

# Statistical Illiteracy Among Clinicians: A review of the evidence and the ethical problems it generates

Dr Adrian Soto-Mota

---

*Abstract. To become a medical doctor, it is necessary to acquire knowledge in many different subjects such as Biochemistry, Public Health and Clinical Examination. Statistical reasoning is a crucial skill for clinicians not only because it is necessary to critically analyse emerging evidence from biomedical sciences but also because it is essential for risk assessments at bedside and for advising patients. Despite this, evidence suggests that even experienced clinicians struggle with assimilating the differences and implications of fundamental statistical concepts such as odds ratio versus absolute risk and sensitivity versus positive post-test probability. On the other hand, useful concepts such as number needed to treat/screen, intention to treat analysis, and Bayesian probability are often overlooked or ignored when making clinical decisions. Even further, some studies report that there are discrepancies between the interventions or treatments which clinicians prescribe and the ones that they undergo when they are patients. This review intends to illustrate how statistical illiteracy generates ethical problems in the patient clinician relationship, summarise existing evidence about this problem and the concerns raised by experts on this matter, and analyse the potential solutions that have been put forward. Finally, a set of recommendations for patients is provided, to help them assess risk more efficiently and make informed decisions.*

\* \* \*

Breast cancer screening reduces mortality by 20%. Therefore, we will save 1 out of every 5 women we screen. Right?

**WRONG.**

Do not worry if you got that wrong, it is very likely your doctor got it wrong too (and **that** is something we should all worry about).

## Introduction

The current COVID-19 pandemic increased the already high demand for evidence-based healthcare (Djulbegovic and Guyatt, 2020). Evidence-Based Medicine (making clinical decisions based on the best scientific evidence available and not on “common sense” or on your mentor’s opinion) has revolutionised and improved medical practice (Sackett et al., 2007). However, it brought a challenge we have not been able to conquer successfully. That challenge is: medical doctors need to learn how to interpret probabilities and statistics because, of course, it is not possible to practise

Evidence-Based Medicine if clinicians cannot understand or at least interpret scientific evidence.

To become a medical doctor, it is necessary to acquire knowledge and skills in many different subjects such as Physiology, Pharmacology, Public Health, and Clinical Examination. Among these skills, statistical reasoning is critical not only because it is essential to analyse emerging evidence from all other biomedical sciences, but also because it is required for risk assessments at bedside and for advising patients.

Although most medical schools recognise the importance of these skills and include Statistics in their curricula, evidence shows that even top and experienced clinicians struggle with assimilating the differences and implications of fundamental statistical concepts such as odds ratio versus absolute risk and sensitivity versus positive post-test probability (Jenny, Keller and Gigerenzer, 2018). Moreover, useful concepts such as absolute risk changes, number needed to treat/screen, intention-to-treat analysis and Bayesian probability are often overlooked when making clinical decisions and when explaining the implications of tests and treatments to patients (Naylor, Chen and Strauss, 1992; Whiting et al., 2015).

As a result, the ethical principles of autonomy and beneficence are threatened because patients are frequently exposed to unnecessary risks (which they frequently ignore), taxpayers' money is wasted, and misinformation is widely spread. Even further, some studies report discrepancies between the interventions and treatments that clinicians prescribe and the ones that they undergo when they face the same diseases (Slevin et al., 1990; Smith et al., 1998; Gallo et al., 2003).

This review intends to illustrate how statistical illiteracy generates ethical problems in the patient-clinician relationship or in public health decision-making, and to summarise existing evidence about this problem. Finally, a set of recommendations for patients is provided to help them assess risk more efficiently and make informed decisions.

## 1. The Origins and Size of the Problem

First, we should highlight that most clinicians want to help their patients. This problem does not originate in them although they are, in many ways, victims of its consequences. Gigerenzer and Gray (2011) proposed “seven sins” in modern medical practice: biased funding; reporting in medical journals, patient pamphlets, and the media; conflicts of interest; defensive medicine; and medical curricula that fail to teach doctors how to interpret health statistics.

