

Risk literacy assessment of general practitioners and medical students using the Berlin Numeracy Test

Hendrik Friederichs (✉ friedeh@uni-muenster.de)

Westfälische Wilhelms-Universität Münster <https://orcid.org/0000-0001-9671-5235>

Roman Birkenstein

Institute of Medical Education and Student Affairs

Jan C. Becker

Institute of Medical Education and Student Affairs

Bernhard Marschall

Institute of Medical Education and Student Affairs

Anne Weissenstein

Marien-Hospital Erftstadt

Research article

Keywords: risk literacy, mass screening, mammography, medical education, predictive value

Posted Date: September 2nd, 2019

DOI: <https://doi.org/10.21203/rs.2.13869/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Version of Record: A version of this preprint was published on July 14th, 2020. See the published version at <https://doi.org/10.1186/s12875-020-01214-w>.

Abstract

Background Responsibility for helping patients understand potential health benefits and risks falls largely to general practitioners (GPs), especially for screening tests. The Berlin Numeracy Test (BNT) specifically measures risk literacy, meaning the range of statistical numeracy associated with accurate interpretation of information about risks. This study explored the association between the level of risk literacy and clinical experience, and risk literacy's impact on evaluating the predictive value of screening tests.

Methods The participants were 84 GPs and 92 third-year medical students who completed the BNT (total score range 0–4 points). For the GPs an additional case scenario on mammography screening was given as a simple measure of performance in applying numeracy skills. Results Despite having an average of 25.9 years clinical experience, GPs scored no better than did medical students on risk literacy (GPs: 2.33 points, 95% CI 2.08–2.59; students: 2.34, 95% CI 2.07–2.61; $P = .983$). Of all GPs, 71.6% ($n = 58$) came to the opposite conclusion of the real predictive value. Conclusions In this study, we found no difference in risk literacy between current students and current GPs. GPs lack risk literacy and in consequence do not fully understand numeric estimates of probability of routine screening procedures.

Background

As previously described in detail [1], several studies suggest that patients want to be included in discussions about their health, diagnostics, and therapy [2]. This patient-centered and shared decision-making approach between patients and doctors is the basis of modern medicine [3-7]. Individual risk calculations are an effective method for risk communication, particularly in patient education for screening programs. Previous research has shown that adequate risk communication can even increase the use of screening tests [8]. But studies have also demonstrated low *numeracy*, the ability to perform basic quantitative calculations, among patients [9]. Because understanding risk requires a certain level of numeracy, patients often must depend on their providers to interpret and communicate risk information. If providers also do not understand the quantitative aspects of risk, patients will not be able to comprehend or use the information well [10]. It is therefore safe to assume that, as a rule, only physicians with adequate training will be able to assist their patients in this way.

To fully understand and communicate numeric estimates of probability, a sufficient competence in risk numeracy is mandatory [11]. The ability to understand statistics and the quantitative aspects of research varies tremendously among physicians and medical students [12,13]. Additionally, because risk information is primarily quantitative, these individual differences arguably affect medical professionals' comprehension of diagnostic tests and other risk-related communications [14]. The Berlin Numeracy Test [BNT [15]] is a multiple-choice instrument that specifically measures *risk literacy*, meaning the range of statistical numeracy that is associated with accurate interpretation of information about risks. The BNT was developed for use with educated and highly educated samples. In a previous study we showed that the average risk literacy in medical students is slightly above average (Cohen's $d = 0.23$) and does not increase during medical studies [1]. Little is known about whether risk literacy increases during postgraduate medical education.

Previous research has suggested suboptimal risk literacy in clinicians. Gigerenzer and colleagues [16] studied 160 gynecologists and found that one-third were ignorant of the benefits of mammography screening. Most of the investigated gynecologists were unable to calculate the positive predictive value, so their probability of choosing the best answer was “slightly less than chance.” In a different report published as a letter, Estrada et al. [17] assessed the numeracy skills of 45 health professionals (including physicians, nurses, doctorate faculty, and medical students). Twenty percent of the participants gave incorrect answers to two or more of six absolutely basic numeracy questions. Bramwell and colleagues [18] reported that only about one-third of obstetricians were able to calculate the probability of an unborn child actually having Trisomie 21, given a positive test result. In interpreting these findings, one must consider that a lack of risk literacy on the one hand appears to account for medical students’ and novice physicians’ repeating unnecessary diagnostics [19] and on the other hand could lead to an overestimation of the value of diagnostic findings.

