

# BMJ Open Assessing statistical literacy in medical students and doctors: a single-centre, cross-sectional survey in South Korea

Sun Young Lee ,<sup>1,2</sup> Soela Kim,<sup>3</sup> Soyun Kim,<sup>3</sup> Yukyung Shin,<sup>4</sup> Jae-Joon Yim ,<sup>5,6</sup> Hyeontaek Hwang,<sup>5,6</sup> Youngsuk Kwon,<sup>7</sup> Un-Na Kim,<sup>8,9</sup> Young Kyung Do <sup>3,8</sup>

**To cite:** Lee SY, Kim S, Kim S, et al. Assessing statistical literacy in medical students and doctors: a single-centre, cross-sectional survey in South Korea. *BMJ Open* 2025;**15**:e095173. doi:10.1136/bmjopen-2024-095173

► Prepublication history and additional supplemental material for this paper are available online. To view these files, please visit the journal online (<https://doi.org/10.1136/bmjopen-2024-095173>).

Received 16 October 2024  
Accepted 20 March 2025



© Author(s) (or their employer(s)) 2025. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ Group.

For numbered affiliations see end of article.

**Correspondence to**  
Dr Young Kyung Do;  
ykdo89@snu.ac.kr

## ABSTRACT

**Objective** Healthcare professionals must possess statistical literacy to provide evidence-based care and engage patients in decision-making. However, there have been concerns about healthcare professionals' inadequate understanding of health statistics. As an initial step in addressing the issue, we assessed the statistical literacy of medical students and doctors in South Korea by evaluating their comprehension of four statistical concepts: (a) single-event probability, (b) relative risk reduction, (c) positive predictive value and (d) 5-year survival rate.

**Design** Cross-sectional survey study.

**Setting** The survey was conducted from October 2018 to January 2019 in one medical school and its affiliated teaching hospital in Seoul, South Korea.

**Participants** 303 medical students from all six grades and 291 doctors from various specialties.

**Primary and secondary outcome measures** The primary outcome measure was the correct answer rate for each question. The secondary outcome measure was the mean number of correct answers across the four statistical literacy questions, calculated for each individual.

**Results** The correct answer rates for basic numeracy questions were close to 100%. Regarding statistical literacy, 95.5% and 83.2% of the participants accurately understood single-event probability and relative risk reduction, respectively. However, only 49.3% and 49.2% of the participants accurately understood the positive predictive value and 5-year survival rate, respectively. The correct answer rates for the question about the 5-year survival rate differed significantly between students (40.9%) and doctors (57.7%) ( $p < 0.001$ ). There were no statistically significant differences in the correct answer rates for other questions, regardless of the student's grade level or the doctor's specialty.

**Conclusions** Medical students and doctors have weaker statistical literacy than their basic numeracy. Therefore, it is essential to implement medical education and professional development programmes that focus on improving their statistical literacy. These programmes should specifically address measures of medical test accuracy and the distinction between a 5-year survival rate and mortality.

## INTRODUCTION

Statistical literacy in healthcare entails the ability to critically assess statistics in health

## STRENGTHS AND LIMITATIONS OF THIS STUDY

- ⇒ This study assessed the statistical literacy among medical students at different stages of their education.
- ⇒ This study assessed the statistical literacy among practising doctors across various clinical experiences and specialties.
- ⇒ We measured statistical literacy using survey questions adapted to a more clinically relevant context.
- ⇒ Participants were recruited from only a single medical school and its affiliated teaching hospital in South Korea using a convenience sampling method.

information and understand statistical concepts in healthcare.<sup>1,2</sup> This competency is essential for healthcare professionals practising evidence-based medicine, where medical decisions are guided by the best available evidence—often numerically represented—alongside clinical expertise and patients' values and preferences.<sup>3,4</sup> For healthcare professionals, statistical literacy serves several critical functions. First, it enables the analysis and interpretation of emerging quantitative evidence about the benefits and risks of various healthcare options. Second, it allows for accurate statistical inferences from test results, aiding in accurate diagnosis and effective treatment planning.<sup>4,5</sup> Third, it facilitates clear explanations of the implications of tests and treatments to patients, supporting informed, shared decision-making.<sup>6</sup> Without statistical literacy, healthcare professionals may struggle to provide optimal care and effectively involve patients in their healthcare decisions.

Despite the critical importance of statistical literacy in healthcare, numerous studies have identified common misunderstandings and errors among health professionals about statistical concepts.<sup>5,7–9</sup> For example, healthcare professionals often struggle to comprehend and explain statistical concepts such as single-event probability (eg, there is



