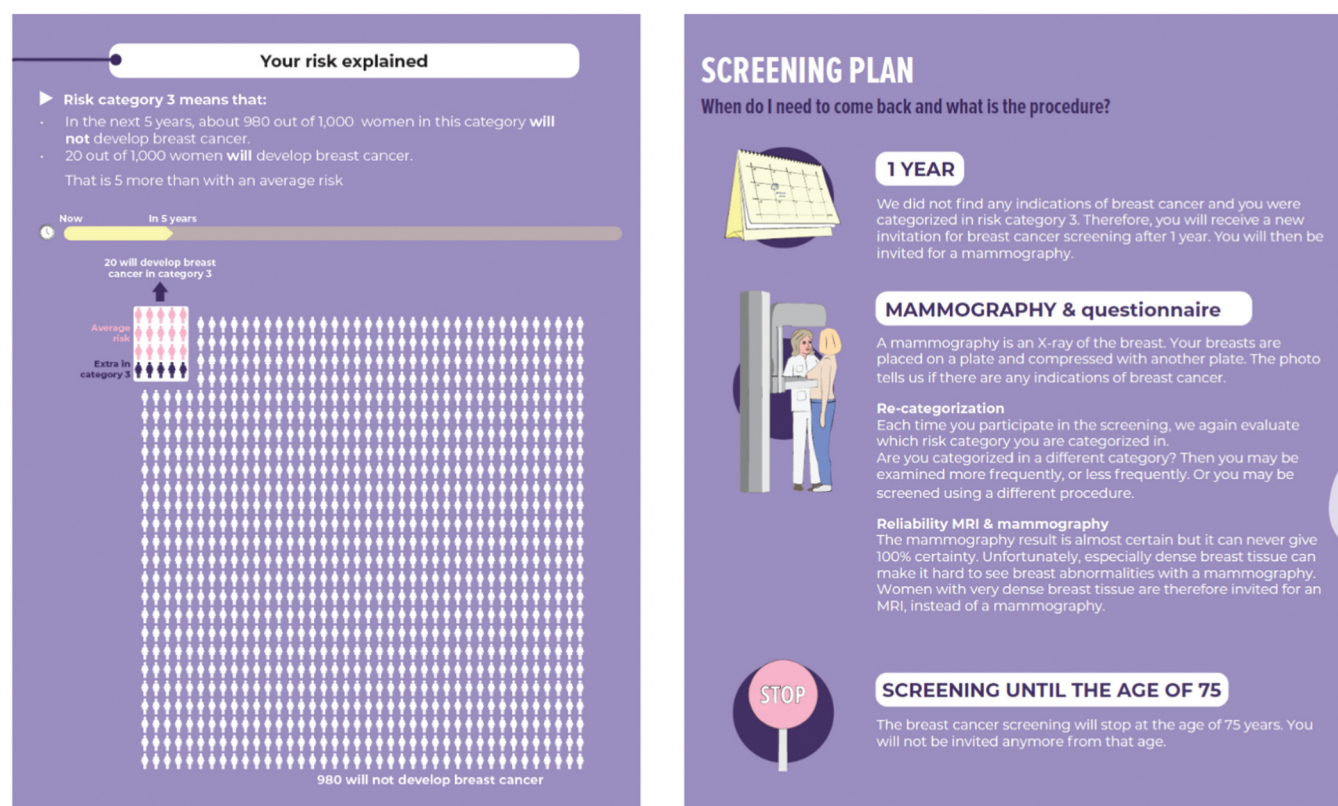


Figure 2. Example of a static quantitative risk visualization (left panel) and a qualitative risk visualization (right panel) for the third risk category.

Primary outcomes

We used a set of measures related to what is often called “gist comprehension” and “verbatim comprehension” in the risk communication and decision-making literature. This distinction stems from Fuzzy Trace Theory (Reyna, 2008), which postulates that gist comprehension -the underlying bottom-line meaning of the information- is especially important for decision making, and that verbatim comprehension -the details of the information- are important for extracting the right gist.⁴

Verbatim comprehension of numerical probability information (used for H1a). Two open-ended questions to assess whether participants remembered the exact absolute risk corresponding to their risk category as well as the exact difference from the average risk, e.g., “The chance of developing breast cancer within 5 years for my risk category is . . . out of 1,000 women.” Items were based on previous research measuring verbatim comprehension (e.g., Hawley et al., 2008; Tait et al., 2010) and tailored to the context of the current study. Answers were coded as correct when exactly correct (1 = correct; 0 = incorrect) and summed to form a composite score, with higher values indicating better comprehension.

Gist comprehension of numerical probability information (used for H1b). Two multiple-choice questions assessing understanding of the bottom line meaning of the assigned risk category relative to the average risk, e.g., “What does the result mean? The result means that my chances of developing breast cancer within 5 years are less than/the sameas/greater

than/I don’t know the chances of an average woman.” Answers were coded as 1 (correct) or 0 (incorrect) and summed, with higher scores representing better comprehension. Using items with one correct response builds on previous work using a similar approach to assess gist comprehension (Blalock & Reyna, 2016; Reyna & Brust-Renck, 2020; Smith et al., 2015).

Verbatim comprehension of the qualitative risk information (used for H2). Six multiple-choice questions related to the different risk concepts (e.g., whether an abnormality was found, the risk category, the corresponding screening plan and method, and risk factors). For example, “What was in the screening plan? In 1/2/4 year I will receive an invitation for breast cancer population screening/I don’t know.” Answers were coded as 1 (correct) or 0 (incorrect) and summed, with higher scores representing better comprehension.

Secondary outcomes

Affective reaction. 10 items of the Short PANAS (Kercher, 1992) to measure participant’s affective reaction after reading the results leaflet. Half of the items measure Positive Affect (PA) and half Negative Affect (NA) on a 5-pointscale (1 = very slightly or not at all, 5 = extremely). Two items were added based on previous research (van Strien-Knippenberg et al., 2022), i.e., worried and overwhelmed.

Perceived likelihood⁵. Two items adapted from Ockhuysen-Vermey et al. (2008) reflecting a cognitive evaluation of the risk information provided (7-pointscale). For example “After

Table 1. Questions associated with the primary and secondary outcomes.

	Variable	Item(s)	Scale/Response categories
Primary outcomes	Verbatim comprehension of numerical probability information	- The chance of developing breast cancer within 5 years for my risk category is - The results leaflet states how many out of 1000 women will develop breast cancer within 5 years at an average risk. What is the difference with YOUR risk group?	... out of 1000 women/I don't know ... more/fewer/as many women in my risk group will develop breast cancer compared to the average risk/I don't know
	Gist comprehension of numerical probability information	- What does the result mean? The result means that my chances of developing breast cancer within 5 years are the chances of an average woman - Anna is classified in risk group 2 (average risk of developing breast cancer within 5 years). Compare Anna's risk with your risk in the results leaflet. Who has the greatest risk?	less than/the same as/greater than/I don't know Anna/I/Our risk is the same/I don't know
	Verbatim comprehension of the qualitative risk information	- What does the result say? - In which risk category are you classified? - Do you have very dense breast tissue according to the result? - What was in the screening plan? - What else was in the screening plan? I will receive an invitation for: - Which factors are related to getting breast cancer? You can tick multiple answers	An abnormality was found/No abnormality was found/I don't know Risk category 1/2 (average risk)/3/4/I don't know Yes/No/I don't know In 1/2/4 years I will receive an invitation for the breast cancer population screening/I don't know A mammography/an MRI/I don't know An unhealthy weight/smoking/never been pregnant/using deodorant/big breasts/first menstruating period at a young age/dense breast tissue/a higher age/wearing an underwired bra/breast cancer in your family
Secondary outcomes	Affective reaction	Below are some words that describe different feelings and emotions. We want to know how you are feeling after reading the information. Please read each word and indicate to what extent you feel this way.	1 (very slightly or not at all) – 5 (extremely)
		- Alert (PA) - Excited (PA) - Enthusiastic (PA) - Nervous (NA) - Inspired (PA) - Distressed (NA) - Determined (PA) - Upset (NA) - Scared (NA) - Afraid (NA) - Worried - Overwhelmed	
	Perceived likelihood	- After reading the result how likely do you think it is that you will develop breast cancer in the next 5 years? - What do you think is the risk of developing breast cancer in the next 5 years?	1 (very unlikely) – 7 (very likely) 1 (very small risk) – 7 (very big risk)
	Affective risk appraisal	- After reading the result, to what extent are you worried about the risk of developing breast cancer in the next 5 years? - After reading the result, to what extent do you feel anxious about the risk of developing breast cancer in the next 5 years?	1 (not worried at all) – 7 (very worried) 1 (not anxious at all) – 7 (very anxious)
	Evaluation of information material	What do you think of the way the results leaflet is presented? - Enjoyable - Creative - Clearly - Pretty - Understandable	1 (strongly disagree) – 7 (strongly agree)

reading the result, how likely do you think it is that you will develop breast cancer in the next 5 years?"

