

# Interpreting the result of a diagnostic test: An experimental study in a social network

Ahmad Sofi-Mahmudi, DDS

Cochrane Iran Associate Centre, National Institute for Medical Research Development (NIMAD), Tehran, Iran.

Postal address: No. 21, West Fatemi St., Tehran, Iran. Postal code: 1419693111

Email address: [a.sofimahmudi@gmail.com](mailto:a.sofimahmudi@gmail.com)

Telephone number: +98 21 6693 8037

Fax number: +98 21 6690 0920

## Abstract

**Background:** Statistical literacy is an important element in informing patients and planning treatments for both clinicians and policymakers. This study aim to investigate and compare the interpretation of a diagnostic test result in medical and non-medical professionals.

**Methods:** A diagnostic test question about positive predictive value was posted on the “Stories” section of five personal Instagram accounts. The collected information included the exact calculated number, educational status (undergraduate or graduated) and field of study (dentistry, medicine, other medical, non-medical). The results were analysed using R version 3.4.4 on Microsoft Windows 10 with the Pearson  $\chi^2$  test.

**Results:** Of 121 participants, 63.6% did not provide a correct answer. The highest and the lowest correct answer ratios were seen in non-medical and dentistry groups, with 52.6% and 22.2%, respectively ( $P = 0.09$ ,  $\chi^2 = 6.464$ ,  $df = 3$ ). Undergraduates showed a more favourable performance than graduates with 41.8% correct answers compared to 20.0% ( $P = 0.03$ ,  $\chi^2 = 4.616$ ,  $df = 1$ ).

**Conclusion:** The statistical interpretation of the medical professionals is lower than the non-medical professionals and graduates showed less favourable results compared to undergraduates. There is a need for more beneficial and continuative statistical education for medical professionals.

**Keywords:** Data interpretation, dentists, graduate, physicians, statistics, undergraduate

## Introduction

Statistical literacy is a concern for patients, practitioners, and policymakers. Inappropriate interpretation of health outcome statistics can cause misunderstandings that may lead to preventable unfavourable health outcomes (1). For example, by under- or over- estimating a health statistic, unnecessary stress can suffer patients or policymakers would invest resources on ineffective health programs.

Quantitative studies are the predominant type of research in current medical literature so that reading most medical sciences research papers requires interpreting statistical results. Although students in

medical universities are expected to learn how to interpret these results during several courses, it seems that there is still a lack of knowledge in graduates (2). The role of practitioners is vital because they have access to health statistics relating to their field and act as an interpreter and carrier of this information to their patients.

Health statistics are presented in various ways including risk (e.g. frequencies, percentages, and probabilities) or risk reduction (e.g. relative risk reduction, absolute risk reduction, and number needed to treat). These ways have their advantages and disadvantages; for example, it seems that people understand natural frequencies better than percentages (3). In this study, interpreting a diagnostic test result presenting by percentage among members of a social network was evaluated.

## Methods

### Participants

We recruited participants from the followers of personal Instagram accounts on 30 September 2018. Participants were from a diverse range of disciplines but mainly dental and medical students studying in Iran.

### Question

A question used in a previous research article regarding a positive screening result presented in percentage format (4) was used. I translated and simplified it and put it on the “Stories” section of five Instagram accounts using a “Question Sticker”. Stories are temporary posts which disappear after 24 hours and participants cannot see each other’s responses. The question and its answer are described in Box 1 and Box 2, respectively.

#### **Box 1. The question asked in “Stories”.**

*The first story:*

One simple question:

About 1% of babies born in a population have Down’s syndrome. In a screening test, if the baby actually has Down’s syndrome, there is a 90% chance that the result will be positive. If the baby does not have Down’s syndrome, there is a 1% chance that the test will be positive.

*The second story:*

A pregnant woman has given this test and the result has been positive. How likely is that her baby will have Down’s syndrome?

Can you solve this question?

#### **Box 2. The answer to the question.**

If 10,000 pregnant women were tested, we would expect 100 (1% of 10 000) to have babies with Down's syndrome.

Of these 100 babies with Down's syndrome, the test result would be positive for 90 (90% of 100) and negative for 10.

Of the 9900 unaffected babies, 99 (1% of 9900) will also have a positive test, and 9801 will have a negative test result.

So, out of the 10,000 pregnant women tested, we would expect to see 189 (90 + 99) positive test results. Only 90 of these actually have babies with Down's syndrome, which is 47.6%.