a 10% chance of an allergic reaction to a medication) and relative risk reduction (RRR, eg, a new drug reduces the risk of having a heart attack by 60% compared with a placebo).<sup>2 5 7 10</sup> Moreover, medical students and doctors frequently find it challenging to understand conditional probabilities like sensitivity and specificity, and how to combine them with disease prevalence to estimate the positive predictive value (PPV: the conditional probability of having a disease when a test result is positive).<sup>5 11–20</sup> Errors in estimating PPV can lead to severe consequences. Overestimating PPV can lead to overdiagnosis and overtreatment, causing unnecessary anxiety, costs and harm to patients.<sup>5 18–21</sup> Conversely, underestimating PPV can result in missed opportunities for early intervention and worsen patient outcomes. Additionally, healthcare professionals often conflate an increase in the 5-year survival rate with a reduction in mortality, even though these statistics measure different aspects of clinical and epidemiological data.<sup>15 22</sup> The 5-year survival rate represents the proportion of individuals who survive 5 years after being diagnosed with an illness, whereas mortality refers to the annual rate of disease-related deaths within a given population. Screening asymptomatic individuals can increase the 5-year survival rate by detecting the disease earlier, but it may not reduce mortality if the disease progression or treatment outcomes are unaffected. Therefore, relying solely on a 5-year survival rate, or confusing it with mortality, can misrepresent the life-saving benefits of screening programmes.<sup>10 15 22 23</sup> As evidenced in these widespread misunderstandings of fundamental statistical concepts, the lack of statistical health literacy among healthcare professionals not only impedes accurate assessment of medical interventions but also potentially compromises the overall effectiveness of evidence-based medicine.<sup>1 5 22 24</sup>

Although statistical literacy, like other medical knowledge and clinical skills, can affect patient health outcomes, the medical education community often overlooks the importance of continuously enhancing formal training to foster statistical literacy. To introduce improved formal training to medical students and doctors, it is imperative to assess their current statistical literacy on major clinical issues highlighted in the literature and identify areas for improvement. For a notable example, a study from Germany assessed the minimum statistical literacy of 169 final-year medical students, measuring their understanding of 10 basic statistics concepts, including sensitivity, specificity, PPV, RRR and mortality.<sup>2</sup> The students' median percentage of correct answers to these questions was 50% before brief training, which increased to 90% afterwards.

We aimed to assess the statistical literacy of medical students and doctors in South Korea, combining the conceptual components of the German study and adapting them to a more clinically relevant context.<sup>2</sup> When assessing statistical literacy, we focused on understanding four key statistical concepts: (a) single-event probability, (b) RRR, (c) PPV and (d) 5-year survival rate.

We also assessed their basic numeracy (ie, an elementary skill to understand and use numbers), which may be a prerequisite for statistical literacy.<sup>5</sup> This assessment will help identify which statistical concepts are most challenging for our target population and guide the development of improved medical education.

## METHODS

### Study setting

Between October 2018 and January 2019, we conducted a cross-sectional survey among a convenience sample of medical students from a single medical school and doctors from its affiliated teaching hospital on the same campus in Seoul, South Korea. South Korean medical education consists of a comprehensive 6-year programme. The medical school in this study admits approximately 135 students per academic year, with students completing their clinical clerkships at the medical school's affiliated teaching hospitals in their fifth and sixth years. The teaching hospital on the same campus has a capacity of 1800 beds and serves a large patient population, handling 2.4 million outpatient visits and 560 000 inpatients each year. Medical students from all 6 years of the programme, trainee doctors (interns and residents) and attending physicians at the hospital were eligible to participate in the study. We contacted student organisations and resident physicians to inform them about the study and sought their assistance in recruiting participants. Co-investigators, who were medical students or doctors-in-training themselves, approached potential participants before and after events, such as meetings, classes and conferences that many students and doctors attended. They explained the study and invited individuals to participate. Participants who agreed were given a questionnaire on the spot. The entire questionnaire included two separate thematic sections: medical statistical literacy (the focus of this article) and patient-centredness (not reported in this article), along with key demographic information, such as gender, age, student's year of study or doctor's grade and specialty. The questionnaire took approximately 5–10 min to complete. Participants received a gift card worth around US\$4 as a token of appreciation.

### Sample

We aimed to obtain a diverse sample of medical students across different years of study and doctors from various specialties within the teaching hospital. For medical students, our goal was to survey approximately one-third of students in each grade of the 6-year medical school programme, targeting 50 students per year for a total student sample of 300. Due to the expected challenges in surveying hospital doctors, we targeted approximately one-fourth of the trainee doctors, including both 1-year interns and 3-year or 4-year residents, as well as 100 attending physicians from a total of nearly 900. Since interns are not yet affiliated with a specialty, specialty information was collected only from residents

and attending physicians. We aimed to survey residents and attending physicians from 13 specialties out of the 23 clinical specialties, which we classified into three main groups: medical (internal medicine, paediatrics, rehabilitation, family medicine and psychiatry), surgical (emergency medicine, obstetrics and gynaecology, orthopaedics, otorhinolaryngology and general surgery) and service (radiology, laboratory medicine and anaesthesiology). The approximate proportion of these groups in our sample was set to 5:3:2 for residents and attending physicians, reflecting the proportion of residency openings in these specialties in 2019. See online supplemental table 1) for the target and actual numbers of participants.