Affective risk appraisal. Two items assessing worry and anxiety about breastcancer risk, adapted from Cameron and Reeve (2006), for example, "After reading the result, to what extent are you worried about the risk of developing breast cancer in the next 5 years?" Items were rated on a 7-point scale.

Evaluation of information material. Five items of the subscales attractiveness and comprehensibility of the Website Satisfaction Scale adapted for the present study (WSS; 7-point scale ranging from strongly disagree to strongly agree) (Bol et al., 2014). An

example is "understandable" in response to the question "What do you think of the way the results leaflet is presented?"

Individual differences

Numeracy was assessed using Schwartz et al.'s 3-item scale (Schwartz et al., 1997), with the first question formulated as in Lipkus et al. (2001). Health literacy was assessed with the 13 multiple-choice items of the SAHL-D (Woudstra et al., 2019). Graph literacy was measured with the four items of the Short Graph Literacy (SGL) Scale (Okan et al., 2019). Answers were coded as 1 (correct) or 0 (incorrect) and summed for each variable, with higher scores representing higher numeracy/HL/GL skills.

Attention and realism check

Effort⁶. One question to assess how carefully participants looked at the information materials (7-point scale ranging from very carefully to not at all carefully). “Have you looked carefully at the result folder?”

Time spent⁴. Registration of the time women spent reading the stimulus material.

Realism check. One item based on Gudde Kuiper et al. (2012) assessing how well someone could identify with the information (7-point scale ranging from totally disagree to totally agree). “I could well pretend that this result folder was about me.”

Statistical analysis

Analyses with comprehension as outcome were performed with cumulative odds ordinal logistic regression with proportional odds due to the ordinal nature of the comprehension variables with limited possible outcomes (i.e., ranging from 0–3 and 0–7)⁷. Assumption testing of proportional odds was tested by the full likelihood ratio test comparing the fit of the proportional odds model to a model with varying location parameters. Multicollinearity was assessed by looking at the condition index. To assess the effects of HL, numeracy, and GL, models with one of these variables were compared with models with all three variables included, i.e., nested models were used. The model with the lowest Akaike Information Criterion (AIC) was reported. Also, models with interaction terms were performed. When not significant, interactions were removed from the final model. Secondary outcomes were analyzed using two-way ANOVAs with information format and risk category as independent variables. Analyses were

performed using SPSS version 26 and significance levels were set at $p < .05$.

Results

Sample characteristics

A total of 1,047 women participated (age 58.2 ± 9.3 years), of whom 431 (39.4%) had a lower educational level. Participant characteristics are displayed in Table 2.

Effects of visualizations on verbatim comprehension of numerical probability information (H1a, RQ1-3)

The number of correct answers to the individual items assessing comprehension of numerical probability information is provided in Supplementary materials A Table A2. The final ordinal logistic regression model with verbatim comprehension of numerical probability information as dependent variable included information format, assigned risk category, educational level, and numeracy. No interaction effects were found. The model without interactions significantly predicted comprehension over and above the intercept-only model $\chi^2(7) = 216.68$, $p < .001$. The assumption of proportional odds was met, $\chi^2(7) = 2.71$, $p = .910$. Information format had a statistically significant effect on respondents' verbatim comprehension of numerical probability information, Wald $\chi^2(2) = 21.85$, $p < .001$. Contrary to our hypothesis, there were no significant differences between those exposed to the “text with risk visualizations” format and the “text with numbers” format, Wald $\chi^2(1) = 2.66$, $p = .103$. The results of the analysis are displayed in Table 3 and Supplementary materials B.

Table 2. Participant characteristics per information format.

Characteristic	Text with numbers (<i>n</i> = 352)	Text with risk visualizations (<i>n</i> = 351)	Interactive (<i>n</i> = 344)	Total (<i>n</i> = 1,047)
Age (years), mean $\pm$ SD	57.4 $\pm$ 9.1	59.4 $\pm$ 9.5	57.7 $\pm$ 9.2	58.2 $\pm$ 9.3
Education ^a , <i>n</i> (%)				
Low	137 (38.9%)	140 (39.9%)	136 (39.5%)	413 (39.4%)
Middle/high	215 (61.1%)	211 (60.1%)	208 (60.5%)	634 (60.6%)
Medical education, <i>n</i> (%)	33 (9.4%)	39 (11.1%)	29 (8.4%)	101 (9.6%)
Heard of BC screening, <i>n</i> (%)	346 (98.3%)	349 (99.4%)	338 (98.3%)	1033 (98.7%)
Family history BC, <i>n</i> (%)				
Yes	96 (27.3%)	89 (25.4%)	84 (24.4%)	269 (25.7%)
No	234 (66.5%)	240 (68.4%)	238 (69.2%)	712 (68.0%)
Don't know	22 (6.3%)	22 (6.3%)	22 (6.4%)	66 (6.3%)
Ever had an invitation, <i>n</i> (%)				
Yes	266 (75.6%)	275 (78.3%)	281 (81.7%)	822 (78.5%)
No	76 (21.6%)	72 (20.5%)	62 (18.0%)	210 (20.1%)
Don't know	10 (2.8%)	4 (1.1%)	1 (0.3%)	15 (1.4%)
Participation in BC screening (when invited), <i>n</i> (%)				
Yes	249 (93.6%)	263 (95.6%)	256 (91.1%)	768 (93.4%)
No	17 (6.4%)	12 (4.4%)	23 (8.2%)	52 (6.3%)
Don't know	—	—	2 (0.7%)	2 (0.2%)
Health literacy (HL), <i>M</i> $\pm$ <i>SD</i> [range 0 – 13] ^b	9.74 $\pm$ 2.8	9.91 $\pm$ 2.8	9.82 $\pm$ 2.9	9.82 $\pm$ 2.8
Numeracy, <i>M</i> $\pm$ <i>SD</i> [range 0 – 3] ^b	1.62 $\pm$ 1.1	1.55 $\pm$ 1.0	1.54 $\pm$ 1.1	1.57 $\pm$ 1.1
Graph literacy (GL), <i>M</i> $\pm$ <i>SD</i> [range 0 – 4] ^b	1.97 $\pm$ 1.1	1.92 $\pm$ 1.0	1.81 $\pm$ 1.1	1.90 $\pm$ 1.1

Abbreviation: BC breast cancer.

^alow education is a maximum of a pre-vocational secondary education degree; high education is a minimum of a secondary vocational education degree (Centraal Bureau voor de Statistiek, 2021).

^bHigher scores represent higher HL/numeracy/GL.

Interactive format (exploratory analysis)

Those exposed to the “interactive” format exhibited worse verbatim comprehension of the numerical probability information compared to those exposed to the “text with numbers” format, OR .48 (95% CI [.35, .65]), Wald $\chi^2(1) = 21.49, p < .001$.