Therefore, 47.6% of pregnant women who have a positive result from the test would actually have a baby with Down's syndrome.

The correct answer to this question is 47.6%. Positive predictive value (PPV) must be calculated to solve it. The original article proposes the solution method described in Box 2. It is worth noting that there is another method using conditional probability ( $P(\text{Down's syndrome} | +)$ ) (5).

Answers ranging from 45 to 50 were considered as correct.

## Data collection

Data were collected from five personal Instagram accounts with 3532 followers in total. The collected information included the exact calculated number, educational status (undergraduate or graduated) and field of study. Field of the study was divided into four categories: dentistry, medicine, other medical, and non-medical. The non-medical category was used as control. To avoid duplication of the results, account names were also checked.

## Data analysis

The results were analysed using R version 3.4.4 on Microsoft Windows 10. The Pearson  $\chi^2$  test was used to compare groups by field of study and educational status.

## Results

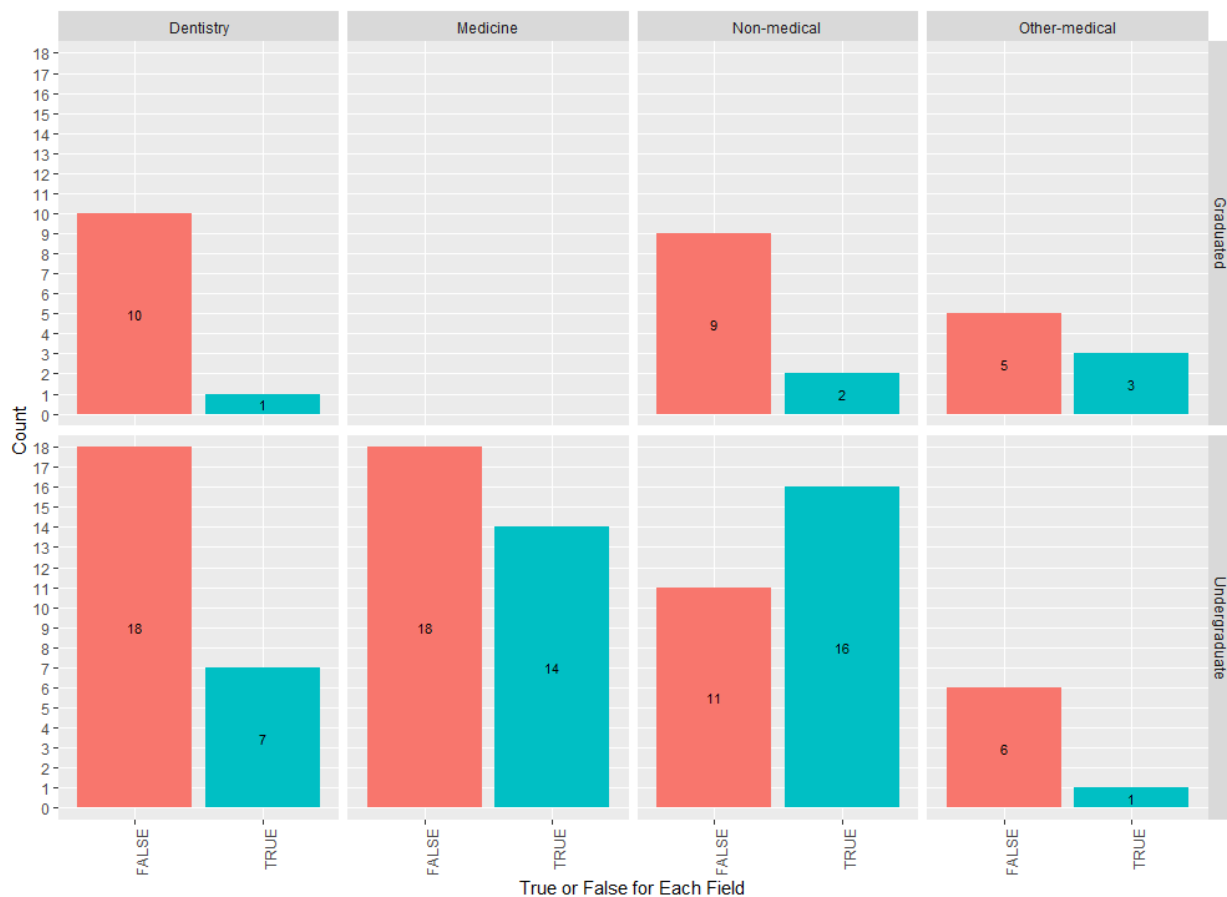
Overall, 121 people answered the question (52.9% female). The response rate was not calculated as many people may simply skip the stories and without reading the text. The details of the participants are summarised in Table 1.