To our knowledge very limited data exist on the risk literacy of general practitioners (GPs), despite the importance of their role in health services in general and in risk communication in particular. GPs are the first line of contact regarding questions about whether to use offered screenings; therefore, they bear the main responsibility for helping patients understand the potential benefits and risks of screenings. For example, in Germany mammography screening is offered to all women from the age of 50 to 69 years. The invitation includes a recommended appointment in one of the certified mammography centers. In a patient-centered study about understanding of risks and benefits of breast cancer screening, most participants overestimated the benefits. Patients with higher numeracy assessed their risk for death with and without mammography more accurately than did patients with lower numeracy [20]. In Germany, the participation rate is about 56%, so nearly half of the invited women decline the recommended appointment [21]. Therefore, an additional issue that needs to be addressed in this context is the consequences of GPs’ risk literacy for their comprehension of and recommendations for the screening test.

Research Questions

We performed a controlled study on GPs and medical students to investigate the following questions:

1. What is the risk literacy among GPs?
2. Is there a difference in risk literacy levels between clinically experienced GPs and medical students?
3. Does the level of risk literacy influence evaluation of the predictive value of screening tests?

Methods

Setting and Subjects

Our study was conducted in the winter semester of 2015/2016 with general practitioners of the Westfalen-Lippe State Chamber of Physicians, which is a registered corporation under public law and represents the interests of 42,000 physicians in this area (the fourth-largest state chamber in Germany). A

postal questionnaire was our only financially viable option, but the response rates of GPs to postal surveys are often relatively low [22]. Expecting a large non-response rate we contacted as many GPs as necessary to reach the *a priori* calculated sample size. The students were recruited at the medical school of the Westphalian Wilhelms University in Münster by approaching them in the context of a curricular course in physical examination. This course prepares the students for their clinical clerkships, so it takes part at the very beginning of their clinical experiences.

Study Design

The participants were asked to complete the BNT voluntarily and anonymously. For the general practitioners, an additional case scenario on mammography screening was given to assess their performance in applying numeracy skills. Students were not asked to complete the scenario because it is already used twice in the medical curriculum to explain two aspects of mass screening (i.e., positive predictive power and overdiagnosis).

We did not ask medical students and general practitioners about their academic background in statistics and probability theory. However, completion of some statistical coursework is required before taking the preliminary medical exam ('Physikum'), which is a prerequisite for entering the clinical phase of medical school. Therefore, we can assume that all participants received statistical training.

All participants had to agree verbally to participate. Additionally, they provided informed consent prior to the study by reading the background information and deciding to fill in data. Ethical approval was given by the Ethics Committee of the Chamber of Physicians at Westfalen-Lippe and the Medical School of Westphalian Wilhelms University in Münster (2015-308-f-S).

Data Collection

In the BNT, the participants were required to solve four short case studies; two gave probabilities in relative number format, and two gave probabilities in absolute number format in the form of Bayesian inferences. To obtain the correct solution for these text problems, statistical information on the prevalence of a certain event had to be inserted into Bayes' theorem. Participants received one point per correct solution; thus, their total score could range between 0 and 4 points. The test format has been previously validated by Cokely and colleagues [15] against the original paper and pencil version of the BNT, which in turn has been benchmarked to several population datasets [15]. The average score (standard deviation) in the original validation study with this test format is 1.6 (1.21) [15]. Even when used with highly numerate individuals, the multiple-choice test format provided good discriminability and remained well balanced [15].

In the additional case scenario for the GPs, we presented information about mammography screening concerning disease prevalence and conditional probabilities (sensitivity and 1-specificity) (see supplementary material in Figure 1). The task was to give a probability estimate for the positive predictive value. Only the correct answer (9.2%) with an error margin from 8 to 10 was scored as correct; all other

responses and missing responses were considered incorrect. Each participant who wrote down a number in this range received one point.

Outcome Measures

Outcome measures were the total test scores students and GPs achieved on the BNT scenarios, and (for GPs), performance on the hypothetical case scenario. The GPs solved five case scenarios, and the students solved four, as mentioned earlier. In addition, baseline characteristics (age, sex, clinical experience for GPs) were obtained from all participants.

Statistical Methods

We used the standard alpha level of .05 for significance and a power level of 0.8. Therefore, we needed a sample size of at least 82 participants to detect an effect size showing a minimally important difference ($d = 0.4$) [23] in risk literacy levels between clinically experienced GPs and medical students (calculated a priori with G*Power 3.1) [24].

