Diagrammatic Reasoning Meets Medical Risk Communication

Angela Brunstein (angela@brunstein.net)
Department of S.A.P.E. - Psychology, American University in Cairo, AUC Ave, PO Box 74, New Cairo 11835, Egypt

Joerg Brunstein (joerg@brunstein.net)
Department of S.A.P.E. - Psychology, American University in Cairo, AUC Ave, PO Box 74, New Cairo 11835, Egypt

Ali Marzuk (AM12334@rcsi-mub.com)
Royal College of Surgeons in Ireland – Bahrain, PO Box 15503, Adliya, Bahrain

Abstract

Informed consent for medical procedures requires that patients understand risks associated with diagnostic and treatment options. Similar to performance for diagrammatic reasoning and system dynamics, patients, physicians and medical students are reported to perform poorly on understanding medical risk-related information. At the same time, different presentation formats seem to support different kind of conclusions across domains. In this research, we investigated different formats of presenting risk information related to a treatment scenario with 22 medical and 50 non-medical students. As expected, medical students performed better than non-medical students for all versions of the problem, while non-medical students could partially compensate missing medical knowledge with displays that reduce complexity and allow reducing cognitive load. This implies that it is possible to support patients' decision making, but also highlights the need to educate patients on potential risks and benefits.

Keywords: medical decision making, diagrammatic reasoning, risk evaluation

Applying Diagrammatic Reasoning to Medical Risk Communication

Diagrammatic reasoning and understanding complex system has been demonstrated to be difficult for several task domains (e.g., Cronin, Gonzalez, & Sterman, 2009). Performance for these kinds of tasks is better if information can be directly read out from diagrams compared to inferring it (e.g., Larkin & Simon, 1987). The other way around, specific presentation formats seem to support different kinds of conclusions. For example, for judging a car's fuel efficiency, presentation of gallons per mile is more promising than presentation of miles per gallon (Larrick & Soll, 2008).

For the medical domain, understanding information related to diagnostic and treatment choices is essential to informed consent, evidence-based medicine and doctor-patient shared decision-making. As for understanding complex systems in general, there is evidence that both, physicians and patients have difficulties understanding risk related information (Mazur & Hickam, 1993; Windish, Huot, & Green, 2007). Similar to the diagrammatic reasoning literature, patients' and undergraduates' performance varies between different presentation formats (Gigerenzer et al., 2007; Shapira, Nattinger, & Mc Auliffe, 2006). Also medical students perform better for displays on accumulation problems than non-medical students for some, but not all medical scenarios (Brunstein, Gonzalez, & Kanter, 2010).

In this study, we aimed to combine the lessons learned from diagrammatic reasoning for understanding a treatment scenario on ventricular fibrillation with medical students and non-medical undergraduates.

For decision whether or not to undergo surgery to get an implantable cardioverter defibrillator (ICD) after surviving a heart attack, patients need to evaluate the risk of having another heart attack (i.e., severity of disease) and how likely an ICD can save their life during that heart attack (i.e., effectiveness of treatment).

In the medical literature, common measures for those values come from clinical trials with number of patients surviving versus dying in treatment, in this case ICD, versus control groups, in this case heart medication only.

For physicians, critical values impacting treatment decisions are absolute and relative risk reduction as estimated proportion of patients surviving due to treatment and number needed to treat as estimates of the number of patients who are exposed to potential side effects for saving one patient's life.

For illustrating these measures, several formats of information presentation are used in the medical literature and on patient information leaflets or websites. These include tables, frequency arrays or bar graphs with numbers or proportions of patients dying versus surviving in different conditions. As for the miles per gallon example, patients perform better for understanding outcomes of clinical trials when presented in terms of frequencies and not in terms of probabilities (Gigerenzer et al., 2007). Unfortunately, the decision whether or not undergo surgery to receive an ICD requires both, understanding the research results and estimating the probability of success for themselves. Also, when presented with several treatment options, comparing several pairs of 100 smileys each can become very confusing. Therefore, it is not evident which kind of display might help patients best to make informed decisions.

For each of these three domains, participants display a tendency for bias or errors. For the miles per gallon illusion (Larrik & Soll, 2008), participants tend to assume a linear instead curvilinear relationship and, therefore, they do not appreciate the increase of efficiency by replacing least efficient cars. For health information, participants tend to neglect base rates or to confuse variables for calculations and, therefore, do not understand the value of screening or underestimate the risks associated with treatment.

Different kinds of visualization do not change the concepts or required calculations, but they do change the likelihood of error in participants' responses. Miles per gallons come with decimals instead naturals and with constant base rates. This makes the comparison much easier. Frequencies instead conditional probabilities have the base rates already integrated and require one step less for processing. At the same time, naturals are more convenient for calculations and comparisons than proportions with varying base rates.

Applying these considerations to medical risk communication, all presentation formats are isomorph and display the same information. All of them allow comparison. At the same time, each of them invites for different strategies:

Tables explicitly display number of patients, invite to calculate, but require numeracy and statistical literacy to get the correct number and to understand that result. It is challenging to visualize patients in different conditions from numbers. And patient numbers need to be converted to estimate personal odds of surviving or dying.

