Intensive Care Med (2019) 45:1447–1450
https://doi.org/10.1007/s00134-019-05715-w

LESS IS MORE IN ICU

When will less monitoring and diagnostic testing benefit the patient more?

Fernando G. Zampieri[1*] and Sharon Einav[2]

The choosing wisely campaign has highlighted for each medical profession the five practices that both physicians and patients should question (http://www.choosingwisely.org/). Achieving informed test selection was named as one of the five challenges our profession should address in the years to come. However, the means of achieving this aim remains unclear. Some call for real-time disclosure of the costs and consequences of excessive testing. Others believe in better education on the topic. Both approaches are challenging, and neither is likely to suffice alone. There is also a need to change the medical system and societal expectations from a good doctor.

We perform many tests simply because we can and because we hesitate to change longstanding routines. We also overtest because we are concerned that we might miss an important finding that will ultimately affect patient survival [1]. However, excessive testing carries a heavy price. Over half of the intensive care unit (ICU) patients are already anemic at the time of ICU admission. "Routine" blood sampling for a complete blood count, short biochemistry and clotting mechanism requires drawing of ~ 10 cc of blood, which corresponds to a decrease of 0.7 g/l (± 95% CI 0.5, 0.9) in hemoglobin. "Routine" sampling over 5 days thus entails a blood loss of ~ 50 cc. Add to this three samples of cardiac enzymes, an expanded biochemistry test and perhaps (by this time) a brief anemia workup, and the patient has lost 200 cc of blood within a week. Anemia during hospital admission is undeniably tied to frequent blood draws [2]. This could be easily overcome with blood conservation devices and smaller test tubes but these entail increased expenditure. Conversely, less blood testing is sound from both an economic and patient outcome perspective. It would also prevent squandering of millions of liters of blood, exceeding more than four times the total volume of blood transfused annually [3].

Overtesting is not only harmful per se. It can lead to harmful consequences. Routine repetitive ICU chest radiographs rapidly increases the probability of a false positive result. Abnormal unexpected findings can trigger additional potentially redundant testing (e.g., patient transfer to computed tomography) or, even worse, procedures. Conversely, the results of redundant tests may be ignored, particularly if they are controversial and may misinform the clinician. Chest radiographs have a low rate of interobserver agreement [4]. It is thus unsurprising that most routine chest radiographs do not alter clinical management even when abnormalities are revealed [4] and eliminating them affects neither ICU nor hospital mortality or length of stay [5]. Similarly, miscalibrated arterial lines (especially when underdamping is present) can be more dangerous than no arterial line at all [6].

Monitoring is a euphemism for high-frequency testing. The same test may be used for monitoring and for diagnosis (e.g., blood pressure, electrocardiography). A test can also be used for disease diagnosis and for monitoring the response to treatment (e.g., CRP). Monitors and tests differ in that monitors are supposed to be highly sensitive, whereas tests are expected to be specific. But the use of different cutoffs (i.e., favoring either specificity or sensitivity) does not elevate the test above the limitations of repetitive testing. Monitoring, by definition, must be (nearly) continuous to detect acute changes. The question is whether simple "test repetition" (e.g., beat-by-beat heart rate and blood pressure monitoring) yields the optimal results. The price of (nearly) continuous monitoring is an exceptionally high rate of false alarms (> 85%) [7]. Signals are more likely

*Correspondence: fgzampieri@gmail.com
[1] Research Institute, HCor-Hospital do Coração, São Paulo, Brazil
Full author information is available at the end of the article