Acquired data were entered and analyzed using the statistical software IBM SPSS Statistics 21 (IBM Corp., Armonk, New York). Descriptive means and standard deviations were calculated for participants' age, and total test scores and frequencies were computed for gender and for solving the case scenarios. Sample means and frequencies were compared to population means and frequencies by one-sample t-tests and chi-squared-tests, respectively. We used a Mann–Whitney U test to assess the difference in risk literacy between medical students and GPs as well as the difference between GPs who could solve the case scenario successfully and those who could not. In addition, Kendall's tau correlation coefficient between BNT scores and successfully solved case scenarios was calculated. To compare the proportions of GPs and medical students achieving the maximum score, those data were analyzed with a chi-squared test. No adjustment for multiple testing was performed.

Results

Recruitment Process and Demographic Characteristics

The recruitment process is shown in Figure 1. We obtained 84 complete data sets (return rate 10.7%) after contacting a total of 785 contacted GPs. The average age of the GPs was 54.7 (SD 8.0) years, and 24 (28.6%) were female. They had an average clinical experience of 25.9 (SD 8.6) years. Ninety-two complete questionnaires from 96 participating medical students were entered into analysis (return rate 95.8%). The average age of the medical students was 23.2 (SD 4.0) years, and 55 (59.8%) were female. The time needed to complete the test was less than 10 minutes.

As Table 1 reveals, the baseline characteristics of the study populations and the analyzed samples show no significant differences.

Primary and secondary Outcomes

After the evaluation of all datasets, three major findings emerged. The first is that participants achieved an average score of 2.34 (SD 1.24) points, which is not meaningfully higher than the average score obtained by Cokely et al. [15], who reported that a score of 2.2 points was achieved by “highly numerate individuals.” There was only a minimal difference between the scores of general practitioners (2.33 points [95% CI 2.08–2.59]) and medical students (2.34 [95% CI 2.07–2.61]); accordingly, the Mann–Whitney U test showed no significance ($P = .983$). Figure 2 shows the distribution of scores by group. The proportion of participants achieving the maximum score (4 points) was higher for the student group (27.2% vs. 19.0%, $P = .20$).

Our second finding was that there were no significant differences between male and female participants ($P \geq .05$).

Finally, GPs who successfully solved the extra scenario ($n = 19$, 23.5%) also achieved significantly higher scores on the BNT (3.00 points) than did GPs who were unable to solve it ($n = 62$, 2.11 points, $P < .01$), which is also reflected in a positive correlation between BNT scores and solving the extra scenario (correlation coefficient Kendall's tau 0.31, $P < .01$). Of the GPs, 58 (71.6%) entered values $> 80\%$, a conclusion that strongly overestimates the real predictive value.

Discussion

Main Findings and Implications

Our results show no difference in risk literacy between current students and current GPs. Although showing above average test results, GPs lack understanding of the underlying success and failure probabilities of mammography screening and therefore will not be able to relate these to their patients accurately.

The first major finding of our study is that GPs showed above-average test results on the BNT, but no significant advantage over medical students could be detected. GPs reached an average risk literacy score of 2.33 points, which according to data of Cokely et al. [15] is not meaningfully better than the 2.2 points considered average for “highly numerate individuals.” The likelihood that a positive result on a diagnostic test is correct (the test's positive predictive value) strongly depends on the prevalence of the disease. If prevalence is low, even the best diagnostic test will not be very effective. GPs require accumulated experience in accurate interpretations of tests and the resulting risks under these conditions. But, despite their work in settings with low prevalence for many diseases, GPs show no better risk literacy than do average medical students.

Although the ability to work with numbers is important for physicians, only a handful of studies have explored their numeracy skills. In these studies, physicians have repeatedly been shown to have poor understanding of statistics and quantitative aspects of research [25-28]. They are particularly lacking in the ability to apply Bayesian inferences that are used to estimate probabilities of a hypothesis in the face of new data [29,30] [see also Wegwarth & Gigerenzer [31] for a good summary]. In the previously

mentioned study by Gigerenzer et al. [16], the majority of gynecologists grossly overestimated the probability of cancer after positive results in mammography screening, and their probability estimates were highly variable. Not surprisingly, studies among medical students have shown similar results. In 2002, Sheridan and Pignone provided 62 first-year medical students with three numeracy questions associated with information about the baseline risk for developing a hypothetical disease [32]. They reported that almost one-quarter of first-year medical students study had trouble performing basic numerical operations. Although showing a relatively high level of numeracy (77% could answer all questions correctly), only 61% of all students correctly interpreted quantitative data for treatment of a hypothetical disease. The authors concluded that medical students do not understand statistical concepts well, and our study suggests that this conclusion also applies to fully trained physicians.