### Measures

To assess basic numeracy, we used three fill-in-the-blank questions previously designed to measure basic numeracy across various populations in prior studies.<sup>5 25</sup> These questions involved converting between percentages and frequencies and interpreting chance outcomes. To assess statistical literacy, we developed four questions based on two previous studies regarding the statistical literacy of medical students and professionals that identified four commonly misunderstood statistical concepts: single-event probability, RRR, PPV and 5-year survival rate.<sup>2 5</sup> Our research team, consisting of experts in clinical medicine and medical education, initially formulated questions to evaluate comprehension of the four concepts based on their definitions. The question on single-event probability was included as it is often confusing due to the lack of a reference class, causing diverse misunderstandings. The RRR question aimed to assess participants' ability to explain this concept in the context of comparing new versus conventional chemotherapy. The PPV question involved calculating the probability of having breast cancer given a positive mammogram result, using information on sensitivity, specificity and prevalence. The question on the 5-year survival rate was framed around the increased survival rates of thyroid cancer in South Korea to evaluate participants' understanding of the distinction between increased survival and reduced mortality.<sup>5 23</sup> The wording of the formulated questions was further refined through an iterative process involving both medical students and physicians to improve readability and real-world clinical relevance. The final set of questions was evaluated and revised until consensus was reached. See [Box 1](#) for an English translation of the exact wording of the questions and response options.

### Statistical analysis

The collected data were converted into an anonymised database and analysed. The percentage of correct answers for each basic numeracy and statistical literacy question was computed. We analysed the percentage of correct answers for medical students and doctors for the four statistical literacy questions, examining variations by students' grades and doctors' specialty. Differences between subgroups were compared using the  $\chi^2$  test for

### Box 1 Questionnaire to assess basic numeracy and medical statistical literacy

#### Basic numeracy (BN)

BN Q1. People who take drug A have a 1% chance of having an allergic reaction. If 1000 people take drug A, how many people are expected to have an allergic reaction?

Answer: \_\_\_\_\_ out of 1000 (*Correct answer: 10*).

BN Q2. 1 out of 1000 people who take drug B may have an allergic reaction. What percentage of people who take drug B are expected to have an allergic reaction?

Answer: \_\_\_\_\_ % (*Correct answer: 0.1*).

BN Q3. Suppose that a coin is tossed 1000 times. How many times do you expect to get heads out of 1000 attempts?

Answer: About \_\_\_\_\_ times out of 1000 (*Correct answer: 500*).

#### Statistical literacy (SL)

SL Q1. Antidepressant C has a 20% risk of causing weight gain. Which of the following is the most correct explanation? (*Single-event probability*)

- Patients with depression who take C have a 20% increase in weight.
- 2 out of 10 patients with depression who take C experience weight gain. \*\*\*
- If you take 10 pills of C, 2 of them have a risk of causing weight gain.
- If you take C for 10 months, you are at risk of weight gain for 2 months.

SL Q2. A new chemotherapy drug reduces the risk of vomiting (as a side effect) by 60% compared with conventional chemotherapy. Which of the following is the most correct explanation? (*Relative risk reduction*)

- When using the new chemotherapy, the risk of vomiting is reduced to 40%.
- When using the new chemotherapy, vomiting occurs in 40 of 100 patients.
- Among 100 patients, the number of patients experiencing vomiting is reduced by 60 when using the new chemotherapy compared with conventional chemotherapy.
- If vomiting occurs in 50 out of 100 patients when using conventional chemotherapy, vomiting occurs in 20 out of 100 patients when using the new chemotherapy. \*\*\*

SL Q3. The prevalence of breast cancer for women in their 60s is 1%. A woman with breast cancer has a 90% chance of being positive on a mammogram, and a woman without breast cancer has a 9% chance of testing positive on a mammogram. Which of the following is the closest to the probability that a woman with a positive mammogram actually has breast cancer? (*Positive predictive value*)

- 81%.
- 9 out of 10.
- 1 out of 10. \*\*\*
- 1%.

SL Q4. The 5-year survival rate of thyroid cancer in South Korea has improved compared with the past. Which of the following is the most correct explanation? (*5-year survival rate*)

- It is possible that the incidence of thyroid cancer has decreased.
- An improvement in the 5-year survival rate of thyroid cancer means an improvement in the cure rate of thyroid cancer.
- An improvement in the 5-year survival rate of thyroid cancer means a reduction in mortality due to thyroid cancer.
- Early detection of thyroid cancer may increase the 5-year survival rate, but may not reduce mortality. \*\*\*

\*\*\* denotes the correct answer. The questionnaire was administered in Korean; this is an English translation. Italicised words were not included in the questionnaire but are shown here for clarity.

**Table 1** Characteristics of study participants

	Total (n=594)		Medical students (n=303)		Doctors (n=291)	
	n	%	n	%	n	%
<i>Age group (years)</i>						
-19	37	6.2	37	12.2	0	0.0
20-29	371	62.5	265	87.5	106	36.4
30-39	140	23.6	1	0.3	139	47.8
40-49	44	7.4	0	0.0	44	15.1
50+	2	0.3	0	0.0	2	0.7
<i>Gender</i>						
Male	355	59.8	179	59.1	176	60.5
Female	239	40.2	124	40.9	115	39.5
<i>Student's grade</i>						
Premedical			100	33.0		
Medical			203	67.0		
<i>Doctor's grade</i>						
Intern					40	13.7
Resident					146	50.2
Attending					105	36.1
<i>Doctor's specialty</i>						
Medical					125	43.0
Surgical					84	28.9
Service					41	14.1
Missing					1	0.3
See the text for the classification of medical specialties.						

two groups and ANOVA for three groups. All statistical analyses were conducted using SAS software, V.9.4 (SAS Institute Inc, Cary, North Carolina, USA). P values were based on a two-sided significance level of 0.05.

### Patient and public involvement

No patients were involved in setting the research question or the outcome measures, nor were they involved in developing plans for design or implementation of the study.