The influence of the assigned risk category and numeracy

The assigned risk category had a statistically significant effect, Wald $\chi^2(3) = 46.60, p < .001$. Women assigned to risk categories 1, 3, or 4 (i.e., slightly decreased, slightly increased, or increased risk), exhibited worse verbatim comprehension of the numerical probability information compared to women assigned to risk category 2 (average risk). An increase in numeracy was associated with an increase in the odds of having a better comprehension, OR 2.03 (95% CI [1.78, 2.32]), Wald $\chi^2(1) = 109.84, p < .001$.

Effects of visualizations on gist comprehension of numerical probability information (H1b, RQ1-3)

The final ordinal logistic regression model with gist comprehension of numerical probability information as dependent variable included information format, assigned risk category, educational level, numeracy, and an interaction effect between numeracy and risk category. This model significantly predicted the gist comprehension of numerical probability information over and above the intercept-only model $\chi^2(10) = 265.00, p < .001$. The assumption of proportional odds was not met, $\chi^2(10) = 51.04, p < .001$. The results for the third and fourth risk categories should be interpreted with caution (for details see Table 3). Contrary to our hypothesis, information format did not affect respondents' gist comprehension of numerical probability information, Wald $\chi^2(2) = .68, p = .710$.

Table 3. Results of the ordinal logistic regression models for the verbatim and gist comprehension of the numerical probability information and the comprehension of the qualitative risk information involved in the multifaceted risk message.

	OR	95% CI of OR	p-value
Verbatim comprehension of numerical probability information			
Information format [text format as reference]			
Text with risk visualizations	.78	.58–1.05	.103
Interactive	.48	.35–.65	<.001
Risk category [risk category 2 as reference]			
Risk category 1	.34	.24–.48	<.001
Risk category 3	.42	.30–.60	<.001
Risk category 4	.40	.28–.57	<.001
Numeracy	2.03	1.78–2.32	<.001
Education level [middle/high as reference]			
Low	.72	.55–.95	.019
Gist comprehension of numerical probability information			
Information format [text format as reference]			
Text with risk visualizations	1.04	.78–1.39	.806
Interactive	.92	.69–1.23	.572
Risk category [risk category 2 as reference] ^a			
Risk category 1	.37	.20–.68	.001
Risk category 3	.09	.05–.17	<.001
Risk category 4	.08	.04–.15	<.001
Numeracy	1.05	.84–1.31	.684
Numeracy * Risk category [risk category 2 as reference]			
Numeracy * Risk category 1	1.69	1.22–2.32	.001
Numeracy * Risk category 3	2.12	1.53–2.92	<.001
Numeracy * Risk category 4	2.09	1.50–2.91	<.001
Education level [middle/high as reference]			
Low	.58	.45–.75	<.001
Comprehension of the qualitative risk information involved in the multifaceted risk message			
Information format [text format as reference] ^b			
Text with risk visualizations	.77	.59–1.00	.052
Interactive	.51	.39–.66	<.001
Risk category [risk category 2 as reference] ^c			
Risk category 1	.76	.56–1.03	.080
Risk category 3	.43	.32–.59	<.001
Risk category 4	.30	.22–.40	<.001
Numeracy	1.56	1.40–1.74	<.001
Education level [middle/high as reference]			
Low	.61	.48–.77	<.001

Note. OR = Odds Ratio; CI = Confidence Interval.

^aThe assumption of proportional odds was not met, as assessed by a full likelihood ratio test comparing the fit of the proportional odds model to a model with varying location parameters, $\chi^2(10) = 51.04, p < .001$. Interpretations of separate binomial logistic regression on cumulative dichotomous dependent variables showed that the odds ratios for the binomial logistic regression may not be tenable for the third risk category (i.e., 3.47 and 8.48) and fourth risk category (i.e., .24 and .56). Therefore, the results for these risk categories should be interpreted with caution.

^bThe assumption of proportional odds was not met, as assessed by a full likelihood ratio test comparing the fit of the proportional odds model to a model with varying location parameters, $\chi^2(35) = 90.54, p < .001$. Interpretations of separate binomial logistic regression on cumulative dichotomous dependent variables showed that the odds ratios for the binomial logistic regression may not be tenable for the information format text with risk visualizations (i.e., lowest odds ratio .34 and highest .85) and the interactive format (i.e., lowest odds ratio .22 and highest .89). Therefore, these results should be interpreted with caution.

^cThe odds ratios for the binomial logistic regression may not be tenable for the first risk category (i.e., lowest odds ratio .60 and highest 2.03), third risk category (i.e., lowest odds ratio .75 and highest 4.36), and fourth risk category (i.e., lowest odds ratio .26 and highest 1.22). Therefore, these results should be interpreted with caution.

Interactive format (exploratory analysis)

There were no significant differences between those exposed to the “interactive” format and those exposed to the “text with numbers” format.

The influence of the assigned risk category and numeracy

The assigned risk category had a statistically significant effect, Wald $\chi^2(3) = 85.54$, $p < .001$. Women assigned to risk categories 1, 3, or 4 (i.e., slightly decreased, slightly increased, or increased risk), exhibited worse gist comprehension of the numerical probability information compared to women assigned to risk category 2 (average risk). There was a significant interaction between numeracy and risk category, Wald $\chi^2(3) = 27.35$, $p < .001$. The interaction showed that with a higher level of numeracy, comprehension was higher for the first, third and fourth risk categories than expected based on the main effects of numeracy and risk category.

Effects of visualizations on verbatim comprehension of the qualitative risk information (H2, RQ1-3)

The number of correct answers to the individual items assessing verbatim comprehension of qualitative risk information is provided in Supplementary materials A Table A3. The final ordinal logistic regression model with verbatim comprehension of qualitative risk information as dependent variable included information format, assigned risk category, educational level, and numeracy. No interaction effects were found. The model without interactions significantly predicted the verbatim comprehension of the qualitative risk information over and above the intercept-only model $\chi^2(7) = 204.31$, $p < .001$. The results of this model are displayed in Table 3 and Supplementary materials B Figure B3. The assumption of proportional odds was not met, $\chi^2(35) = 90.54$, $p < .001$. The results for the text with risk visualizations and the interactive information formats should be interpreted with caution just as the results for the first, third, and fourth risk categories (for details see Table 3). Information format had a statistically significant effect on respondents' comprehension of the qualitative risk information, Wald $\chi^2(2) = 25.10$, $p < .001$. Contrary to our expectations, there were no significant differences between the “text with risk visualizations” format and the “text with numbers” format, Wald $\chi^2(1) = 3.77$, $p = .052$.

Interactive format (exploratory analysis)

Those exposed to the “interactive” format showed worse comprehension compared to those exposed to the “text with numbers” format, OR .51 (95% CI [.39, .66]), Wald $\chi^2(1) = 24.70$, $p < .001$.

The influence of the assigned risk categories and numeracy

The assigned risk category had a statistically significant effect, Wald $\chi^2(3) = 70.03$, $p < .001$. Women assigned to risk categories 3 or 4 (i.e., slightly increased or increased risk), exhibited worse comprehension compared to women assigned to risk category 2 (average risk). An increase in numeracy was further associated with an increase in the odds of having a higher comprehension, with an odds ratio of 1.56 (95% CI [1.40, 1.74]), Wald $\chi^2(1) = 63.32$, $p < .001$.

Secondary outcomes (RQ4)

No main effects of format and assigned risk category or interactions between these factors were found on respondents' affective reaction.

There was a main effect of risk category on perceived likelihood, $F(3, 1035) = 19.10$, $p < .001$, partial $\eta^2 = .05$. Pairwise comparisons showed a significant difference between all risk categories, except between risk categories 3 and 4 (i.e., slightly increased and increased risk). Perceived likelihood was lowest for those in risk category 1 ($2.83 \pm .08$) followed by the average risk category ($3.24 \pm .08$), risk category 4 ($3.58 \pm .08$), and 3 ($3.59 \pm .08$).