**Table 1. Number of participants in each field and their graduation status.**

Field	Undergraduate	Graduated
Dentistry	25	11
Medicine	32	0
Pharmacy	3	0
Dental prosthesis technician	1	2
Public health	1	2
Occupational therapy	1	0

<b>Nursing</b>	1	0
<b>Health information technology</b>	0	1
<b>Midwifery</b>	1	1
<b>Laboratory sciences</b>	0	1
<b>Electrical engineering</b>	10	2
<b>Computer engineering</b>	7	3
<b>Architecture</b>	1	2
<b>Biotechnology</b>	1	0
<b>Chemistry</b>	1	1
<b>Civil engineering</b>	1	1
<b>Economics</b>	0	1
<b>Industrial engineering</b>	2	0
<b>Law</b>	1	0
<b>Mathematics</b>	1	0
<b>Physics</b>	0	1
<b>Physical education</b>	1	0
<b>Psychology</b>	1	0

Figure 1 shows the correct and incorrect response by field and educational status. The majority of responses (52.1%) were between 0-10% and 90-100%. Overall, 63.6% of the participants did not provide a correct answer. Dental graduates provided the least correct answers ratio among all groups (90.9%). The highest and the lowest correct answer ratios were seen in non-medical and dentistry groups, with 52.6% and 22.2%, respectively ( $P = 0.09$ ,  $\chi^2 = 6.464$ ,  $df = 3$ ). Only non-medical undergraduates had a correct answer ratio over 50%; however, the difference between undergraduates was not statistically significant ( $P = 0.05$ ,  $\chi^2 = 7.570$ ,  $df = 3$ ). Undergraduates showed a more favourable performance than graduates with 41.8% correct answers compared to 20.0% ( $P = 0.03$ ,  $\chi^2 = 4.616$ ,  $df = 1$ ).



**Figure 1. Results of responses to the question, by field of study**

## Discussion

This study indicated that the status of interpretation of statistics among health care professionals is not favourable, especially among participants with dental background. Participants outside the medical field provided the best results. Furthermore, undergraduates provided better results than graduates did. This may demonstrate that graduation has a negative effect on the statistical interpretation ability as graduates deal less with health statistics in their daily practices.

In the study which the question of this study was derived from, midwives and gynaecologists had the lowest percentage of correct answers among other groups (4). However, the results of participants with medical sciences backgrounds and overall results were lower than our study. Still, it seems necessary to improve this situation among students of medical universities in Iran.

Most responses were either very high or very low. Two incorrect answers of 1.0% and 99.0% were the most frequent answers. Respondents who answered 1.0% just used the base rate of Down's syndrome in the population and those who reached 99.0% considered the conditional probability of observing a positive test in a child without Down's syndrome. This may show that these participants have not completely understood the concept of PPV and how to calculate it.

## Limitations

Online tests can simplify the study design and save time and other resources. However, controlling the study conditions such as time and use of other sources to answer the question may be a concern. Yet the results of this study can be used to raise awareness about the problem and to design studies of higher quality and at a larger scale.

## Conclusion

This study showed that the statistical interpretation of the medical professionals is lower than the non-medical professionals and graduates showed less favourable results compared to undergraduates. Based on these results, more attention should be paid to the interpretation of medical statistics in medical universities of Iran and the educations should be continued after graduation.

## Acknowledgments

The author would like to thank Pouria Iranparvar, Hemn Rasouli, Soroush Sadr, and Melika Mohammadi for their help in data collection.

## References

1. Gigerenzer G, Gaissmaier W, Kurz-Milcke E, Schwartz LM, Woloshin S. Helping doctors and patients make sense of health statistics. *Psychol. Sci. Public Interest.* 2007;8(2):53-96.
2. Steurer J, Fischer JE, Bachmann LM, Koller M, ter Riet G. Communicating accuracy of tests to general practitioners: a controlled study. *BMJ.* 2002;324(7341):824-6.
3. 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(3).
4. Bramwell R, West H, Salmon P. Health professionals' and service users' interpretation of screening test results: experimental study. *BMJ.* 2006;333(7562):284.
5. Molinaro AM. Diagnostic tests: how to estimate the positive predictive value. *Neurooncol Pract.* 2015;2(4):162-6.