### RESULTS

A total of 303 medical students and 291 doctors participated in the survey. [Table 1](#) presents the characteristics of the participants.

The correct answer rate for all three basic numeracy questions was close to 100% in both the student and doctor groups ([figure 1](#)). The first two statistical literacy questions—the single-event probability question and the RRR question—also had high correct answer rates, approximately 95% and 83%, respectively. However, the PPV question and the 5-year

survival rate question had much lower correct answer rates, approximately 49%. There was no notable difference in correct answer rates between medical students and doctors, except for the 5-year survival rate question, where doctors had a higher correct answer rate (57.7%) than medical students (40.9%,  $p < 0.001$ ). See online supplemental table 2 for the distributions of answer choices selected for each statistical literacy question.

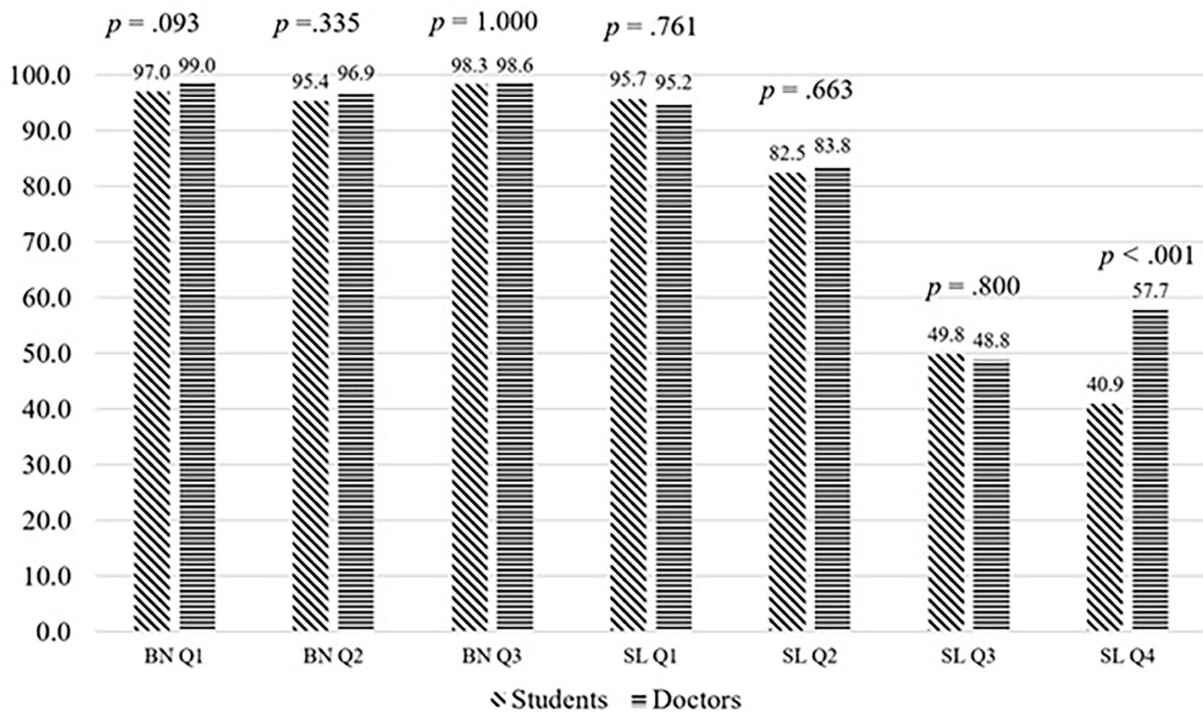
[Figure 2](#) presents the percentage of correct answers given by medical students at different grade levels for statistical literacy questions. In general, the correct answer rate increased with grade level, although the observed differences were not statistically significant. For instance, 46.6% of the third-year and fourth-year medical students answered the 5-year survival rate question correctly compared with 38.0% of the premedical and first-year and second-year medical students.

Doctors in service specialties consistently demonstrated higher correct answer rates than those in medical or surgical specialties, although these differences were not statistically significant ([figure 3](#)). Notably, the correct answer rate for the PPV question was higher among doctors in service specialties (56.1%) compared with their peers in medical (48.8%) and surgical (47.6%) specialties. The overall mean score on the four statistical literacy questions was also slightly higher among doctors in service specialties (online supplemental table 3).

### DISCUSSION

Despite their high basic numeracy, the medical students and doctors in this study demonstrated areas for improvement in key aspects of statistical literacy. While over 80% of participants correctly answered questions on single-event probability and RRR, correct answer rates were substantially lower for the PPV and 5-year survival rate questions. Both groups performed similarly on most questions, except for the 5-year survival question, where medical students had a significantly lower correct answer rate than doctors. Notably, increasing years in medical school did not result in considerably higher correct answer rates.