We also found a main effect of the assigned risk category on affective risk appraisal, $F(3, 1035) = 23.08$, $p < .001$, partial $\eta^2 = .06$. Pairwise comparisons showed a significant difference between all risk categories, except between risk categories 3 and 4. Risk appraisal was lowest for those in risk category 1 ($2.73 \pm .09$) followed by risk category 2 ($3.12 \pm .09$), 3 ($3.51 \pm .09$), and 4 ($3.71 \pm .09$).

There was a main effect of format on the evaluation of information material, $F(2, 1035) = 4.20$, $p = .015$, partial $\eta^2 = .01$. Pairwise comparisons showed a more positive evaluation for the “interactive” format ($5.58 \pm .06$) compared to the “text with numbers” format ($5.33 \pm .06$), $p = .012$.

Attention and realism check

No main effects of format and assigned risk category or interactions between these factors were found on respondents' effort (smallest p -value = .25), suggesting that self-reported carefulness in reviewing the materials did not differ significantly across experimental conditions.

Similarly, the time spent reading the stimulus materials did not differ significantly across presentation formats or assigned risk categories (smallest p -value = .32). However, there were extreme outliers, defined as values lying more than 3xIQR above the third quartile or below the first quartile. After removing these outliers there was a main effect of format, $F(2, 955) = 4.430$, $p = .012$, partial $\eta^2 = .009$. Time spent was lowest in the interactive format (232.01 ± 6.60) compared to the text with risk visualizations format (258.15 ± 6.45). There were no differences in time spent between the text with numbers format (238.08 ± 6.50) and the other two formats.

There was a main effect of risk category on the realism check, $F(3, 1035) = 4.510$, $p < .005$, partial $\eta^2 = .013$. Pairwise comparisons showed that participants reported being less able to identify with the information presented for risk category 4 ($5.49 \pm .08$) as compared to risk category 1 ($5.84 \pm .08$) and risk category 2 ($5.85 \pm .08$).

A correlation matrix of the variables treated as continuous variables is included in Supplementary materials C Table C2.

Discussion

In this experimental study, we investigated whether visualizations of quantitative and qualitative risk information could improve women's understanding of information about hypothetical risk-based breast cancer screening results,

relative to a text format with numerical probability information in a frequency format, with all formats designed based on evidence-based principles. Specifically, we examined the effect of visualizing quantified breast cancer risk using an icon array and visualizing more “qualitative” risk information using simplified drawings co-created with women in a previous study. The impact of different risk categories was also investigated. We found that the tested visualizations did not improve comprehension of the information. Moreover, exploratory analyses revealed that the interactive format hindered verbatim comprehension of numerical probability information and qualitative risk information. Our results also showed that women assigned to risk categories other than the average risk category exhibited worse comprehension.

A potential explanation for the absence of a significant effect of icon arrays representing breast cancer risk on understanding numerical probability information may be that our “text with numbers” format was also designed using evidence-based principles for communicating numerical information. For example, the “text with numbers” format used an X per 1,000 format with a 5-year reference class (Trevena et al., 2021) and aimed to reduce cognitive effort by explicitly stating the difference between the assigned and the average risk category (Bonner et al., 2021). Previous studies have shown that icon arrays can facilitate comprehension by making part-to-whole relationships, calculations, and comparisons visible (Ancker et al., 2006; Garcia-Retamero & Cokely, 2017; Stone et al., 2018). Since we also included such elements in the “text with numbers” format, e.g., making part-to-whole relationships clear –albeit without a visualization–, there may be no additional benefit from adding an icon array. The structured presentation of numbers in this format may also have been important. This would be in line with previous research on prostate cancer screening where no increase in comprehension was found between presenting structured information in so-called “fact boxes” and adding icon arrays to this information (McDowell et al., 2019), and research on presenting survival rates for breast cancer treatments, where visual formats had no added value on comprehension compared to a well-designed structured numerical textual format (van Strien-Knippenberg, Timmermans, et al., 2024). Similarly, a study among counselees with a family history of breast cancer found no difference in comprehension between women presented with their breast cancer risk in a frequency format with or without an icon array (Henneman et al., 2013). Yet, it should be noted that the way the visualizations were designed could also have influenced the results. Although we co-created the information with women from the target group, in collaboration with a UX designer, and based on evidence-based principles and guidelines for communicating numerical probability information, it is possible that a different representation of the icon array would have yielded different results.

Another explanation may relate to the amount and potential complexity of the information in our stimulus materials. Recall of the numerical probability information was not optimal in all formats. Almost 40% of women answered both gist comprehension questions correctly and an average of 13% answered both verbatim comprehension questions correctly.

This may be because breast cancer screening, like most other cancer screening, involves small absolute disease risks, known to be difficult to understand (Ancker et al., 2006). Furthermore, in the context of breast cancer screening, women are known to expect higher absolute probabilities (Buxton et al., 2003; de Jonge et al., 2009; Jones et al., 2011), possibly hindering comprehension. In addition, it may well be that visualizations have positive effects when larger probabilities are involved (Trevena et al., 2021) because larger probabilities can be displayed using icon arrays with a smaller total number of icons (e.g., 10 or 100 instead of 1,000). For example, research has shown that a 2% probability displayed in an icon array influences people’s risk perception, but a 0.2% probability displayed in an icon array does not (Fraenkel et al., 2018). At the same time, another study found that the effects of visualizations displaying only numerators (vs. both numerators and denominators) on increasing risk perception and risk aversion were reduced for larger probabilities (Okan et al., 2020). Further research is needed on the role of probability size on outcomes including comprehension and risk perception across different numerical and visual displays. Another aspect is that the icon array was added to textual information, increasing the amount of information to process and possibly offsetting any positive effect on comprehension. The current study contributes to the growing literature indicating that icon arrays cannot be assumed to consistently enhance comprehension. Their effectiveness seems to depend on specific design features, the cognitive demands of the task, and the broader context in which they are presented.

Concerning visualizations of more “qualitative” risk information, prior studies show that they can support comprehension of medical information (Brotherstone et al., 2006; Schubbe et al., 2020). In our study, however, visualizing, for example, the screening result, risk factors, and the screening plan did not increase comprehension. The amount and complexity of the information may have also contributed, as comprehension was generally low across formats, with an average of 3.75 correct answers out of 6 for verbatim comprehension of qualitative risk information.

In the “interactive” format, the icon array was presented dynamically and initially only basic information was displayed on the screen, with respondents being able to choose whether to see additional information by pressing a button. This format was associated with worse verbatim comprehension of the numerical probability information compared to the “text with numbers” format. A recent secondary analysis of a systematic review showed that there is no strong evidence that interactivity or animation is more effective than static visualizations in supporting comprehension of numerical probability information (Ancker et al., 2024b). However, there is some evidence that adding dynamic elements specifically designed to promote active information processing can enhance risk comprehension (Okan et al., 2015). One possible explanation for the absence of a positive effect of our dynamic icon array on comprehension is that the different elements appeared on the screen quickly one after the other. As a result, the respondents’ attention may not have been sufficiently focused on the various components. A possible explanation for the detrimental impact of the interactive format on

comprehension of qualitative risk information is that the “interactive” format may have required additional cognitive effort in deciding whether or not to click for more details, which may have hampered comprehension. Furthermore, after removing outliers, defined as values lying more than 3xIQR above the third quartile or below the first quartile, participants spent less time on the interactive format than on the text with risk visualization format. However, self-reported carefulness in reviewing the materials did not differ between formats.