Evidence-Based Medicine has evolved and become increasingly complex in the last two decades (Djulbegovic and Guyatt, 2017). The rules and conventions for carrying out and reporting different types of medical studies (The EQUATOR Network | Enhancing the QUALity and Transparency Of Health Research) or for evaluating their quality (Higgins et al., 2011) have been updated many times in the last two decades. Nowadays, the critical assessment of a study requires sound methodological knowledge, practice, and time. Even when most clinicians can correctly identify different types of studies, different risk outcomes or different hypothesis tests, studies indicate that many do not

understand key concepts and can be manipulated by misleading statistical formats (Jenny, Keller and Gigerenzer, 2018).

Of course, Science faces its own crisis today (Ioannidis, 2005; Baker and Penny, 2016) and the lack of good quality evidence on many clinical questions has direct implications in medical practice. However, elaborating on these biomedical-research-specific problems is outside the scope of this review.

Even if we naively assume that a certain research has been flawless, many problems arise when we report its results. Let's say that a treatment reduces the probability of getting disease "A" from 10 to 5 in 1,000, while it increases the risk of disease "B" from 5 to 10 in 1,000. Very frequently, a journal article reports the benefit as a 50% risk reduction and the harm as an increase of 5 in 1,000, that is, 0.5%. According to Sedrakyan and Shih (2007), this kind of mismatch (reporting some things as relative risks while reporting others as absolute risks) is present in 33% of papers in top medical journals and influences the way clinicians assimilate these data.

One may think that a smart clinician would not miss these differences. However, most final-year students from a top medical school failed an exam evaluating their competence when applying these concepts in practical scenarios (Jenny, Keller and Gigerenzer, 2018). Or, one may think that experience will eventually teach clinicians how to interpret data correctly. However, senior gynaecologists also fail to interpret the real risk meaning that a mammogram result implies (Anderson et al., 2014).

Alternatively, one may think that these problems concern exclusively new or infrequent diseases and treatments. However, evidence of risk misunderstanding by clinicians has been found in scenarios as frequent in everyday medical practice as cancer screening (Wegwarth and Gigerenzer, 2018).

Analogous to the difference between strict illiteracy and functional illiteracy (Tóth, 2001), the problem is not that clinicians lack statistical training or that they ignore the concepts required to read a scientific study. Most clinicians are familiar with the theory behind the statistical methods that the studies in their field use. The problem is that many clinicians struggle with efficiently incorporating the emerging evidence they read into their every-day practice. Clinicians need to learn not only the necessary statistical concepts and lexicon to read a study, but they also need to learn how to apply what they read when advising their patients.

## **2. Clinicians' statistical illiteracy exposes patients to unnecessary risks (which they frequently ignore)**

Unarguably, it would be unethical to hide potential adverse effects when obtaining informed consent for a medical procedure (Braschi et al., 2020). The ethical principle of autonomy protects a patient's right to make all decisions concerning their health. Clinicians are supposed to facilitate information about the potential risks and benefits of available interventions to their patients so they can make an informed decision (Entwistle et al., 2010). In practice, this is involuntarily hampered if clinicians ignore or do not understand the risks involved in the interventions they offer.

Cancer screening is an illustrative scenario to continue developing the previous example about the interpretation of different risk markers. Now, let's say – and this has already been argued in scientific publications – that performing breast cancer screening using mammograms has been reported to reduce breast cancer mortality by 20% (Elmore et al., 2005).

Does this imply that we will save 1 out of every 5 women who undergo the test? Well, many clinicians think so (Anderson et al., 2014). However, this is another example where reporting relative and not absolute risk reduction is misleading. It is indeed true that 20% is the relative risk reduction that corresponds to an absolute risk reduction from 5 to 4 out of every 1,000 women. However, saying the former sounds more impactful than the latter. In reality, we need to test 1,000 women to save 1 and some studies have estimated this number to be as high as 2,000 tests for every woman saved (Gøtzsche and Jørgensen, 2013).

Saving 1 woman in every 2,000 tests could still be considered a success because breast cancer is the most common cancer in women (Torre et al., 2017). However, mammograms also have risks, of which the most important is overdiagnosis (Nelson et al., 2016). Surprising as it may seem, it is almost 10 times more likely that a positive (abnormal) mammogram is a false positive than a true positive.