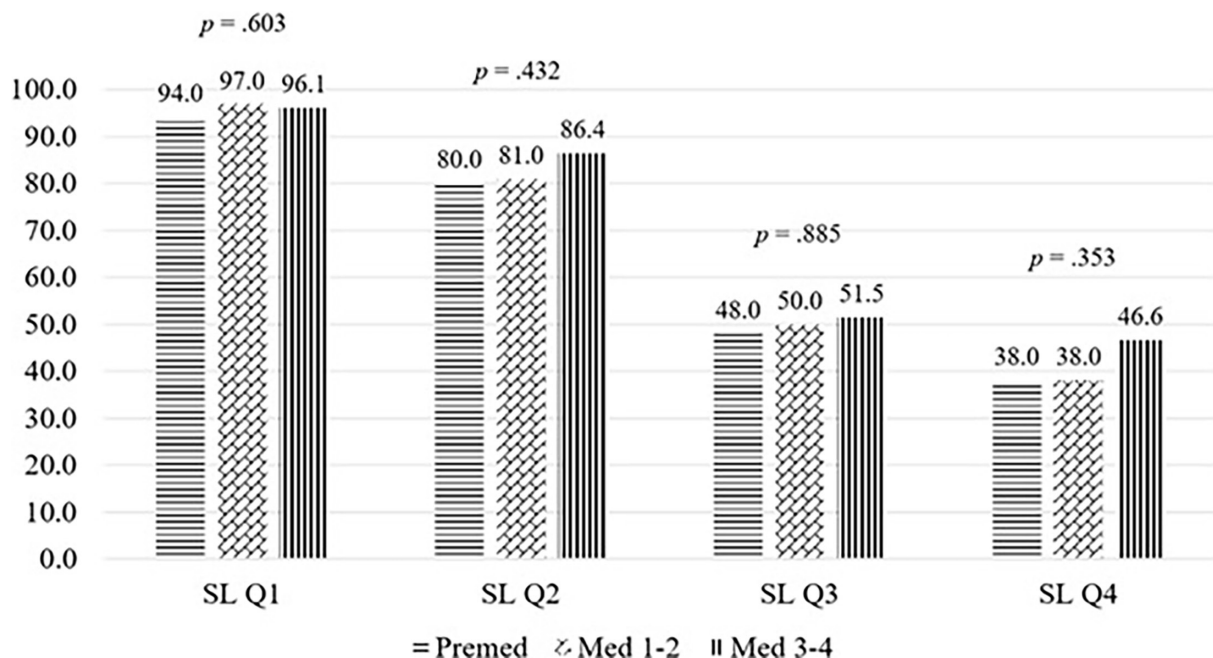
While this study may offer only a snapshot of the statistical literacy of Korean medical students and doctors, it serves as a window to examine the current state of medical education concerning statistical literacy in South Korea and beyond. Since the students and doctors in this study possessed high basic numeracy skills, their underdeveloped statistical literacy cannot be attributed to their basic numeracy. The two questions about statistical literacy, single-event probability and RRR, had high correct answer rates. This could be mainly because they are more closely related to basic numeracy compared with the other two questions, PPV and 5-year survival rate. These are areas where medical education can play an important role. Previous studies have linked insufficient statistical literacy in doctors to a non-transparent presentation of



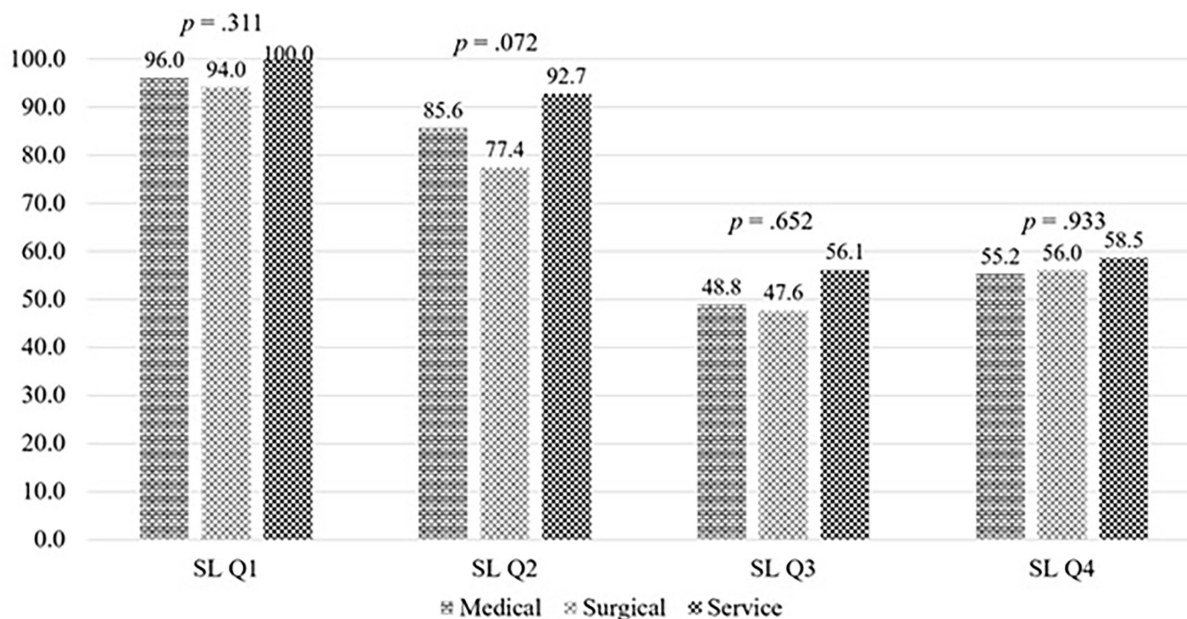
**Figure 1** Correct answer rates for basic numeracy (BN) and medical statistical literacy (SL) questions. BN Q1: converting a per cent into a proportion; BN Q2: converting a proportion into a per cent; BN Q3: familiarity with chance outcome; SL Q1: single-event probability; SL Q2: relative risk reduction; SL Q3: positive predictive value; SL Q4: 5-year survival rate.