Bar graphs illustrate proportions of patients and are therefore closest to personal risk. They invite to estimate, but they can be problematic given reported difficulties with probabilities (e.g., Gigerenzer et al., 2008). As for tables, it is difficult to imagine number of patients in different conditions.

Frequency arrays as favored by Gigerenzer and colleagues (2008) work with naturals and do not require considering base rates. They illustrate numbers of patients and are intuitive to understand. As for tables, number of patients need to be converted to estimate personal risk.

Similarly, separate or integrated displays are associated with different advantages and disadvantages. Separate displays can be directly mapped onto treatment conditions, but require comparing and integrating displays to derive conclusions on treatment effectiveness. Integrated displays highlight the difference and take processing away from participants, but make it more difficult to read the display. This holds especially for the part of patients that have survived in treatment condition, while the corresponding number of patients had died in control condition.

1. Based on the diagrammatic reasoning and system dynamics literature, we expected that medical students should outperform non-medical students due to their greater knowledge associated with medical risk communication and treatment options for that disease.
2. Given that background knowledge, we expected little impact of presentation formats on medical students' performance for evaluating outcomes of clinical trials.
3. In contrast, we expected, differences in performance for different presentation formats for non-medical students. Because tables, arrays and bar graphs are differently suited to support performance on proportion versus number of patients, but also vary in complexity and in how intuitive they are to interpret, we had no specific hypotheses which presentation should be best for this task.

Method

Participants

Fifty non-medical students (26 female, 24 male), mean age 21 years (SD =3) participated for course credit in this study and were randomly assigned to one of five presentation formats for a scenario on treating ventricular fibrillation with ICD (10 per condition).

Forty medical students (19 female, 16 male, 18 dropped before demographic information) mean age 21 years (SD = 2) agreed to participate in this study. Due to high dropout, only 22 participants (4-5 per condition) completed the scenario. All participants who quitted before completing the survey, did so *before* performing the risk information scenario at the consent page or at the demographics page. Data from these participants were excluded from analysis.

Design

This study implemented a 5 (versions of visualization) x 2 (medical knowledge) between participants design for answering 4 questions associated with severity of disease and effectiveness of treatment.

Materials

Participants were presented with scenario on 5-year survival of patients with ventricular fibrillation from a clinical trial (based on Moss et al., 1996). According the description, one hundred patients had an ICD implanted in addition to traditional heart medication (treatment condition). The remaining 100 patients received heart medication only (control condition). This was the topic of a group project for first year medical students at RCSI. Therefore, medical students were familiar with the topic, but not with the specific data.

Data were presented in one of five presentation formats: a table (see Figure 1a) or frequency arrays for number of patients dying versus surviving in treatment and control conditions (see Figure 1d and 1e) or bar graphs on proportions patients surviving versus dying (see Figure 1b and 1c). In addition for arrays and bar graphs, we either presented a pair of individual displays of patients surviving versus dying per condition or a combined display for both conditions.

Participants answered four questions on simplified versions of risk reduction and number needed to treat (NNT), on severity of disease and estimated effectiveness of treatment. These questions are relevant for patients, for example, with ventricular fibrillation, for making informed decisions on their treatment. The first question asked for a number, the remaining questions were true/false statements:

	Dead	Alive	Total
With ICD implant	20	80	100
With heart medication only	40	60	100
Total	60	140	200

Figure 1: Displays for patients surviving versus dying in treatment condition (ICD) and control condition (heart medication only) as (a) table, (b) integrated bar graph, (c) separate bar graphs, (d) integrated frequency array, and (e) separate frequency arrays.

Table 1: Average performance (and standard deviations) for medical students (med; N = 22) and non-medical students (non-med; N =40) for different presentation formats (maximum score was 4).

	Table	**1 Array**	**2 Arrays**	**1 Bar**	**2 Bars**	**Total**
med	3.0 (1.0)	3.2 (0.8)	3.7 (0.6)	3.0 (0)	3.3 (0.5)	3.2 (0.3)
non-med	1.6 (0.7)	2.5 (0.7)	2.0 (0.7)	2.1 (0.9)	1.7 (0.5)	2.0 (0.4)

The first question asked for the difference of survivors between both conditions as a proxy for treatment effectiveness and as a component of absolute and relative risk reduction. Correct answer is 20.

The second question asked whether more patients die than survive as a proxy for the severity of the disease or for the necessity to treat. Correct answer is "false".

The third question asked whether more patients die with implant than without. The correct answer is "false".

These 3 questions are needed to calculate the number needed to treat to save one patient's life (NNT). For this scenario, 1 of 5 patients survives due to treatment and the remaining 4 of 5 suffer from potential side effects of treatment without benefit. One of five will die in both conditions and 3 of 5 will live in both conditions. The correct answer is "true".