How is this possible? For women in their 40s, the sensitivity of a mammogram is 75% and their false positive rate is 10% (Medical Advisory Secretariat, 2007). Does this mean that if the test is positive, the probability of having cancer is 75%? No. How likely you are to have a disease if you had a positive result in a test is known as the “positive predictive value”, which is not the same as how likely you are to test positive in a test if you have a disease. Positive predictive value is heavily influenced by how prevalent a disease is, in this case, 1.4% for women in their 40s, and is very often mistaken by clinicians as the sensitivity of a test (Whiting et al., 2015).

In other words, in a group of 1,000,000 women, 14,000 have breast cancer. Therefore, 986,000 of them do not have breast cancer. Of the 14,000 women who have breast cancer, 75% (10,500) will be detected by the mammogram. However, of the 986,000 women without breast cancer, 10% (98,600) will be told that they have breast cancer when they do not. Thus, after performing 1,000,000 tests, there will be 10,500 true positive and 98,600 false positive tests. Therefore, a positive result is almost 10 times more likely to belong to the 98,600 group rather than to the 10,500 group.

Apart from stress and anxiety, a false positive test also entails biopsies, potential surgery, and even more false positive results. According to Elmore et al. (2015), pathologists (medical doctors who are experts in analysing biopsies and the current gold standard for diagnosing breast cancer) disagree 25% of the time when they analyse the same breast biopsies.

Prostate cancer screening is a similar scenario. The number needed to screen to save 1 man is 1,254 (Loeb et al., 2011), and 7-10 out of 100 men who undergo a biopsy will require to be hospitalised due to complications of the procedure (Brewster et al., 2017).

Does this mean that breast and prostate cancer screenings are useless and that we should stop them? No, because if your mother and your aunts had breast cancer or if your father and your uncles had prostate cancer, getting screened could save your life

(Mitra et al., 2011; Bae et al., 2020). This just means that even seemingly harmless procedures need to be individually assessed in terms of benefit/risk ratio; which, of course, is impossible if clinicians do not understand them. Additionally, and as discussed in the next section, risk misunderstanding results in unnecessary risks and expenses at the population level as well.

### 3. Clinicians become (or are relied upon by) decision makers

“I had prostate cancer five, six years ago. My chance of surviving prostate cancer — and, thank God, I was cured of it — in the United States? Eighty-two percent. My chance of surviving prostate cancer in England? Only 44 percent under socialized medicine.” (Bosman, 2007).

The difference between a five-year survival rate, mortality, lethality, and overall survival is very frequently misunderstood by decision makers (as in the example above) and by primary care physicians (Wegwarth et al., 2012).

Yes, all these concepts and indicators are related to death and cancer, but they are not equivalent or even correlated because you can be diagnosed with cancer but die of something else. Additionally, tumours have very different rates of prognosis. For example, you can falsely inflate five-year survival rates just by diagnosing earlier and without raising the number of people not dying of cancer.

The problems that arise from confusions like these go beyond the misuse of technical language. These indicators have confused decision makers about the net benefits of screening, and there are well documented examples of taxpayers’ money being misspent (Iacobucci, 2018). Of course, this money could have been spent in more effective practices to prevent cancer-related deaths.

### 4. Statistical illiteracy makes clinicians break ‘the Golden Rule’: *‘Treat others as you would like to be treated’*

Studies show that clinicians often choose to be treated differently to the way they treat patients with the same diseases or under the same circumstances (Smith et al., 1998; Gallo et al., 2003). Interestingly, clinicians who undergo the diseases that they typically treat change their practice significantly after they recover (Cen, 2015).

As most medical doctors genuinely want to help their patients, it is unlikely that they willingly mislead them when explaining treatments. At the core of these discrepancies in the treatments that are chosen, we find a different form of statistical illiteracy. In this case, clinicians are not confused by statistical jargon, they simply cannot assess the risk/benefit ratio adequately due to the lack of patient-oriented evidence.

When designing clinical trials, we usually choose hard clinical outcomes as the primary objective of our study. We opt to look at survival rate, hospitalisation rate, and years in clinical remission, while often ignoring softer outcomes such as patient satisfaction. Therefore, we tend to base our recommendations on evidence including clinically oriented outcomes and not patient-oriented outcomes.