statistical information and to medical schools that do not give adequate attention to the importance of teaching risk communication.<sup>2 5</sup> Therefore, it is crucial to introduce medical education and professional development programmes that enhance statistical literacy among medical students and doctors. In these programmes, the main focus should be on enhancing the ability to make

statistical inferences from medical test results and to acquire proficiency in using relevant medical statistics to critically evaluate the effects of illness and the life-saving advantages of medical treatments. The following discussion focuses on the two areas—PPV and 5-year survival rate, where there is a large room for improvement, as demonstrated in this study.



**Figure 2** Correct answer rates for medical statistical literacy (SL) questions by student grade. Q1: single-event probability; Q2: relative risk reduction; Q3: positive predictive value; Q4: 5-year survival rate.



**Figure 3** Correct answer rates for medical statistical literacy (SL) questions by specialty. Q1: single-event probability; Q2: relative risk reduction; Q3: positive predictive value; Q4: 5-year survival rate. See the text for the classification of medical specialties. Interns are not included in this analysis. The response from one attending physician who did not specify specialty was excluded from the analysis.

Although the medical students and doctors in this study had a significantly higher accuracy rate (approximately 49%) in answering the PPV question compared with their counterparts in previous German studies (approximately 20%),<sup>2,5</sup> this rate is still far from satisfactory, particularly from a medical education standpoint. Our results indicated that being in higher years in medical school and even currently practising medicine as a doctor was not associated with higher correct answer rates for the PPV question, suggesting both undergraduate and postgraduate medical education could improve significantly. Participants might have confused the PPV with sensitivity,<sup>13</sup> which was presented as 90% in the question. Alternatively, the tendency to overestimate PPVs, as observed in previous studies,<sup>12 16 26 27</sup> might have led to the incorrect answer. Regardless of which explanation is more plausible, it is important to remember that overestimating PPVs of medical tests can lead to further tests, unnecessary treatments and potential patient harm.<sup>5</sup> There is ample evidence that presenting statistical information in the form of natural frequencies rather than probabilities can improve conditional probabilistic reasoning as it helps with an intuitive understanding of conditional probabilities.<sup>8 17 28–31</sup> The observed effect of the natural frequency format was evident in individuals with both high and low numeracy.<sup>29 32</sup> Furthermore, studies have shown that teaching medical students and doctors how to translate relevant statistical information presented in probabilities into natural frequencies also facilitates conditional probabilistic reasoning.<sup>5 28 33 34</sup> However, given that statistical literacy skills improved after training can deteriorate within 1–2 months without reinforcement, medical schools and boards should implement

regular statistical training and assessments to maintain these crucial competencies.<sup>9 35</sup> It would be helpful to incorporate these research findings more actively when developing training programmes to improve medical students' and doctors' ability to estimate the predictive values of medical tests.

Another major weak area in statistical literacy identified in this study was the confusion between the 5-year survival rate and mortality. Almost half of medical students and one-third of doctors in this study incorrectly believed that an increase in the 5-year survival rate of thyroid cancer indicates a decrease in mortality from the disease. The relatively higher correct answer rate among doctors compared with medical students is likely due to their experience in clinical practice, where the concepts of 5-year survival rates and mortality are frequently used and compared. Nevertheless, the correct answer rate was still less than 60% among doctors and even lower among medical students, which is concerning because these two concepts must be carefully distinguished when assessing the impact of illness and the life-saving benefits of medical interventions.<sup>5 23</sup> Otherwise, healthcare professionals may overestimate the life-saving advantages of cancer screening, which could explain the overutilisation of low-value cancer screenings and the overdiagnosis of cancer. This misunderstanding has far-reaching consequences. Indeed, the overuse of low-value cancer screenings is contributing to cancer overdiagnosis and overtreatment globally, with South Korea being a particularly notable example of this trend.<sup>36–38</sup> When teaching about the 5-year survival rate and mortality, it is critical to emphasise their key difference: the denominators used in their calculation. For the 5-year survival rate, the

denominator is the number of people diagnosed with the disease, whereas for mortality rate, it is the number of people in the general population.<sup>23 39</sup> Understanding this distinction is essential for correctly interpreting these statistics in medical contexts. Using visual aids such as bar charts or pie charts can be helpful. These graphics could separately show (a) the proportion of people who survived for 5 years after being diagnosed with a certain disease (5-year survival rate) and (b) the annual rate of disease-related deaths in the population (mortality rate). Such visual representations can also facilitate comparing these rates across different populations or groups.<sup>21</sup>

This study has several limitations. First, we recruited participants from only a single medical school and its affiliated teaching hospital in South Korea using a convenient sampling method. This limits the generalisability of our findings beyond these specific institutions to the broader population of medical students and doctors across South Korea. Even within the single centre, participants who chose to respond to our survey might have different statistical literacy characteristics than those who did not participate, potentially leading to an overestimation of statistical literacy if those who felt more confident in their abilities were more likely to participate. Additionally, the exact response rate could not be calculated as the survey was conducted with a target number of participants based on the proportion of students and doctors, and data collection ended once the target was reached. Second, while our work focused on key weak areas of statistical literacy among medical students and doctors as highlighted in previous research,<sup>2 5</sup> it must be explicitly acknowledged that the questionnaire does not comprehensively cover all aspects of statistical literacy relevant to medical practice, as one would expect from a validated instrument. Third, we did not investigate the factors that may influence statistical literacy beyond examining its associations with basic characteristics available in the data.