Because all four questions ask for aspects that are relevant for the treatment decision and because of the small number of participants among medical students, we report accumulated scores below. For patients, the next question would be whether they want to have an ICD implanted.

Procedure

The study was conducted as an online experiment on surveymonkey.com. After providing informed consent and demographic information, participants answered the four questions on the treatment scenario. Total time was about 10 to 15 min.

The IRB/ethics boards of the American University in Cairo and the Royal College of Surgeons in Ireland – Bahrain had approved this research.

Results

As expected, medical students performed better than non-medical students, $F(1, 62) = 47.12$, $p < .001$, $\eta2 = .43$.
Given the low number of participants for the group of medical students, we analyzed the effect of versions on participants' performance for the four questions separately for groups. For medical students, presentation format had no impact on performance, $F(4, 17) = 0.64$, $\eta2 = .13$. In tendency, medical students performed better for separate displays than for integrated displays (Bonferoni, all p's > .10). However, presentation format impacted performance of non-medical students, $F(4, 45) = 2.96$, $p < .05$, $\eta2 = .21$. Non-medical students performed better for the array than 2 bars or table (p's < .01) and in tendency better for integrated displays than for separate displays (p's > .10, see Table 1). This indicates that displays that reduce complexity and have potential for visual imagery better support non-medical students' performance for evaluating medical risk related information.

Discussion

Understanding medical risk communication is essential for informed consent for treatment choices, evidence-based medicine and physician-patient shared decision-making. At the same time, literature on medical decision-making indicates that patients' performance is far from perfect. This is matched by reported difficulties for diagrammatic reasoning and system dynamics. For our study, medical students performed better for evaluating treatment options than non-medical students.

However, not all medical students scored 4 of 4 questions correctly. As for diagrammatic reasoning literature, in our study different formats were differently supportive for evaluating medical risk information in non-medical students. Potentially, knowledge of the task domain and presentation formats that match the required task can help with understanding risk information. For our task that required understanding of research results in terms of frequencies and estimating the patient's own chances of success in terms of probability, there is not one single format of diagram that serves all aspects of the task best. Therefore, it seems that displays that are intuitive (arrays of patient numbers) and displays that allow reducing cognitive load (a combined bar graph on patient proportions) seem to foster non-medical students' performance. In contrast, medical students tended to profit from separate displays that can serve as external memory when calculating the statistical values. This means we should not leave it to the doctor to choose the presentation format for the patient because what is best for the expert does not match what is best for the layperson.

For our study, dropout rates for medical students (18 of 40 dropped) were very different from non-medical students' dropout rate. If eliminating the corresponding proportion of low-performing non-medical students, the effect of presentation format becomes weaker, but does not disappear completely. In addition, even the best performing group of non-medical students performs worse than any of the groups of medical students.

This indicates that presentation format can promote performance for medical risk evaluation, but cannot replace domain knowledge for understanding implications of illustrated data. Therefore, when supporting patients' treatment choices, we will need to educate them on potential risks and benefits in addition to provide intuitive displays.

Acknowledgements

This research was made possible by the Royal College in Ireland – Bahrain's generous support. The data collection on medical students at RCSI Bahrain was conducted when Dr. Angela Brunstein was a Senior Lecturer in Psychology at RCSI Bahrain. The contents of this report are solely the responsibility of the authors and do not necessarily represent the official views of RCSI Bahrain.

References

Brunstein, A, Gonzalez, C, & Kanter, S. (2010). Effects of domain experience in the stock-flow failure. *System Dynamics Review, 26,* 347–354.

Cronin, M., Gonzalez, C., & Sterman, J.D. (2009). Why don't well-educated adults understand accumulation? A challenge to researchers, educators, and citizens. *Organizational Behavior and Human Decision Processes, 108,* 116–130.

Gigerenzer, G., Gaissmaier, W., Kurz-Milcke, E., Schwartz, L.M., & Woloshin, S. (2007). Helping doctors and patients make sense of health statistics. *Psychological Science in the Public Interest, 8,* 53-96.

Larkin, J.H., & Simon, H.A. (1987).Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science, 11,* 65-99.

Larrik, R.P., & Soll, J.B. (2008). The MPG illusion. Science, 320, 1593-94.

Mazur, D.J, & Hickam, D.H. (1993). Patients' and physicians' interpretations of graphic data displays. *Medical Decision Making, 13,* 59–63.

Moss, A.J., Jackson Hall, W., Cannon, D.S., Daubert, J.P., Higgins, S.L., et al. (1996). Improved survival with an implanted defibrillator in patients with coronary disease at high risk for ventricular arrhythmia. *New England Journal of Medicine, 335,* 1933-1940.

Shapira, M.M., Nattinger, A.B., McAuliffe, T.L. (2006). The influence of graphical format on breast cancer risk communication. *Journal of Health Communication, 11*, 569-82.

Windish, D.M., Huot, S.J., & Green, M.L. (2007). Medicine residents' understanding of the biostatistics and results in the medical literature. *JAMA, 298,* 1010–1022.