As a consequence, GPs performed poorly in applying numeracy skills in the given case scenario. Nearly three-quarters of the GPs entered values of $> 80\%$, which denotes a conclusion quite disparate from the real predictive value. Risk literacy correlated significantly and positively with the ability to solve the case scenario. This is not only a problem with physicians. In the mammography study by Schwartz et al. mentioned earlier, even the most arithmetically able patients made accurate assessments of risk less than half the time [20]. In reality, physicians and patients are often confronted with additional irrelevant screening evidence such as improved 5-year survival and increased early detection. Wegwarth et al. demonstrated the unfortunate finding that “many physicians did not distinguish between irrelevant and relevant screening evidence” [33], a situation that almost certainly hampers understanding of benefits and risks of screening. Finally, it can be concluded that GPs lack risk literacy and in consequence do not fully understand the numeric estimates of probability associated with routine screening procedures.

In a Shared Decision Making approach, the treatment decision is made jointly by the doctor and the patient [34], with the patient’s wishes and needs about the therapy receiving considerable weight. This is the core element of patient-centered medicine [34]. Adequate communication of risks of disease and treatment modalities is an important factor for ensuring successful decision-making. Decision-making based on correct understanding of the underlying risks and benefits is central to the physician-patient relationship and is therefore an essential determinant of patient satisfaction and compliance [11,35]. We were able to show that physicians clearly do not have the necessary numeracy to interpret cancer screening statistics themselves. As Moyer (2012) states, “Expecting them to communicate this information to patients is a stretch.” [36].

Training in risk literacy may help to enhance students’ numeracy skills, but this sometimes does not seem to be the case [37]. GPs who successfully solved the case scenario attained significantly higher BNT scores (3.00 points) than participants who did not. This is also reflected in the positive correlation of solving the extra scenario with BNT scores. According to Cohen, a correlation coefficient of > 0.3 reflects a medium effect size [38]. So, the BNT can help to identify individual weaknesses in numeracy. The described BNT is a quick questionnaire for detecting individual deficiencies in risk literacy; therefore, it may enhance the effectiveness of trainings. Several tools exist to increase understanding of risks. [39,40]

However, the introduction of this new training in an already crowded medical studies curriculum remains a challenge.

The most promising way forward is the use of alternative statistical formats, in particular, natural frequencies (such as *100 of 1000* instead of *10%*) – for presenting the risks for screening tests. In consequence, developers of patient decision aids are recommended to preferentially use a natural frequency format for test and screening decisions [41]. The natural frequency method is very effective (effect size Cohen's d 0.69; 95% CI 0.45–0.93) and should be applied with patients *and* medical doctors [42]. Authors of publications about screenings should be encouraged by medical journal editors to ensure that results are presented in natural frequencies to avoid misinterpretation. For all other cases, students should be trained to generate natural frequencies from offered data by themselves.

Limitations

The main limitation of this study is the different return rates of the questionnaire by GPs and students. Whereas over 90% of the medical students were recruited successfully, far fewer of the GPs completed the questionnaire. Because the population of interest was a large, geographically dispersed one, the postal questionnaire was our only financially viable option. Although we tried to increase the response rate [43], only about 10% of the contacted GPs replied. Non-response reduces the effective sample size and may have introduced selection bias.

Selection bias is a severe limitation for interpretation of empirical findings. Baseline characteristics give no indication of systematic differences between the study populations and our analyzed samples. In our study, it is likely that if selection bias occurred, it would work *against* the reported findings; i.e., there is every chance that GPs who answered the postal survey were much more confident about their numeracy skills than were those who did not. Our results may therefore actually overestimate the competence of GPs in risk literacy and their understanding of routine screening procedures.

Conclusions

In this study, we found no difference in risk literacy between current students and current GPs. GPs lack risk literacy and as a consequence do not fully understand numeric estimates of probability of routine screening procedures.

Practice Points

- Responsibility for helping patients understand potential health benefits and risks falls largely to general practitioners (GPs).
- In this study, we found no difference in numeracy skills between current students and current GPs.
- There is pressing need to help physicians understand and correctly interpret risk information.
- Training on risk literacy during medical school may help, but we also suggest presenting risk information as natural frequencies.