Despite these limitations, this study provides the first assessment of the statistical literacy of Korean medical students in varying grades and doctors with varying clinical experience and specialties. Based on the findings of this study, we designed and carried out an educational intervention aimed at improving medical students' statistical literacy that is necessary to understand medical statistics and critically assess the available scientific evidence. We believe that this study will inspire further research in improving medical education regarding statistical literacy in South Korea and other countries.

#### Author affiliations

<sup>1</sup>Public Healthcare Center, Seoul National University Hospital, Jongno-gu, Seoul, Korea (the Republic of)

<sup>2</sup>Department of Human Systems Medicine, Seoul National University College of Medicine, Jongno-gu, Seoul, Korea (the Republic of)

<sup>3</sup>Institute of Health Policy and Management, Seoul National University Medical Research Center, Jongno-gu, Seoul, Korea (the Republic of)

<sup>4</sup>Centre for the Sociology of Organisations, Sciences Po, Paris, Île-de-France, France

<sup>5</sup>Department of Internal Medicine, Seoul National University College of Medicine, Seoul, Korea (the Republic of)

<sup>6</sup>Division of Pulmonary and Critical Medicine, Department of Internal Medicine, Seoul National University Hospital, Jongno-gu, Seoul, Korea (the Republic of)

<sup>7</sup>Seoul National University Hospital, Jongno-gu, Seoul, Korea (the Republic of)

<sup>8</sup>Department of Health Policy and Management, Seoul National University College of Medicine, Jongno-gu, Seoul, Korea (the Republic of)

<sup>9</sup>Korea Ministry of Health and Welfare, Sejong, Sejong, Korea (the Republic of)

**Acknowledgements** We thank the medical students and doctors who voluntarily participated in this study.

**Contributors** SYL, U-NK and YKD developed the questionnaire. SYL, YS, J-JY, HH, YK and YKD contributed to data collection. SYL conducted the statistical analysis and wrote the first draft of the manuscript. SYL, SoeK, SoyK and YKD critically reviewed and revised earlier drafts. All authors reviewed and provided input during the revision process. YKD secured funding for this study and is the guarantor of the study. An AI language model, ChatGPT, was used to assist with grammar checking and text refinement during the revision of this manuscript. The AI was employed solely for proofreading and enhancing language clarity.

**Funding** This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korean government (MSIT) (NRF-2018R1A2B2009257). The funder was not involved in the design of the study, the collection, analysis, and interpretation of data, or in writing the submitted work.

**Competing interests** None declared.

**Patient and public involvement** Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

**Patient consent for publication** Not applicable.

**Ethics approval** This study involves human participants and was approved by the Institutional Review Board of Seoul National University Hospital (IRB no. 1808-185-969). Participants gave informed consent to participate in the study before taking part.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Data availability statement** Data are available upon reasonable request. The data sets used in the current study are available from the corresponding author upon reasonable request.

**Supplemental material** This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

**Open access** This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

#### ORCID iDs

Sun Young Lee <http://orcid.org/0000-0002-1626-2721>

Jae-Joon Yim <http://orcid.org/0000-0002-9605-0074>

Young Kyung Do <http://orcid.org/0000-0001-5024-8264>

#### REFERENCES

- 1 Gigerenzer G, Gray JAM. Launching the century of the patient. In: *Better doctors, better patients, better decisions*. The MIT Press, 2011: 3–28.
- 2 Jenny MA, Keller N, Gigerenzer G. Assessing minimal medical statistical literacy using the Quick Risk Test: a prospective observational study in Germany. *BMJ Open* 2018;8:e020847.
- 3 Monahan J. *A prerequisite for evidence-based medicine*. SAGE Publications Sage CA: Los Angeles, CA, 2007.
- 4 Schmidt FM, Zottmann JM, Sailer M, et al. Statistical literacy and scientific reasoning & argumentation in physicians. *GMS J Med Educ* 2021;38:Doc77.