Although our study suggests that interactive formats may not necessarily be well-suited to support understanding, it is worth noting that studies focusing on the same target group show that women wanted layered information when receiving results of risk-based breast cancer screening (van Strien-Knippenberg, Arjangi-Babetti, et al., 2024). In the current study, people also had a more positive evaluation of the “interactive” than the “text with numbers” format, despite poorer performance with the “interactive” format. This mismatch between user evaluations and understanding is in line with previous literature showing that people’s preferences for presentation formats are not always aligned with comprehension (Gaissmaier et al., 2023; Garcia-Retamero & Cokely, 2017).

The hypothetical risk category to which people were assigned also affected information comprehension, with those assigned to a risk category deviating from the average breast cancer risk exhibiting a worse gist and verbatim comprehension of the numerical probability information. For gist comprehension, this effect was more prevalent for those with lower numeracy, which may be related to the fact that more numerical information had to be processed in the risk categories deviating from the average. However, women’s likelihood perceptions and affective risk appraisal were lowest among those in the below-average risk category and highest in the above-average risk categories, suggesting they did grasp the meaning of the risk categories to some extent. Participants assigned to above-average risk categories exhibited worse comprehension of qualitative risk information than those assigned to the average risk category. These differences may be related to the fact that the information for the average risk category most closely matches what is currently provided in population-based breast cancer screening. Perhaps an information leaflet is not sufficient to help women fully understand the concept of risk-based screening, including risk factors and risk category assignment. For example, previous research in the context of risk-based screening has shown that those assigned to risk categories higher than the average risk would like to get in touch with a healthcare provider (Rainey et al., 2020; Taylor et al., 2023; van Strien-Knippenberg, Arjangi-Babetti, et al., 2024).

Limitations and future research

A study strength is that we investigated visualizations of both quantitative and qualitative risk information in one total package of information, resembling information provision in breast cancer screening in practice (e.g., RIVM, 2025). This similarity with practice contributes to the ecological validity of the study. The visualizations were compared to a textual format with numerical probability information developed in co-creation with women from the target group and based on evidence-

based principles. Comprehension of both the numerical probability information and qualitative risk information was assessed, although it should be noted that the combined use of the visualizations of quantified risk and more qualitative risk information prevents us from examining the effect of the icon array and qualitative risk visualizations separately. Future research could focus on the separate effects of both visualizations in risk-based screening. In doing so, it is important to use evidence-based design principles and communication guidelines and to have the visualizations designed by a professional designer. For visualizations of qualitative risk information, it would also be interesting to investigate the impact of designs that reduce the amount of textual information, as visualizations requiring little textual explanation appear to be more effective than those requiring more textual explanation (Schubbe et al., 2020).

Concerning the use of “interactive” formats, future research could focus on what information should be displayed by default and what should be provided as additional information. Subsequently, it could be investigated what additional information is actually viewed and in what order, for example by registering how people click through the information or by eye-tracking or think-aloud interviews. In addition, future research into “interactive” formats should explicitly focus on how such formats can best promote active information processing.

Regarding the concept of information comprehension, it is important to note that what comprehension entails varies in risk communication and decision making research, has been measured with different measures, and can involve different tasks (Ancker et al., 2022). For example, what we call “verbatim comprehension” in this study could also be referred to as “recall” and what we call “gist comprehension” as “cognitive activities” (Ancker et al., 2022). It should also be noted that verbatim comprehension of the qualitative risk information was measured in this study with multiple choice questions, whereas verbatim comprehension of the numerical probability information was assessed with open-ended questions. Gist comprehension of numerical probability was measured with two multiple-choice questions with one answer coded as correct for each item. While several studies used a similar approach (e.g., Blalock & Reyna, 2016; Reyna & Brust-Renck, 2020; Smith et al., 2015), others assessed gist comprehension differently, for example, using statements where people rate their agreement a given statement on a scale (e.g., Reyna & Brainerd, 2011), or by accepting a specified range around the exact number when coding responses (Okan et al., 2025). Gist comprehension of the qualitative risk information was not measured, due to the fact that there is currently no paradigmatic way to do so. Future research could benefit from the development of more consistent measures of gist comprehension and its various components, such as categorical representations and ordinal representations of relative magnitudes (Brust-Renck & Reyna, 2023). In this regard, the taxonomy of outcome measures proposed by Ancker et al. (2022) should also be considered to promote consistency in naming and measurement across studies.

In this study, various interaction effects with numeracy were also investigated. However, this study may not have

been sufficiently powered to investigate the three-way interaction information format $\times$ risk category $\times$ numeracy. Moreover, initial power calculations were based on factorial ANOVA designs, while ordinal logistic regression analyses were performed. This may have affected the power.

Another strength of this study is that we used information materials based on how risk-based breast cancer screening might look in the future, although this implied that women did not receive their actual risk, but a hypothetical risk. Another limitation concerns the sample's representativeness, as approximately 25% reported a family history of breast cancer. However, sensitivity analyses excluding all women who reported a family history of breast cancer yielded a similar pattern of results, see Supplementary materials C, Table C1. It should also be noted that more than 90% of the participants had participated in breast cancer screening, while the national average in the Netherlands is just over 70% (Bevolkingsonderzoek Nederland, 2024).

Conclusion

The visualizations of risk information used in this study, consisting of both an icon array and visualizations of qualitative risk information, did not significantly enhance comprehension of risk-based breast cancer screening results compared to a text format with numerical probability information. All formats used in this study were developed based on evidence-based design principles and guidelines for communicating numerical probability information. Presenting the information in an “interactive” format hindered comprehension. Regardless of the format, comprehension of risk-based screening results seems to be difficult. Women assigned to a risk category other than the average risk category exhibited worse comprehension of the information compared to women assigned to the average risk category.

Notes

1. An icon array is a visual representation of numerical information using repeated icons, where subgroups are differentiated by varying colors to depict a ratio.
2. A pictogram refers to a pictorial symbol for a word or phrase.
3. Instead of ANOVAs, ordinal logistic regression analyses were performed which may have affected the power of the analyses.
4. Within the risk communication literature, comprehension is a broad concept that is measured with different measures and can involve different tasks. What we call “verbatim comprehension” - measured when participants no longer saw the stimulus materials - has also been called “recall.” What we call “gist comprehension” has also been classified under “cognitive activities” in a recent taxonomy of outcome measures (Ancker et al., 2022).
5. Referred to in the pre-registration as “cognitive risk appraisal.”
6. Referred to in the pre-registration as “attention check.”
7. This approach deviates from our pre-registration, which outlined the use of ANOVAs. Following further statistical consultation, we opted to analyze the comprehension variables using a more appropriate approach that preserves their ordinal nature, rather than treating them as interval data.

Acknowledgements

We would like to thank our institutional statisticians for their advice and support in the analyses. Also, all women who participated in the pretest or the experiment are thanked for their participation.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

The work was supported by the ZonMw [50–53125-98–163]; Yasmina Okan's time was funded by a Ramón y Cajal Fellowship [RYC2020-028857-I] financed by MICIU/AEI/10.13039/501100011033 and by The FSE Invests in your Future.

ORCID

Inge S. van Strien-Knippenberg  <http://orcid.org/0000-0002-0886-9224>
 Danielle R.M. Timmermans  <http://orcid.org/0000-0002-3602-3875>
 Yasmina Okan  <http://orcid.org/0000-0001-7963-1363>
 Carla H. van Gils  <http://orcid.org/0000-0003-0817-7567>
 Mireille J.M. Broeders  <http://orcid.org/0000-0002-8741-8148>
 Olga C. Damman  <http://orcid.org/0000-0002-4482-5042>

Data availability statement

The data that support the findings of this study are available on request from the corresponding author. Data cannot be shared publicly because the Ethics Committee requires that the collected data remain securely stored and not be shared publicly.