Declarations

Author bios: Roman Birkenstein works as a doctoral student at the simulation center in Münster. Dr. Marschall is the dean of the medical faculty and the head of the Institute of Medical Education and Student Affairs at the University of Münster. Dr. Becker is Vice Head of the Institute of Medical Education and Student Affairs. Dr. Weissenstein works as a senior resident at the Department of Internal Medicine at the Marien-Hospital Erftstadt. Dr. Friederichs is the head of the simulation center at the medical school in Münster.

Notes on Contributors

HF conceived of the study and participated in its design and coordination. RB participated in the data acquisition and data analysis. JCB and BM participated in the study design. AW participated in the design and coordination of the study. All authors helped to draft the manuscript.

Acknowledgments

The authors wish to thank Saima Ghazal and Mirta Galesic for the provision of the Berlin Numeracy Test, and Gerd Gigerenzer for the suggestion of using a case scenario in the evaluation of GPs.

Funding/Support

None

Declaration of Interest

The authors have no conflicts of interest to declare.

Disclaimer

None

Word count

3051 words

Previous Presentations

None

Data Availability Statement

The original data underlying this manuscript can be found at <https://uni-muenster.sciebo.de/index.php/s/KqWs7nE2rGfnYgr>.

References

1. Friederichs H, Schölling M, Marschall B. Assessment of Risk Literacy Among German Medical Students: A Cross-Sectional Study Evaluating Numeracy Skills. *Human and Ecological Risk Assessment (HERA)*. 2014;20:1139–47.
2. Kiesler DJ, Auerbach SM. Optimal matches of patient preferences for information, decision-making and interpersonal behavior: evidence, models and interventions. *Patient Educ Couns*. 2006;61:319–41.
3. Kashaf MS, McGill E. Does Shared Decision Making in Cancer Treatment Improve Quality of Life? A Systematic Literature Review. *Medical decision making : an international journal of the Society for Medical Decision Making*. 2015;35:1037–48.
4. Serrano V, Rodriguez Gutierrez R, Hargraves I, Gionfriddo MR, Tamhane S, Montori VM. Shared decision-making in the care of individuals with diabetes. *Diabetic Medicine*. 7 ed. 2016;33:742–51.
5. Sheridan SL, Harris RP, Woolf SH, Shared Decision-Making Workgroup of the U.S. Preventive Services Task Force. Shared decision making about screening and chemoprevention. a suggested approach from the U.S. Preventive Services Task Force. *American journal of preventive medicine*. 2004;26:56–66.
6. Picano E. Informed consent and communication of risk from radiological and nuclear medicine examinations: how to escape from a communication inferno. *BMJ. British Medical Journal Publishing Group*; 2004;329:849–51.
7. Apter AJ, Paasche-Orlow MK, Remillard JT, Bennett IM, Ben-Joseph EP, Batista RM, et al. Numeracy and communication with patients: they are counting on us. *Journal of general internal medicine : official journal of the Society for Research and Education in Primary Care Internal Medicine*. Springer-Verlag; 2008;23:2117–24.
8. Edwards AGK, Naik G, Ahmed H, Elwyn GJ, Pickles T, Hood K, et al. Personalised risk communication for informed decision making about taking screening tests. Edwards AG, Edwards AG, editors. *Cochrane Database Syst Rev*. Chichester, UK: John Wiley & Sons, Ltd; 2013;2:CD001865.
9. Schwartz LM, Woloshin S, Welch HG. Can patients interpret health information? An assessment of the medical data interpretation test. *Medical decision making : an international journal of the Society for Medical Decision Making*. 2005;25:290–300.
10. Ancker JS, Kaufman D. Rethinking health numeracy: a multidisciplinary literature review. *Journal of the American Medical Informatics Association : JAMIA*. The Oxford University Press; 2007;14:713–21.
11. Ghosh AK, Ghosh K. Translating evidence-based information into effective risk communication: current challenges and opportunities. *J. Lab. Clin. Med*. 2005;145:171–80.
12. Bryant GD, Norman GR. Expressions of probability: words and numbers. *N Engl J Med*. 1980;302:411–2.