- 5 Gigerenzer G, Gaissmaier W, Kurz-Milcke E, *et al*. Helping Doctors and Patients Make Sense of Health Statistics. *Psychol Sci Public Interest* 2007;8:53–96.
- 6 Osborn CY. Shared Decision Making Requires Statistical Literacy. *AMA J Ethics* 2009;11:301–5.
- 7 Gigerenzer G. Better doctors, better patients, better decisions. 2014.
- 8 Lakhlifi C, Lejeune F-X, Rouault M, *et al*. Illusion of knowledge in statistics among clinicians: evaluating the alignment between objective accuracy and subjective confidence, an online survey. *Cogn Res Princ Implic* 2023;8:23.
- 9 Soto-Mota A, Maravilla EC, Fragoso JLC, *et al*. Evaluation of statistical illiteracy in Latin American clinicians and the piloting evaluation of a short course across multiple timepoints. *BMC Med Educ* 2022;22:54.
- 10 Gigerenzer G. Making sense of health statistics. *Bull World Health Organ* 2009;87:567.
- 11 Bramwell R, West H, Salmon P. Health professionals' and service users' interpretation of screening test results: experimental study. *BMJ* 2006;333:284.
- 12 Casscells W, Schoenberger A, Graboys TB. Interpretation by physicians of clinical laboratory results. *N Engl J Med* 1978;299:999–1001.
- 13 Eddy DM. Probabilistic reasoning in clinical medicine: problems and opportunities. In: *Judgment under uncertainty Heuristics and biases*. Cambridge: Cambridge University Press, 1982: 249–67.
- 14 Wegwarth O, Gigerenzer G. US gynecologists' estimates and beliefs regarding ovarian cancer screening's effectiveness 5 years after release of the PLCO evidence. *Sci Rep* 2018;8:17181.
- 15 Wegwarth O, Schwartz LM, Woloshin S, *et al*. Do physicians understand cancer screening statistics? A national survey of primary care physicians in the United States. *Ann Intern Med* 2012;156:340–9.
- 16 Whiting PF, Davenport C, Jameson C, *et al*. How well do health professionals interpret diagnostic information? A systematic review. *BMJ Open* 2015;5:e008155.
- 17 Hoffrage U, Gigerenzer G. Using natural frequencies to improve diagnostic inferences. *Acad Med* 1998;73:538–40.
- 18 Aronowitz R, Greene JA. Contingent Knowledge and Looping Effects - A 66-Year-Old Man with PSA-Detected Prostate Cancer and Regrets. *N Engl J Med* 2019;381:1093–6.
- 19 Gigerenzer G. HIV screening: helping clinicians make sense of test results to patients. *BMJ* 2013;347:bmj.f5151.
- 20 Gigerenzer G, Hoffrage U, Ebert A. AIDS counselling for low-risk clients. *AIDS Care* 1998;10:197–211.
- 21 Fagerlin A, Zikmund-Fisher BJ, Ubel PA. Helping Patients Decide: Ten Steps to Better Risk Communication. *JNCI Journal of the National Cancer Institute* 2011;103:1436–43.
- 22 Wegwarth O, Gaissmaier W, Gigerenzer G. Deceiving numbers: survival rates and their impact on doctors' risk communication. *Med Decis Making* 2011;31:386–94.
- 23 Welch HG, Schwartz LM, Woloshin S. Are increasing 5-year survival rates evidence of success against cancer? *JAMA* 2000;283:2975–8.
- 24 Petrova D, Mas G, Navarrete G, *et al*. Cancer screening risk literacy of physicians in training: An experimental study. *PLoS ONE* 2019;14:e0218821.
- 25 Schwartz LM, Woloshin S, Black WC, *et al*. The role of numeracy in understanding the benefit of screening mammography. *Ann Intern Med* 1997;127:966–72.
- 26 Manrai AK, Bhatia G, Strymish J, *et al*. Medicine's uncomfortable relationship with math: calculating positive predictive value. *JAMA Intern Med* 2014;174:991–3.
- 27 Lyman GH, Balducci L. Overestimation of test effects in clinical judgment. *J Cancer Educ* 1993;8:297–307.
- 28 Feufel MA, Keller N, Kendel F, *et al*. Boosting for insight and/or boosting for agency? How to maximize accurate test interpretation with natural frequencies. *BMC Med Educ* 2023;23:75.
- 29 Galesic M, Gigerenzer G, Straubinger N. Natural frequencies help older adults and people with low numeracy to evaluate medical screening tests. *Med Decis Making* 2009;29:368–71.
- 30 Gigerenzer G. What are natural frequencies? *BMJ* 2011;343:bmj.d6386.
- 31 McDowell M, Jacobs P. Meta-analysis of the effect of natural frequencies on Bayesian reasoning. *Psychol Bull* 2017;143:1273–312.
- 32 Chapman GB, Liu J. Numeracy, frequency, and Bayesian reasoning. *Judgm Decis Mak* 2009;4:34–40.
- 33 Sedlmeier P, Gigerenzer G. Teaching Bayesian reasoning in less than two hours. *J Exp Psychol Gen* 2001;130:380–400.
- 34 Kurzenhäuser S, Hoffrage U. Teaching Bayesian reasoning: an evaluation of a classroom tutorial for medical students. *Med Teach* 2002;24:516–21.
- 35 Kim S, Kim S, Choi Y-J, *et al*. Natural frequency tree- versus conditional probability formula-based training for medical students' estimation of screening test predictive values: a randomized controlled trial. *BMC Med Educ* 2024;24:1207.
- 36 Li M, Maso LD, Vaccarella S. Global trends in thyroid cancer incidence and the impact of overdiagnosis. *Lancet Diabetes Endocrinol* 2020;8:468–70.
- 37 Ahn HS, Kim HJ, Welch HG. Korea's thyroid-cancer "epidemic"-- screening and overdiagnosis. *N Engl J Med* 2014;371:1765–7.
- 38 Lee J-H, Shin SW. Overdiagnosis and screening for thyroid cancer in Korea. *The Lancet* 2014;384:1848.
- 39 Gigerenzer G, Wegwarth O. Five year survival rates can mislead. *BMJ* 2013;346:bmj.f548.