References

- Ancker, J. S., Benda, N. C., Sharma, M. M., Johnson, S. B., Weiner, S., & Zikmund-Fisher, B. J. (2022). Taxonomies for synthesizing the evidence on communicating numbers in health: Goals, format, and structure. *Risk Analysis*, 42(12), 2656–2670. <https://doi.org/10.1111/risa.13875>
- Ancker, J. S., Benda, N. C., & Zikmund-Fisher, B. J. (2024a). Do you want to promote recall, perceptions, or behavior? The best data visualization depends on the communication goal. *Journal of the American Medical Informatics Association*, 31(2), 525–530. <https://doi.org/10.1093/jamia/ocad137>
- Ancker, J. S., Benda, N. C., & Zikmund-Fisher, B. J. (2024b). Insufficient evidence for interactive or animated graphics for communicating probability. *Journal of the American Medical Informatics Association*, 31(11), 2760–2765. <https://doi.org/10.1093/jamia/ocae123>
- Ancker, J. S., Senathirajah, Y., Kukafka, R., & Starren, J. B. (2006). Design features of graphs in health risk communication: A systematic review. *Journal of the American Medical Informatics Association*, 13(6), 608–618. <https://doi.org/10.1197/jamia.M2115>
- Berkman, N. D., Sheridan, S. L., Donahue, K. E., Halpern, D. J., & Crotty, K. (2011). Low health literacy and health outcomes: An updated systematic review. *Annals of Internal Medicine*, 155(2), 97–107. <https://doi.org/10.7326/0003-4819-155-2-201107190-00005>
- Bevolkingsonderzoek Nederland. (2024, September). *Bevolkingsonderzoek borstkanker*. <https://www.bevolkingsonderzoeknederland.nl/borstkanker/>
- Blalock, S. J., & Reyna, V. F. (2016). Using fuzzy-trace theory to understand and improve health judgments, decisions, and behaviors: A literature review. *Health Psychology*, 35(8), 781–792. <https://doi.org/10.1037/hea0000384>
- Bol, N., van Weert, J. C., de Haes, H. C., Loos, E. F., de Heer, S., Sikkels, D., & Smets, E. M. (2014). Using cognitive and affective illustrations to

- enhance older adults' website satisfaction and recall of online cancer-related information. *Health Communication*, 29(7), 678–688. <https://doi.org/10.1080/10410236.2013.771560>
- Bonner, C., Trevena, L. J., Gaissmaier, W., Han, P. K. J., Okan, Y., Ozanne, E., Peters, E., Timmermans, D., & Zikmund-Fisher, B. J. (2021). Current best practice for presenting probabilities in patient decision aids: Fundamental principles. *Medical Decision Making*, 41(7), 821–833. <https://doi.org/10.1177/0272989X21996328>
- Bresciani, S., & Eppler, M. J. (2015). The pitfalls of visual representations: A review and classification of common errors made while designing and interpreting visualizations. *Sage Open*, 5(4). <https://doi.org/10.1177/2158244015611451>
- Broeders, M., & Paci, E. (2015). The balance sheet of benefits and harms of breast cancer population-based screening in Europe: Outcome research, practice and future challenges. *Women's Health*, 11(6), 883–890. <https://doi.org/10.2217/whe.15.34>
- Brotherstone, H., Miles, A., Robb, K. A., Atkin, W., & Wardle, J. (2006). The impact of illustrations on public understanding of the aim of cancer screening. *Patient Education & Counseling*, 63(3), 328–335. <https://doi.org/10.1016/j.pec.2006.03.016>
- Brust-Renck, P. G., & Reyna, V. F. (2023). Individual differences in numerical representations of risk in health decision making: A fuzzy-trace theory approach. *Risk Analysis*, 43(3), 548–557. <https://doi.org/10.1111/risa.13914>
- Buxton, J. A., Bottorff, J. L., Balneaves, L. G., Richardson, C., McCullum, M., Ratner, P. A., & Hack, T. (2003). Women's perceptions of breast cancer risk: Are they accurate? *Canadian Journal of Public Health*, 94(6), 422–426. <https://doi.org/10.1007/bf03405078>
- Cameron, L. D., & Reeve, J. (2006). Risk perceptions, worry, and attitudes about genetic testing for breast cancer susceptibility. *Psychology & Health*, 21(2), 211–230. <https://doi.org/10.1080/14768320500230318>
- Centraal Bureau voor de Statistiek. (2021). *Invulling praktisch en theoretisch opgeleiden*. <https://www.cbs.nl/nl-nl/longread/discussie-papers/2021/invulling-praktisch-en-theoretisch-opgeleiden/3-indeling-van-opleidingen-op-basis-van-niveau-en-oriëntatie>
- Clift, A. K., Dodwell, D., Lord, S., Petrou, S., Brady, S. M., Collins, G. S., & Hippisley-Cox, J. (2022). The current status of risk-stratified breast screening. *British Journal of Cancer*, 126(4), 533–550. <https://doi.org/10.1038/s41416-021-01550-3>
- Davis, T. C., Williams, M. V., Marin, E., Parker, R. M., & Glass, J. (2002). Health literacy and cancer communication. *CA: A Cancer Journal for Clinicians*, 52(3), 134–149. <https://doi.org/10.3322/canjclin.52.3.134>
- de Jonge, E. T., Vlasselaer, J., Van de Putte, G., & Schobben, J. C. (2009). The construct of breast cancer risk perception: Need for a better risk communication? *Facts, Views & Vision in ObGyn*, 1(2), 122–129.
- Fagerlin, A., Zikmund-Fisher, B. J., & Ubel, P. A. (2007). If I'm better than average, then I'm ok? Comparative information influences beliefs about risk and benefits. *Patient Education & Counseling*, 69(1–3), 140–144. <https://doi.org/10.1016/j.pec.2007.08.008>
- Fitzgerald, R. C., Antoniou, A. C., Fruk, L., & Rosenfeld, N. (2022). The future of early cancer detection. *Nature Medicine*, 28(4), 666–677. <https://doi.org/10.1038/s41591-022-01746-x>
- Fraenkel, L., Reyna, V., Cozmuta, R., Cornell, D., Nolte, J., & Wilhelms, E. (2018). Do visual aids influence patients' risk perceptions for rare and very rare risks? *Patient Education & Counseling*, 101(11), 1900–1905. <https://doi.org/10.1016/j.pec.2018.06.007>
- Franconeri, S. L., Padilla, L. M., Shah, P., Zacks, J. M., & Hullman, J. (2021). The science of visual data communication: What works. *Psychological Science in the Public Interest*, 22(3), 110–161. <https://doi.org/10.1177/15291006211051956>
- Fransen, M. P., Dekker, E., Timmermans, D. R. M., Uiters, E., & Essink-Bot, M.-L. (2017). Accessibility of standardized information of a national colorectal cancer screening program for low health literate screening invitees: A mixed method study. *Patient Education & Counseling*, 100(2), 327–336. <https://doi.org/10.1016/j.pec.2016.09.004>
- Gaissmaier, W., Tiede, K. E., & Garcia-Retamero, R. (2023). The lure of beauty: People select representations of statistical information largely based on attractiveness, not comprehensibility. *Medical Decision Making*, 43(7–8), 774–788. <https://doi.org/10.1177/0272989X231201579>
- Galesic, M., & Garcia-Retamero, R. (2011). Graph literacy: A cross-cultural comparison. *Medical Decision Making*, 31(3), 444–457. <https://doi.org/10.1177/0272989X10373805>
- Garcia-Retamero, R., & Cokely, E. T. (2017). Designing visual aids that promote risk literacy: A systematic review of health research and evidence-based design heuristics. *Human Factors*, 59(4), 582–627. <https://doi.org/10.1177/0018720817690634>
- Garcia-Retamero, R., Galesic, M., & Gigerenzer, G. (2010). Do icon arrays help reduce denominator neglect? *Medical Decision Making*, 30(6), 672–684. <https://doi.org/10.1177/0272989X10369000>
- Gudde Kuiper, P., Brabers, A., Reitsma van Rooijen, M., Zijlstra, C., & Jong, J. D. (2012). *Angst: een bijwerking van de bijsluiters? Het effect van twee verschillende presentaties van bijwerkingen op angst, veronderstelde vatbaarheid en ernst*. NIVEL. <https://www.nivel.nl/nl/publicatie/angst-een-bijwerking-van-de-bijsluiters-het-effect-van-twee-verschillende-presentaties>
- Hawley, S. T., Zikmund-Fisher, B., Ubel, P., Jancovic, A., Lucas, T., & Fagerlin, A. (2008). The impact of the format of graphical presentation on health-related knowledge and treatment choices. *Patient Education & Counseling*, 73(3), 448–455. <https://doi.org/10.1016/j.pec.2008.07.023>
- Henneman, L., Oosterwijk, J. C., van Asperen, C. J., Menko, F. H., Ockhuysen-Vermey, C. F., Kostense, P. J., Claassen, L., & Timmermans, D. R. M. (2013). The effectiveness of a graphical presentation in addition to a frequency format in the context of familial breast cancer risk communication: A multicenter controlled trial. *BMC Medical Informatics & Decision Making*, 13, 55. <https://doi.org/10.1186/1472-6947-13-55>
- Hersch, J., Jansen, J., & McCaffery, K. (2018). Decision-making about mammographic screening: Pursuing informed choice. *Climacteric*, 21(3), 209–213. <https://doi.org/10.1080/13697137.2017.1406912>
- Houts, P. S., Doak, C. C., Doak, L. G., & Loscalzo, M. J. (2006). The role of pictures in improving health communication: A review of research on attention, comprehension, recall, and adherence. *Patient Education & Counseling*, 61(2), 173–190. <https://doi.org/10.1016/j.pec.2005.05.004>
- Jones, S. C., Magee, C. A., Barrie, L. R., Iverson, D. C., Gregory, P., Hanks, E. L., Nelson, A. E., Nehill, C. L., & Zorbas, H. M. (2011). Australian women's perceptions of breast cancer risk factors and the risk of developing breast cancer. *Women's Health Issues*, 21(5), 353–360. <https://doi.org/10.1016/j.whi.2011.02.004>
- Kercher, K. (1992). Assessing subjective well-being in the old-old: The PANAS as a measure of orthogonal dimensions of positive and negative affect. *Research on Aging*, 14(2), 131–168. <https://doi.org/10.1177/0164027592142001>
- Koffka, K. (1922). Perception: An introduction to the gestalt-theorie. *Psychological Bulletin*, 19(10), 531–585. <https://doi.org/10.1037/h0072422>
- Lipkus, I. M., & Peters, E. (2009). Understanding the role of numeracy in health: Proposed theoretical framework and practical insights. *Health Education and Behavior*, 36(6), 1065–1081. <https://doi.org/10.1177/1090198109341533>
- Lipkus, I. M., Samsa, G., & Rimer, B. K. (2001). General performance on a numeracy scale among highly educated samples. *Medical Decision Making*, 21(1), 37–44. <https://doi.org/10.1177/0272989X0102100105>
- Mbanda, N., Dada, S., Bastable, K., Ingalill, G.-B., & Ralf, W. S. (2021). A scoping review of the use of visual aids in health education materials for persons with low-literacy levels. *Patient Education & Counseling*, 104(5), 998–1017. <https://doi.org/10.1016/j.pec.2020.11.034>
- McDowell, M., Gigerenzer, G., Wegwarth, O., & Rebitschek, F. G. (2019). Effect of tabular and icon fact box formats on comprehension of benefits and harms of prostate cancer screening: A randomized trial. *Medical Decision Making*, 39(1), 41–56. <https://doi.org/10.1177/0272989X18818166>
- Newman, E. J., & Schwarz, N. (2024). Misinformed by images: How images influence perceptions of truth and what can be done about it. *Current Opinion in Psychology*, 56, 101778. <https://doi.org/10.1016/j.copsyc.2023.101778>
- Ockhuysen-Vermey, C. F., Henneman, L., van Asperen, C. J., Oosterwijk, J. C., Menko, F. H., & Timmermans, D. R. M. (2008). Design of the BRISC study: A multicentre controlled clinical trial to optimize the communication of breast cancer risks in genetic