An example of this is the very high regret rate of patients undergoing dialysis (Davison, 2010). We tend to recommend it based on the real and well-documented clinical benefits without mentioning the also real and well-documented high proportion of patients who regret undergoing dialysis.

In other words, healthcare workers are treated differently because, when choosing treatments for themselves, they can (at least subjectively) weigh in these patient-oriented outcomes based on what they see in their practice (Slevin et al., 1990). On the other hand, when advising their patients, clinicians cannot incorporate these factors in their risk assessments. They need to adhere to the available evidence, and patient-oriented outcomes are understudied or underreported.

### **5. Clinicians can learn when taught properly**

After documenting that most final-year students fail a “translating evidence into practice” test, researchers showed that the same students can ace a similar test after a short course (Jenny, Keller and Gigerenzer, 2018). Additionally, evidence suggests that graphic aids improve the way surgeons communicate procedures’ risks and benefits (Garcia-Retamero et al., 2016).

Thus, since clinicians are usually meticulous students and well-intended people, there is fertile ground to improve this situation. Again, Gigerenzer and Gray (2011), when launching “The century of the patient”, proposed seven goals: funding for research relevant for patients; transparent and complete reporting in medical journals, health pamphlets, and the media; incentive structures that minimise conflicts of interest; promoting better practice instead of defensive medicine; and doctors who understand health statistics.

Patients have little influence on the way research funds are allocated or on the way the media displays scientific news. However, they can overcome most of the problems mentioned above by improving communication with their clinicians and by seeking critical information that is often overlooked.

Statistical illiteracy among clinicians is frequent and has widespread repercussions. It threatens the autonomy principle in shared decision-making and the beneficence principle when allocating tax expenditure for healthcare or when funding medical research that looks into the patients’ best interests with patient-oriented outcomes. Medical schools, medical journals, patients, and the clinicians themselves should acknowledge this problem and do their part in solving it.

### **6. Advice for patients**

1. Be an active patient, do not be afraid to do your own research on whatever illness or health question you may have. Ask your clinician about reliable patient-education resources. If your doctor is not open to questions or cannot admit that they do not know something, get a different one.
2. Ask your clinician to disclose risks and benefits as “the number needed to” (i.e. number needed to treat, to help (or to harm) one person). When

undergoing any diagnostic test, ask about positive and negative predictive values, not about sensitivity or specificity.

3. Ask if there is available evidence about regret rate or would-do-it-again rate and studies including patient-oriented outcomes.
4. Accept the fact that, even in Evidence-Based Medicine, uncertainty is very common and sometimes, the best we have is “an educated guess”.
5. Be patient with your doctor, they do not have every piece of evidence at the top of their head and they might struggle with counter-intuitive statistical concepts.
6. Be patient with scientists, you will find that many questions relevant to your specific case have not been answered yet. Volunteer for research whenever you can!
7. Be patient with yourself, being ill or being unsure about medical decisions is perfectly normal. Keeping a record of your questions and feelings can be very useful for you and your doctor when facing a difficult decision.

*Department of Physiology, Anatomy and Genetics*

## **Bibliography**

- Anderson, B. L. *et al.* (2014) ‘Statistical Literacy in Obstetricians and Gynecologists’, *Journal For Healthcare Quality*, 36(1), pp. 5–17. doi: 10.1111/j.1945-1474.2011.00194.x.
- Bae, M. S. *et al.* (2020) ‘Survival Outcomes of Screening with Breast MRI in Women at Elevated Risk of Breast Cancer’, *Journal of Breast Imaging*. Oxford University Press (OUP), 2(1), pp. 29–35. doi: 10.1093/jbi/wbz083.
- Baker, M. and Penny, D. (2016) ‘Is there a reproducibility crisis?’, *Nature*. Nature Publishing Group, pp. 452–454. doi: 10.1038/533452A.
- Bosman, J. (2007) ‘Giuliani’s Prostate Cancer Figure Is Disputed’, *The New York Times*. Available at: <https://www.nytimes.com/2007/10/31/us/politics/31prostate.html> Published on October 31<sup>st</sup>, 2007.
- Braschi, E. *et al.* (2020) ‘Evidence-based medicine, shared decision making and the hidden curriculum: a qualitative content analysis’, *Perspectives on Medical Education*. Bohn Stafleu van Loghum, 9(3), pp. 173–180. doi: 10.1007/s40037-020-00578-0.
- Brewster, D. H. *et al.* (2017) ‘Risk of hospitalization and death following prostate biopsy in Scotland’, *Public Health*. Elsevier B.V., 142, pp. 102–110. doi: 10.1016/j.puhe.2016.10.006.
- Cen, P. (2015) ‘What my Cancer Taught Me’, *Baylor University Medical Center Proceedings*. Informa UK Limited, 28(4), pp. 526–527. doi: 10.1080/08998280.2015.11929332.