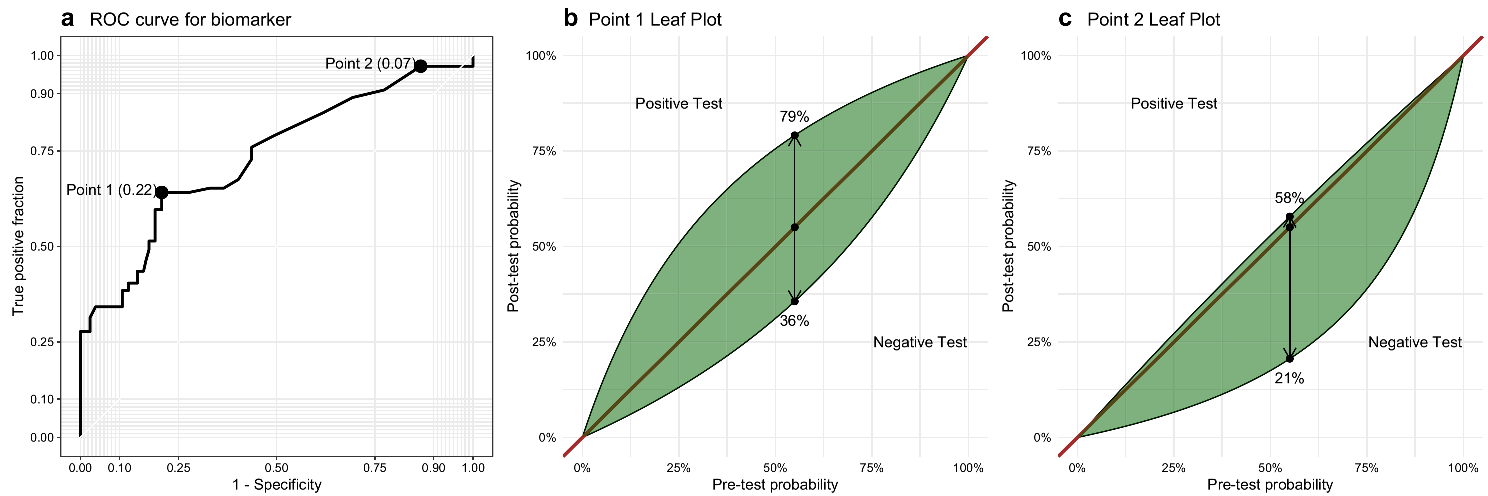

Fig. 1 Interpretation of biomarker levels: a receiver operating curve reflecting the accuracy of a theoretical biomarker for predicting an outcome is shown in panel (**a**). Biomarker levels are continuous variables, therefore each point selected on this graph (i.e. a selected cutoff) has its intrinsic sensitivity, specificity and negative and positive likelihood ratios. We randomly selected two points on the hypothetical ROC curve to illustrate how selection of different cutoffs can lead to different conclusions. The first point (point 1; biomarker value of 0.22) marks the "best" cutoff as defined by Youden´s index. The second point (point 2, biomarker value of 0.07) marks the cutoff that penalizes false negative results five times more than false positive results. Leaf plots display the relationship between the assumed pretest probability (on the *x*-axis) and the pretest probability at the cutoff selected (on the *y*-axis). Two curves are presented, one for positive test results (upper half) and one for negative test results (lower half); a useless test would stand in a perfect 45° line (pretest equals posttest). We constructed leaf plots [13] for points 1 and 2. These two leaf plots (panels **b** and **c,** respectively) were created using the same dataset; only the cutoff point differs. We then (also randomly) selected a pretest probability of 55% to illustrate the following: panel **b** demonstrates the posttest probability considering the baseline 55% probability and a positive test (posttest probability of 79%) or negative text (posttest probability of 36%). Panel **c** demonstrates the other extreme in which testing is used to "rule out" the presence of a finding. In this scenario, a negative result typically reduces the posttest probability (from 55 to 21%), but a positive finding is of little value (from 55 to 58%)

to be missed when they appear frequently [8], particularly in the presence of noise [9]. Compound this with a very human tendency to ignore that which we would prefer not to see [10] and it becomes unsurprising that separate monitoring of multiple parameters eventually leads to alarm fatigue [11]. In other words, although a medical test or a monitor signal can be the tool for justifying a choice of action, it can also yield findings that are entirely irrelevant or unrelated to the reason it was used in the first place. Humans also have a working memory limited to 3–4 interacting variables and this ability also depends on the cognitive load at any given time [12]. Hopefully, in the future, better algorithms will be developed to pool data from multiple sources to a single alarm and to distinguish signal from noise.

Testing and monitors both serve the same purpose: provision of grounds for decisions. A test or a monitor should therefore be capable of changing a prior probability to a posterior probability [13]. The capacity to affect such a change should drive the decision to use any test or monitor. For this reason, physicians should be capable of estimating pretest probability; in fact, this is among the noblest actions in patient care. Whether a test can have a clinically meaningful effect on treatment can be estimated. An example calculation of appropriate versus inappropriate testing is shown in Fig. 1 with the accompanying supplemental R script used to draw a leaf plot (Fig. 1). Ideally, tools such as this should be made user friendly and be provided to the clinician for bedside use. As it is, rarely is the capacity to change a prior probability at the forefront of our thought when choosing to use a test or a monitor. The roots of this oversight run deep; it is testing and monitoring that led to the development of intensive care as we know it today rather than vice versa.

When exponential developments in medical devices ushered in the era of modern intensive care units and critical illness, the very edge of life suddenly became measurable, interpretable, and almost transparent. The addition of every new device was accompanied by receipt of fascinating new information. As survival from critical illness gradually became more than a fluke of chance, it also seemed logical that more data would foster better outcomes. With the arrival of computers, the ICU quickly became not only the location of scientific information about patient condition but also a major repository of data [14]. Yet, the conjuncture of human drama and almost unlimited data with our very human cognitive limitations and biases eventually created habits and beliefs that we are still struggling to banish. Among these is the habit of torturing ourselves (and our patients) with excessive data, information overload and massive amounts of potentially false results.