13. Ghosh AK, Ghosh K, Erwin PJ. Do medical students and physicians understand probability? QJM : monthly journal of the Association of Physicians. 2004;97:53–5.
14. Johnston BC, Alonso-Coello P, Friedrich JO, Mustafa RA, Tikkinen KAO, Neumann I, et al. Do clinicians understand the size of treatment effects? A randomized survey across 8 countries. CMAJ: Canadian Medical Association Journal. Canadian Medical Association; 2016;188:25–32.
15. Cokely ET, Galesic M, Schulz E, Ghazal S, Garcia-Retamero R. Measuring risk literacy: The Berlin numeracy test. Judgment and Decision Making. Society for Judgment and Decision Making; 2012;7:25–47.
16. Gigerenzer G, Gaissmaier W, Kurz-Milcke E, Schwartz LM, Woloshin S. Helping Doctors and Patients Make Sense of Health Statistics. Psychol Sci Public Interest. SAGE Publications; 2007;8:53–96.
17. Estrada C, Barnes V, Collins C, Byrd JC. Health literacy and numeracy. JAMA. 1999;282:527.
18. Bramwell R, West H, Salmon P. Health professionals’ and service users’ interpretation of screening test results: experimental study. BMJ. British Medical Journal Publishing Group; 2006;333:284–0.
19. Noguchi Y, Matsui K, Imura H, Kiyota M, Fukui T. Quantitative evaluation of the diagnostic thinking process in medical students. Journal of general internal medicine : official journal of the Society for Research and Education in Primary Care Internal Medicine. 2002;17:839–44.
20. Schwartz LM, Woloshin S, Black WC, Welch HG. The role of numeracy in understanding the benefit of screening mammography. Ann Intern Med. 1997;127:966–72.
21. Kooperationsgemeinschaft Mammographie Deutsches Mammographie Screening. Jahresbericht Evaluation 2012. Deutsches Mammographie Screening Programm. Berlin.
22. Cottrell E, Roddy E, Rathod T, Thomas E, Porcheret M, Foster NE. Maximising response from GPs to questionnaire surveys: do length or incentives make a difference? BMC medical research methodology. BioMed Central; 2015;15:1.
23. Hattie J. Visible Learning. Routledge; 2008.
24. Faul F, Erdfelder E, Lang A-G, Buchner A. G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behav Res Methods. 2007;39:175–91.
25. Wulff HR, Andersen B, Brandenhoff P, Guttler F. What do doctors know about statistics? Stat Med. 1987;6:3–10.
26. Windish DM, Huot SJ, Green ML. Medicine residents' understanding of the biostatistics and results in the medical literature. JAMA. American Medical Association; 2007;298:1010–22.
27. Berwick DM, Fineberg HV, Weinstein MC. When doctors meet numbers. Am J Med. 1981;71:991–8.
28. Gramling R, Irvin JE, Nash J, Sciamanna C, Culpepper L. Numeracy and medicine: key family physician attitudes about communicating probability with patients. The Journal of the American Board of Family Practice / American Board of Family Practice. 2004;17:473.
29. Hoffrage U, Lindsey S, Hertwig R, Gigerenzer G. Medicine. Communicating statistical information. Science. 2000;290:2261–2.

30. Jakob N, Sprenger M. Erhebung des Zahlenverständnisses sowie der Risk Literacy von Health Professionals in Österreich. 2014.
31. Wegwarth O, Gigerenzer G. Statistical illiteracy in doctors. Better doctors, better patients, better decisions : Envisioning health care 2020. MIT Press; 2011. pp. 137–51.
32. Sheridan SL, Pignone M. Numeracy and the medical student's ability to interpret data. *Effective clinical practice : ECP.* 2002;5:35–40.
33. Wegwarth O, Schwartz LM, Woloshin S, Gaissmaier W, Gigerenzer G. Do physicians understand cancer screening statistics? A national survey of primary care physicians in the United States. *Ann Intern Med.* 2012;156:340–9.
34. Weston WW. Informed and shared decision-making: the crux of patient-centered care. *CMAJ : Canadian Medical Association journal = journal de l'Association medicale canadienne.* 2001;165:438–9.
35. Epstein RM, Alper BS, Quill TE. Communicating evidence for participatory decision making. *JAMA.* 2004;291:2359–66.
36. Moyer VA. What we don't know can hurt our patients: physician innumeracy and overuse of screening tests. *Ann Intern Med.* 2012;156:392-393.
37. Noguchi Y, Matsui K, Imura H, Kiyota M, Fukui T. A traditionally administered short course failed to improve medical students' diagnostic performance. A quantitative evaluation of diagnostic thinking. *Journal of general internal medicine : official journal of the Society for Research and Education in Primary Care Internal Medicine.* Springer-Verlag; 2004;19:427–32.
38. Cohen J. *Statistical Power Analysis for the Behavioral Sciences.* Routledge; 2013.
39. Friederichs H, Ligges S, Weissenstein A. Using Tree Diagrams without Numerical Values in Addition to Relative Numbers Improves Students' Numeracy Skills: A Randomized Study in Medical Education. *Medical Decision Making.* 2014;34:253–7.
40. Garcia-Retamero R, Hoffrage U. Visual representation of statistical information improves diagnostic inferences in doctors and their patients. *Soc Sci Med.* 2013;83:27–33.
41. Trevena LJ, Zikmund-Fisher BJ, Edwards A, Gaissmaier W, Galesic M, Han PKJ, et al. Presenting quantitative information about decision outcomes: a risk communication primer for patient decision aid developers. *BMC medical informatics and decision making.* BioMed Central; 2013;13 Suppl 2:S7.
42. Akl EA, Oxman AD, Herrin J, Vist GE, Terrenato I, Sperati F, et al. Using alternative statistical formats for presenting risks and risk reductions. *Cochrane Database Syst Rev.* 2011;:CD006776.
43. Edwards PJ, Roberts I, Clarke MJ, DiGuseppi C, Wentz R, Kwan I, et al. Methods to increase response to postal and electronic questionnaires. Edwards PJ, editor. *Cochrane Database Syst Rev.* Chichester, UK: John Wiley & Sons, Ltd; 2009;:MR000008.