- Davison, S. N. (2010) 'End-of-life care preferences and needs: Perceptions of patients with chronic kidney disease', *Clinical Journal of the American Society of Nephrology*. American Society of Nephrology, 5(2), pp. 195–204. doi: 10.2215/CJN.05960809.
- Djulgovic, B. and Guyatt, G. (2020) 'Evidence-based medicine in times of crisis', *Journal of Clinical Epidemiology*. Elsevier, 0(0). doi: 10.1016/j.jclinepi.2020.07.002.
- Djulgovic, B. and Guyatt, G. H. (2017) 'Progress in evidence-based medicine: a quarter century on', *The Lancet*. Lancet Publishing Group, pp. 415–423. doi: 10.1016/S0140-6736(16)31592-6.
- Elmore, J. G. *et al.* (2005) 'Screening for breast cancer', *Journal of the American Medical Association*. NIH Public Access, pp. 1245–1256. doi: 10.1001/jama.293.10.1245.
- Elmore, J. G. *et al.* (2015) 'Diagnostic concordance among pathologists interpreting breast biopsy specimens', *JAMA - Journal of the American Medical Association*. American Medical Association, 313(11), pp. 1122–1132. doi: 10.1001/jama.2015.1405.
- Entwistle, V. A. *et al.* (2010) 'Supporting patient autonomy: The importance of clinician-patient relationships', *Journal of General Internal Medicine*. Springer, pp. 741–745. doi: 10.1007/s11606-010-1292-2.
- Gallo, J. J. *et al.* (2003) 'Life-Sustaining Treatments: What Do Physicians Want and Do They Express Their Wishes to Others?', *Journal of the American Geriatrics Society*, 51(7), pp. 961–969. doi: 10.1046/j.1365-2389.2003.51309.x.
- Garcia-Retamero, R. *et al.* (2016) 'Improving risk literacy in surgeons', *Patient Education and Counseling*. Elsevier Ireland Ltd, 99(7), pp. 1156–1161. doi: 10.1016/j.pec.2016.01.013.
- Gigerenzer, G. and Gray, J. A. M. (2011) *Launching the Century of the Patient, Better Doctors, Better Patients, Better Decisions*. Edited by G. Gigerenzer and J. A. M. Gray. The MIT Press. doi: 10.7551/mitpress/9780262016032.001.0001.
- Gøtzsche, P. C. and Jørgensen, K. J. (2013) 'Screening for breast cancer with mammography', *Cochrane Database of Systematic Reviews*. John Wiley and Sons Ltd. doi: 10.1002/14651858.CD001877.pub5.
- Higgins, J. P. T. *et al.* (2011) 'The Cochrane Collaboration's tool for assessing risk of bias in randomised trials', *BMJ (Online)*. British Medical Journal Publishing Group, 343(7829). doi: 10.1136/bmj.d5928.
- Iacobucci, G. (2018) 'Conservative conference: May announces new cancer strategy to boost survival rates', *BMJ (Clinical research ed.)*. NLM (Medline), 363, p. k4198. doi: 10.1136/bmj.k4198.
- Ioannidis, J. P. A. (2005) 'Why Most Published Research Findings Are False', *PLoS Medicine*. Springer International Publishing, 2(8), p. e124. doi: 10.1371/journal.pmed.0020124.
- Jenny, M. A., Keller, N. and Gigerenzer, G. (2018) 'Assessing minimal medical statistical literacy using the Quick Risk Test: A prospective observational study in Germany', *BMJ Open*. BMJ Publishing Group, 8(8), p. e020847. doi: 10.1136/bmjopen-2017-020847.