The reasons for this are many: first, we believe that correcting abnormalities will improve prognosis. It therefore seems logical that if we can correct as many abnormalities as possible, the better off the patient will be [15]. However, the evidence does not point towards such an association. Second, like most humans, intensivists are averse to dread [16]. The medical version of this phenomenon is manifested in the ordering of multiple tests in an unconscious effort to obtain relief from the anxiety of missing a piece of information that could be pivotal for patient outcome [17]. Third, the ideal of the totipotent physician who cares zealously for his patients is easily translatable to constant testing and monitoring [18].

Changing these concepts requires development of appropriate decision-support tools for testing. It also requires educating the next generation of intensivists regarding the risk–benefit ratios of testing, allowing generation of realistic and practical expectations from test results. These need to be accompanied by public (and legal) education. At the same time, care must be taken to avoid replacing a dogma of "more is better" with a dogma of "less is better". A smooth transition from overtesting to effective and enough testing rather than just efficient testing can only be performed with generation of strong scientific evidence showing exactly less of what is safe and for whom.

Electronic supplementary material

The online version of this article (https://doi.org/10.1007/s00134-019-05715-w) contains supplementary material, which is available to authorized users.

Author details

[1] Research Institute, HCor-Hospital do Coração, São Paulo, Brazil. [2] General Intensive Care Unit of the Shaare Zedek Medical Center, Faculty of Medicine, The Hebrew University, Jerusalem, Israel.

Acknowledgements

The authors would like to thank Lucas P Damiani for his statistical assistance in preparing this manuscript.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 18 May 2019 Accepted: 25 July 2019
Published online: 2 August 2019

References

1. Eaton KP, Levy K, Soong C, Pahwa AK, Petrilli C, Ziemba JB, Cho HJ, Alban R, Blanck JF, Parsons AS (2017) Evidence-based guidelines to eliminate repetitive laboratory testing. JAMA Intern Med 177(12):1833–1839
2. Ranasinghe T, Freeman WD (2014) 'ICU vampirism'-time for judicious blood draws in critically ill patients. Br J Haematol 164(2):302–303
3. Levi M (2014) Twenty-five million liters of blood into the sewer. J Thromb Haemost 12(10):1592

4. Ruza GC, Moritz RD, Machado FO (2012) Routine chest radiography in intensive care: impact on decision-making. Rev Bras Ter Intensiva 24(3):252–257
5. Oba Y, Zaza T (2010) Abandoning daily routine chest radiography in the intensive care unit: meta-analysis. Radiology 255(2):386–395
6. Romagnoli S, Ricci Z, Quattrone D, Tofani L, Tujjar O, Villa G, Romano SM, De Gaudio AR (2014) Accuracy of invasive arterial pressure monitoring in cardiovascular patients: an observational study. Crit Care 18(6):644
7. Drew BJ, Harris P, Zègre-Hemsey JK, Mammone T, Schindler D, Salas-Boni R et al (2014) Insights into the problem of alarm fatigue with physiologic monitor devices: a comprehensive observational study of consecutive intensive care unit patients. PLoS One 9(10):e110274
8. Jerison HJ, Pickett RM (1964) Vigilance: the importance of the elicited observing rate. Science 143(3609):970–971
9. Hockney GR (1978) Effects of noise on human work efficiency. In: May DN (ed) Handbook of noise assessment. Van Nostrand Reinhold, New York
10. Teo G, Schmidt T, Szalma J, Hancock G, Hancock P (2014) The effects of individual differences on vigilance training and performance in a dynamic vigilance task. Proc Hum Factors Ergon Soc Annu Meet 58(1):964–968
11. Citing reports of alarm-related deaths, the Joint Commission issues a sentinel event alert for hospitals to improve medical device alarm safety. ED Manag. 2013;26(6):suppl 1–3
12. Halford GS, Baker R, McCredden JE, Bain JD (2005) How many variables can humans process? Psychol Sci 16(1):70–76
13. Coulthard MG, Coulthard T (2019) The leaf plot: a novel way of presenting the value of tests. Br J Gen Pract 69(681):205–206
14. McIntosh N (2002) Intensive care monitoring: past, present and future. Clin Med 2(4):349–355
15. Kavanagh BP, Meyer LJ (2005) Normalizing physiological variables in acute illness: five reasons for caution. Intensive Care Med 31(9):1161–1167
16. Dawson Chris, de Meza David Emmanuel (2018). Wishful thinking, prudent behavior: the evolutionary origin of optimism, loss aversion and disappointment aversion. https://ssrn.com/abstract=3108432 or http://dx.doi.org/10.2139/ssrn.3108432. Accessed 24 Jan 2018
17. Salkovskis P (1991) The importance of behaviour in the maintenance of anxiety and panic: a cognitive account. Behav Psychother 19(1):6–19
18. Kravitz RL, Callahan EJ, Paterniti D, Antonius D, Dunham M, Lewis CE (1996) Prevalence and sources of patients' unmet expectations for care. Ann Intern Med 125(9):730–737