- counselling. *BMC Cancer*, 8, 283. <https://doi.org/10.1186/1471-2407-8-283>
- Okan, Y., Garcia-Retamero, R., Cokely, E. T., & Maldonado, A. (2015). Improving risk understanding across ability levels: Encouraging active processing with dynamic icon arrays. *The Journal of Experimental Psychology: Applied*, 21(2), 178–194. <https://doi.org/10.1037/xap0000045>
- Okan, Y., Janssen, E., Galesic, M., & Waters, E. A. (2019). Using the short graph literacy scale to predict precursors of health behavior change. *Medical Decision Making*, 39(3), 183–195. <https://doi.org/10.1177/0272989X19829728>
- Okan, Y., Stone, E. R., & Bruine de Bruin, W. (2018). Designing graphs that promote both risk understanding and behavior change. *Risk Analysis*, 38(5), 929–946. <https://doi.org/10.1111/risa.12895>
- Okan, Y., Stone, E. R., Parillo, J., Bruine de Bruin, W., & Parker, A. M. (2020). Probability size matters: The effect of foreground-only versus foreground+background graphs on risk aversion diminishes with larger probabilities. *Risk Analysis*, 40(4), 771–788. <https://doi.org/10.1111/risa.13431>
- Okan, Y., Stone, E. R., Petrova, D., & Bruine de Bruin, W. (2025). Communicating probabilities of cervical cancer screening results with icon arrays or tree diagrams: A longitudinal experiment. *Health Communication*, 40(9), 1757–1769. <https://doi.org/10.1080/10410236.2024.2421612>
- Oudhoff, J. P., & Timmermans, D. R. M. (2015). The effect of different graphical and numerical likelihood formats on perception of likelihood and choice. *Medical Decision Making*, 35(4), 487–500. <https://doi.org/10.1177/0272989X15576487>
- Pander Maat, H., Essink-Bot, M.-L., Leenaars, K. E. F., & Fransen, M. P. (2014). A short assessment of health literacy (SAHL) in the Netherlands. *BMC Public Health*, 14, 990. <https://doi.org/10.1186/1471-2458-14-990>
- Petrova, D., Garcia-Retamero, R., & Cokely, E. T. (2015). Understanding the harms and benefits of cancer screening: A model of factors that shape informed decision making. *Medical Decision Making*, 35(7), 847–858. <https://doi.org/10.1177/0272989x15587676>
- Rainey, L., van der Waal, D., & Broeders, M. J. M. (2020). Dutch women's intended participation in a risk-based breast cancer screening and prevention programme: A survey study identifying preferences, facilitators and barriers. *BMC Cancer*, 20, 965. <https://doi.org/10.1186/s12885-020-07464-2>
- Recchia, G. L., Bles, A. M. V., & Freeman, A. L. (2020). Predict: The potential pitfalls of visualisations of risk. *Breast Cancer Research and Treatment*, 180(2), 577–578. <https://doi.org/10.1007/s10549-019-05514-3>
- Reyna, V. F. (2008). A theory of medical decision making and health: Fuzzy trace theory. *Medical Decision Making*, 28(6), 850–865. <https://doi.org/10.1177/0272989x08327066>
- Reyna, V. F., & Brainerd, C. J. (2011). Dual processes in decision making and developmental neuroscience: A fuzzy-trace model. *Developmental Review*, 31(2), 180–206. <https://doi.org/10.1016/j.dr.2011.07.004>
- Reyna, V. F., & Brust-Renck, P. G. (2020). How representations of number and numeracy predict decision paradoxes: A fuzzy-trace theory approach. *Journal of Behavioral Decision Making*, 33(5), 606–628. <https://doi.org/10.1002/bdm.2179>
- Reyna, V. F., Edelson, S., Hayes, B., & Garavito, D. (2022). Supporting health and medical decision making: Findings and insights from fuzzy-trace theory. *Medical Decision Making*, 42(6), 741–754. <https://doi.org/10.1177/0272989x221105473>
- Reyna, V. F., Nelson, W. L., Han, P. K., & Dieckmann, N. F. (2009). How numeracy influences risk comprehension and medical decision making. *Psychological Bulletin*, 135(6), 943–973. <https://doi.org/10.1037/a0017327>
- Richter, R., Jansen, J., Bongaerts, I., Damman, O., Rademakers, J., & van der Weijden, T. (2023). Communication of benefits and harms in shared decision making with patients with limited health literacy: A systematic review of risk communication strategies. *Patient Education & Counseling*, 116, 107944. <https://doi.org/10.1016/j.pec.2023.107944>
- RIVM. (2023, March). *Bevolkingsonderzoek borstkanker - professionals: Actuele ontwikkelingen*. National Institute for Public Health and the Environment. <https://www.rivm.nl/bevolkingsonderzoek-borstkanker/professionals/actuele-ontwikkelingen>
- RIVM. (2025). *Invitation leaflet breast cancer population screening*. National Institute for Public Health and the Environment. <https://www.rivm.nl/documenten/folder-bevolkingsonderzoek-borstkanker-nl>
- Schubbe, D., Scalia, P., Yen, R. W., Saunders, C. H., Cohen, S., Elwyn, G., van den Muijsenbergh, M., & Durand, M. A. (2020). Using pictures to convey health information: A systematic review and meta-analysis of the effects on patient and consumer health behaviors and outcomes. *Patient Education & Counseling*, 103(10), 1935–1960. <https://doi.org/10.1016/j.pec.2020.04.010>
- Schwartz, L. M., Woloshin, S., Black, W. C., & Welch, H. G. (1997). The role of numeracy in understanding the benefit of screening mammography. *Annals of Internal Medicine*, 127(11), 966–972. <https://doi.org/10.7326/0003-4819-127-11-1997121010-00003>
- Smith, S. G., Raine, R., Obichere, A., Wolf, M. S., Wardle, J., & von Wagner, C. (2015). The effect of a supplementary ('gist-based') information leaflet on colorectal cancer knowledge and screening intention: A randomized controlled trial. *Journal of Behavioral Medicine*, 38(2), 261–272. <https://doi.org/10.1007/s10865-014-9596-z>
- Stellamanns, J., Ruetters, D., Dahal, K., Schillmoeller, Z., & Huebner, J. (2017). Visualizing risks in cancer communication: A systematic review of computer-supported visual aids. *Patient Education & Counseling*, 100(8), 1421–1431. <https://doi.org/10.1016/j.pec.2017.02.003>
- Stone, E. R., Bruine de Bruin, W., Wilkins, A. M., Boker, E. M., & MacDonald Gibson, J. (2017). Designing graphs to communicate risks: Understanding how the choice of graphical format influences decision making. *Risk Analysis*, 37(4), 612–628. <https://doi.org/10.1111/risa.12660>
- Stone, E. R., Reeder, E. C., Parillo, J., Long, C., & Walb, L. (2018). Salience versus proportional reasoning: Rethinking the mechanism behind graphical display effects. *Journal of Behavioral Decision Making*, 31(4), 473–486. <https://doi.org/10.1002/bdm.2051>
- Syrowatka, A., Krömker, D., Meguerditchian, A. N., & Tamblyn, R. (2016). Features of computer-based decision aids: Systematic review, thematic synthesis, and meta-analyses. *Journal of Medical Internet Research*, 18(1), e20. <https://doi.org/10.2196/jmir.4982>
- Tait, A. R., Voepel-Lewis, T., Zikmund-Fisher, B. J., & Fagerlin, A. (2010). The effect of format on parents' understanding of the risks and benefits of clinical research: A comparison between text, tables, and graphics. *Journal of Health Communication*, 15(5), 487–501. <https://doi.org/10.1080/10810730.2010.492560>
- Taylor, L. C., Hutchinson, A., Law, K., Shah, V., Usher-Smith, J. A., & Dennison, R. A. (2023). Acceptability of risk stratification within population-based cancer screening from the perspective of the general public: A mixed-methods systematic review. *Health Expectations*, 26(3), 989–1008. <https://doi.org/10.1111/hex.13739>
- Timmermans, D. R. M., Ockhuysen-Vermey, C. F., & Henneman, L. (2008). Presenting health risk information in different formats: The effect on participants' cognitive and emotional evaluation and decisions. *Patient Education & Counseling*, 73(3), 443–447. <https://doi.org/10.1016/j.pec.2008.07.013>
- Trevena, L. J., Bonner, C., Okan, Y., Peters, E., Gaissmaier, W., Han, P. K. J., Ozanne, E., Timmermans, D., & Zikmund-Fisher, B. J. (2021). Current challenges when using numbers in patient decision aids: Advanced concepts. *Medical Decision Making*, 41(7), 834–847. <https://doi.org/10.1177/0272989X21996342>
- van Strien-Knippenberg, I. S., Arjangi-Babetti, H., Timmermans, D. R. M., Schrauwen, L., Fransen, M. P., Melles, M., & Damman, O. C. (2024). Communicating the results of risk-based breast cancer screening through visualizations of risk: A participatory design approach. *BMC Medical Informatics & Decision Making*, 24(1), 78. <https://doi.org/10.1186/s12911-024-02483-6>
- van Strien-Knippenberg, I. S., Boshuizen, M. C. S., Determann, D., de Boer, J. H., & Damman, O. C. (2022). Cocreation with Dutch patients of decision-relevant information to support shared decision-making about adjuvant treatment in breast cancer care. *Health Expectations*, 25(4), 1664–1677. <https://doi.org/10.1111/hex.13510>