- Loeb, S. *et al.* (2011) 'What is the true number needed to screen and treat to save a life with prostate-specific antigen testing?', *Journal of Clinical Oncology*. American Society of Clinical Oncology, 29(4), pp. 464–467. doi: 10.1200/JCO.2010.30.6373.
- Medical Advisory Secretariat (2007) 'Screening mammography for women aged 40 to 49 years at average risk for breast cancer: an evidence-based analysis.', *Ontario health technology assessment series*. Health Quality Ontario, 7(1), pp. 1–32. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/23074501>.
- Mitra, A. V. *et al.* (2011) 'Targeted prostate cancer screening in men with mutations in BRCA1 and BRCA2 detects aggressive prostate cancer: Preliminary analysis of the results of the IMPACT study', *BJU International*. BJU Int, 107(1), pp. 28–39. doi: 10.1111/j.1464-410X.2010.09648.x.
- Naylor, C. D., Chen, E. and Strauss, B. (1992) 'Measured enthusiasm: Does the method of reporting trial results alter perceptions of therapeutic effectiveness?', *Annals of Internal Medicine*. Ann Intern Med, 117(11), pp. 916–921. doi: 10.7326/0003-4819-117-11-916.
- Nelson, H. D. *et al.* (2016) 'Harms of breast cancer screening: Systematic review to update the 2009 U.S. Preventive services task force recommendation', *Annals of Internal Medicine*. American College of Physicians, pp. 256–267. doi: 10.7326/M15-0970.
- Sackett, D. L. *et al.* (2007) 'Evidence based medicine: what it is and what it isn't. 1996.', *Clinical orthopaedics and related research*. British Medical Journal Publishing Group, 455(7023), pp. 3–5. doi: 10.1136/bmj.312.7023.71.
- Sedrakyan, A. and Shih, C. (2007) 'Improving depiction of benefits and harms: Analyses of studies of well-known therapeutics and review of high-impact medical journals', *Medical Care*, 45(10 SUPPL. 2), pp. S23-8. doi: 10.1097/MLR.0b013e3180642f69.
- Slevin, M. L. *et al.* (1990) 'Attitudes to chemotherapy: Comparing views of patients with cancer with those of doctors, nurses, and general public', *British Medical Journal*. BMJ Publishing Group, 300(6737), pp. 1458–1460. doi: 10.1136/bmj.300.6737.1458.
- Smith, T. J. *et al.* (1998) 'Would oncologists want chemotherapy if they had non-small-cell lung cancer?', *ONCOLOGY*, pp. 360–365.
- The EQUATOR Network | Enhancing the QUALity and Transparency Of Health Research* (no date). Available at: <https://www.equator-network.org/>.
- Torre, L. A. *et al.* (2017) 'Global cancer in women: Burden and trends', *Cancer Epidemiology Biomarkers and Prevention*. American Association for Cancer Research Inc., pp. 444–457. doi: 10.1158/1055-9965.EPI-16-0858.
- Tóth, I. G. (2001) 'Literacy and Illiteracy, History of', in *International Encyclopedia of the Social & Behavioral Sciences*. Elsevier, pp. 8961–8967. doi: 10.1016/b0-08-043076-7/02738-8.
- Wegwarth, O. *et al.* (2012) 'Do physicians understand cancer screening statistics? A national survey of primary care physicians in the United States', *Annals of Internal Medicine*. American College of Physicians, 156(5), pp. 340–349. doi: 10.7326/0003-4819-156-5-201203060-00005.

Wegwarth, O. and Gigerenzer, G. (2018) 'The barrier to informed choice in cancer screening: Statistical illiteracy in physicians and patients', in *Recent Results in Cancer Research*. Springer New York LLC, pp. 207–221. doi: 10.1007/978-3-319-64310-6\_13.

Whiting, P. F. *et al.* (2015) 'How well do health professionals interpret diagnostic information? A systematic review', *BMJ Open*. BMJ Publishing Group. doi: 10.1136/bmjopen-2015-008155.