Table

Table 1: Baseline characteristics of study populations and analyzed samples

	Study Population	Analyzed Sample	
General Practitioners	Total of the state chamber	Participants	P-Value
n	4,848	84	
Age in years, mean (SD)	55.1 (8.7)	54.7 (8.0)	.65
Female sex, % (n)	33.3 (1,613)	28.6 (24)	.36
Clinical experience in years, mean (SD)		25.9 (8.6)	-
Medical Students	Total in the cohort	Participants	P-Value
n	102	92	
Age in years, mean (SD)	23.6 (4.5)	23.2 (4.0)	.34
Female sex, % (n)	60.8 (62)	59.8 (55)	.89

Additional File

Supplementary Figure 1: Case scenario of mammography screening

Figures

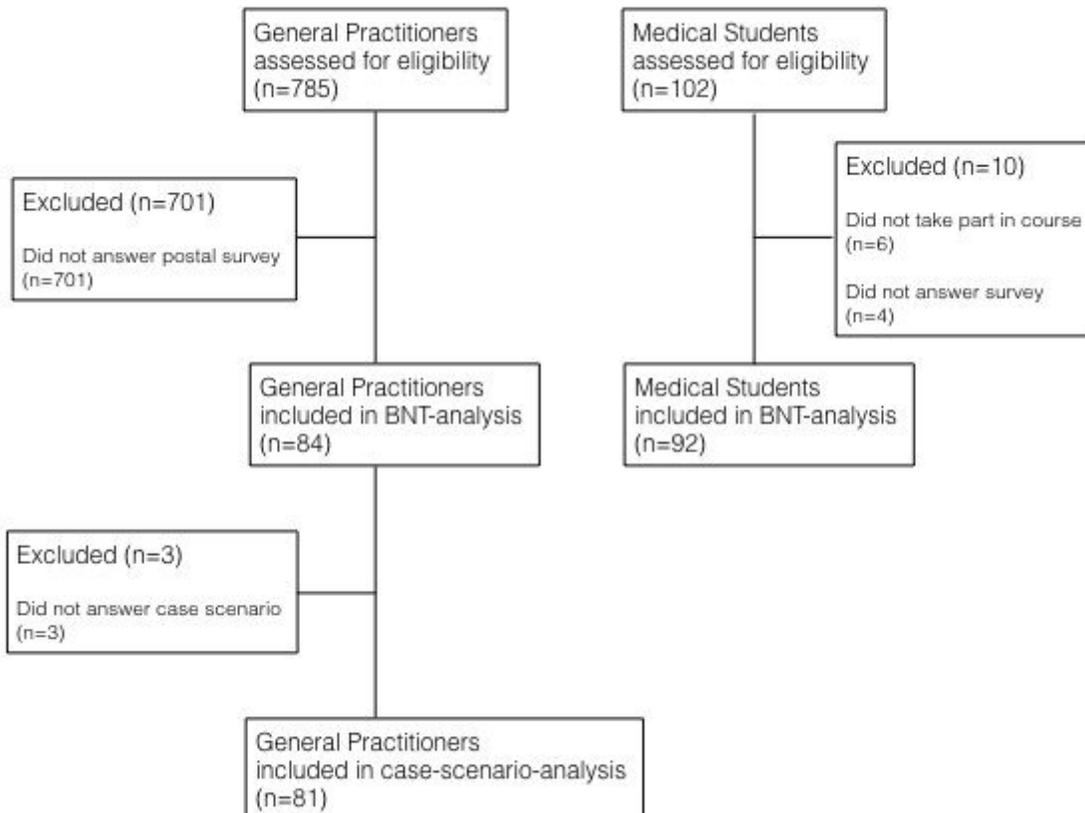


Figure 1

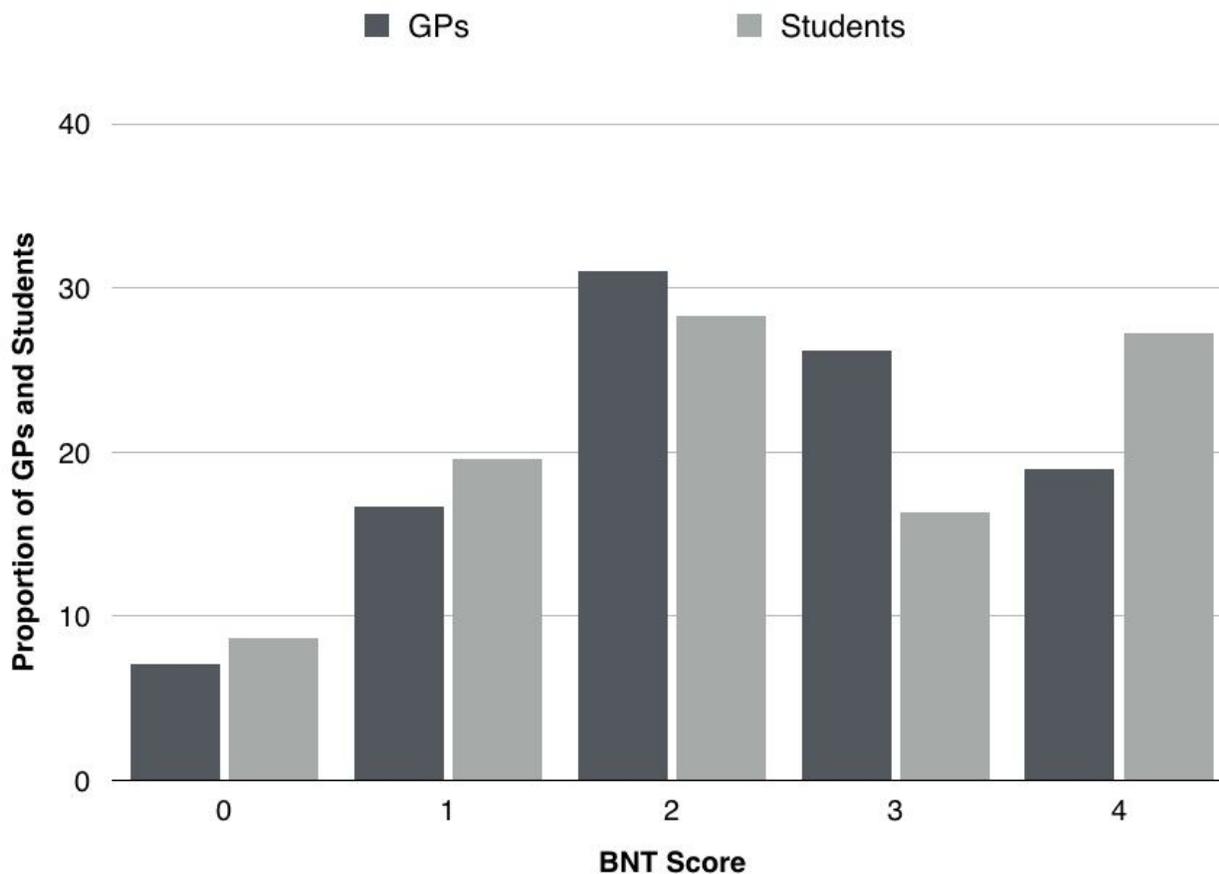


Figure 2

Distribution of BNT scores in GPs and medical students

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [supplement1.jpg](#)
- [supplement2.pdf](#)