- van Strien-Knippenberg, I. S., Timmermans, D. R. M., Engelhardt, E. G., Konings, I. R. H. M., & Damman, O. C. (2024). Presenting decision-relevant numerical information to Dutch women aged 50–70 with varying levels of health literacy: Case example of adjuvant systemic therapy for breast cancer. *PLOS ONE*, 19(9). <https://doi.org/10.1371/journal.pone.0309668>
- Waters, E. A., Foust, J. L., Scherer, L. D., McQueen, A., & Taber, J. M. (2021). Adherence of internet-based cancer risk assessment tools to best practices in risk communication: Content analysis. *Journal of Medical Internet Research*, 23(1). <http://www.jmir.org/2021/1/e23318/>
- Windschitl, P. D., Martin, R., & Flugstad, A. R. (2002). Context and the interpretation of likelihood information: The role of intergroup comparisons on perceived vulnerability. *Journal of Personality and Social Psychology*, 82(5), 742–755. <https://doi.org/10.1037/0022-3514.82.5.742>
- Woudstra, A. J., Meppelink, C. S., Pander Maat, H., Oosterhaven, J., Fransen, M. P., & Dima, A. L. (2019). Validation of the short assessment of health literacy (SAHL-D) and short-form development: Rasch analysis. *BMC Medical Research Methodology*, 19(1), 122. <https://doi.org/10.1186/s12874-019-0762-4>
- Zikmund-Fisher, B. J., Dickson, M., & Witteman, H. O. (2011). Cool but counterproductive: Interactive, web-based risk communications can backfire. *Journal of Medical Internet Research*, 13(3), e60. <https://doi.org/10.2196/jmir.1665>
- Zipkin, D. A., Umscheid, C. A., Keating, N. L., Allen, E., Aung, K., Beyth, R., Kaatz, S., Mann, D. M., Sussman, J. B., Korenstein, D., Schardt, C., Nagi, A., Sloane, R., & Feldstein, D. A. (2014). Evidence-based risk communication: A systematic review. *Annals of Internal Medicine*, 161(4), 270–280. <https://doi.org/10.7326/M14-0295>