



# Rationality, practice variation and person-centred health policy: a threshold hypothesis

Benjamin Djulbegovic MD PhD,<sup>1,2,3</sup> Robert M. Hamm PhD,<sup>4</sup> Thomas Mayrhofer PhD,<sup>5,6</sup> Iztok Hozo PhD<sup>7</sup> and Jef Van den Ende MD<sup>8</sup>

<sup>1</sup>Distinguished Professor, USF Health Comparative Effectiveness Research Core, Division of Evidence Based Medicine, Department of Internal Medicine, Morsani College of Medicine, University of South Florida, Tampa, FL, USA

<sup>2</sup>Distinguished Professor, H Lee Moffitt Cancer Center and Research Institute, Tampa, FL, USA

<sup>3</sup>Distinguished Professor, Tampa General Hospital, Tampa, FL, USA

<sup>4</sup>Professor, Department of Family and Preventive Medicine, University of Oklahoma Health Sciences Center, Oklahoma City, OK, USA

<sup>5</sup>Professor, Cardiac MR PET CT Program, Department of Radiology, Massachusetts General Hospital and Harvard Medical School, Boston, MA, USA

<sup>6</sup>Professor, School of Business Studies, University of Applied Sciences Stralsund, Stralsund, Germany

<sup>7</sup>Professor, Department of Mathematics, Indiana University Northwest, Gary, IN, USA

<sup>8</sup>Professor, Institute of Tropical Medicine, Antwerp, Belgium

## Keywords

epistemology, health policy, person-centred medicine

## Correspondence

Dr. Benjamin Djulbegovic  
USF Health Comparative Effectiveness  
Research Core, Division of Evidence Based  
Medicine, Department of Internal Medicine,  
Morsani College of Medicine  
University of South Florida  
12901 Bruce B. Downs Blvd, MDC02  
Tampa, FL 33612  
USA  
E-mail: bdjulbeg@health.usf.edu

Accepted for publication: 2 November 2015

doi:10.1111/jep.12486

## Abstract

Variation in practice of medicine is one of the major health policy issues of today. Ultimately, it is related to physicians' decision making. Similar patients with similar likelihood of having disease are often managed by different doctors differently: some doctors may elect to observe the patient, others decide to act based on diagnostic testing and yet others may elect to treat without testing. We explain these differences in practice by differences in disease probability *thresholds* at which physicians decide to act: contextual social and clinical factors and emotions such as regret affect the threshold by influencing the way doctors integrate objective data related to treatment and testing. However, depending on a theoretical construct each of the physician's behaviour can be considered rational. In fact, we showed that the current regulatory policies lead to predictably low thresholds for most decisions in contemporary practice. As a result, we may expect continuing motivation for overuse of treatment and diagnostic tests. We argue that rationality should take into account both formal principles of rationality and human intuitions about good decisions along the lines of Rawls' 'reflective equilibrium/considered judgment'. In turn, this can help define a threshold model that is empirically testable.

Current clinical practice is characterized by large variation: similar patients under similar conditions are frequently managed differently [1]. Although some practice variation is warranted [2], much of it results in either underuse or overuse of health interventions, yielding uneven health outcomes among similar patients, waste and high health care costs [3]. Aiming to reduce variation in patient care, current policy initiatives target exogenous factors such as improvement in coordination of care, appropriate use of medical technologies and financial incentives. Although these factors are undoubtedly important, most variation in care is a result of the way physicians make their decisions. The realization that human judgment is of profound importance for health and social policy [4] has also been highlighted by the US Institute of Medicine's recent report on variation of care, which concluded 'target

decision making, not geography'. [5] It follows then that if we want to improve health care and reduce costs, we should understand and improve the way doctors make decisions [4]. We hypothesize that the observed practice variation is related to individual physicians' differences in action *thresholds*, particularly when they act under conditions of diagnostic uncertainty. The threshold approach to decision making indicates that when faced with uncertainty about whether to order a test, or treat a patient who may or may not have a disease, there must exist some probability at which a physician is indifferent between administering versus not administering treatment [6], or observing the patient versus ordering a diagnostic test versus treating without testing [7]. According to the threshold model, the physicians should act when benefits of action (say, treatment or testing) outweigh its harms [6,7]. However,

integration of diagnostic accuracy information with treatment benefits and harms within the framework of the threshold model can occur via different cognitive and decision-making mechanisms that give rise to different theoretical accounts of the threshold model [8]. The question then becomes which one of these theoretical accounts can be accepted as the most rational behaviour in clinical practice?

There are numerous theories of decision making, generally grouped into three classes: normative, descriptive and prescriptive. *Normative* theories rely on mathematical analyses to help derive the optimal course of action. They typically employ expected utility theory (EUT), the basis of applied decision analysis. (Expected utility is the weighted average: combined utility of all possible health outcomes following a decision, weighted by the corresponding probability of these outcomes.) According to EUT, *rational choice* is associated with selection of the alternative with higher expected utility, such as higher quality-adjusted life years. Importantly, EUT is the only theory of choice that satisfies all mathematical axioms of rational decision making. Thus, it appears to make sense to use decision analysis at bedside and in policy decision making, as the US Preventive Services Task Force did in identifying the optimal test for colorectal cancer screening [9].

The problem is that decades of empirical research have demonstrated that people routinely violate the precepts of EUT, and thus do not behave according to EUT's standard of rationality. The *descriptive* theories of decision making attempt to explain why people may act differently from what they normatively should do ('is' vs. 'should'). In clinical medicine, this is often because decision making relies on intuition, heuristic cognitive strategies [10] and habit, and is shaped by emotions and other experiential and contextual factors integral to every clinical encounter.

Recognizing how people think, *prescriptive* decision-making theories seek ways to assist them in making approximately normative decisions. Thus, in medicine, prescriptive theories guide attempts to inculcate EUT into clinical practice. A convenient formulation of EUT, which accommodates physicians' assessment of the likelihood of disease and balancing of treatment's benefits and harms, is to act according to the 'threshold model' [7]. When considering ordering a diagnostic test [7], this model creates three decision fields, separated by two thresholds: if the disease probability is below the test threshold physicians should observe the patient, between the testing and the treatment thresholds they should order a test and act accordingly, and above the treatment threshold they should treat without further testing.

This framework is more tractable for clinical practice than more complex decision analysis, as it summarizes multiple considerations into a small set of ideas. The EUT threshold model stipulates that as the therapeutic benefit/harm ratio increases, the threshold probability at which treatment is justified is lowered [7]. Conversely, when treatment's benefit/harm ratio is smaller, the required threshold for therapeutic action will be higher [7]. For example, the benefit/harm ratio for treating someone with suspected pulmonary tuberculosis is anywhere between 15 and 36 [11]. With that EUT threshold, it is rational for physicians to administer anti-tuberculosis treatment when the probability of tuberculosis exceeds 2–6% [11,12]. However, as discussed earlier, doctors do not behave according to the EUT standard of rationality and physicians most frequently indicate that they would not treat a patient suspected of tuberculosis below a threshold between 20%

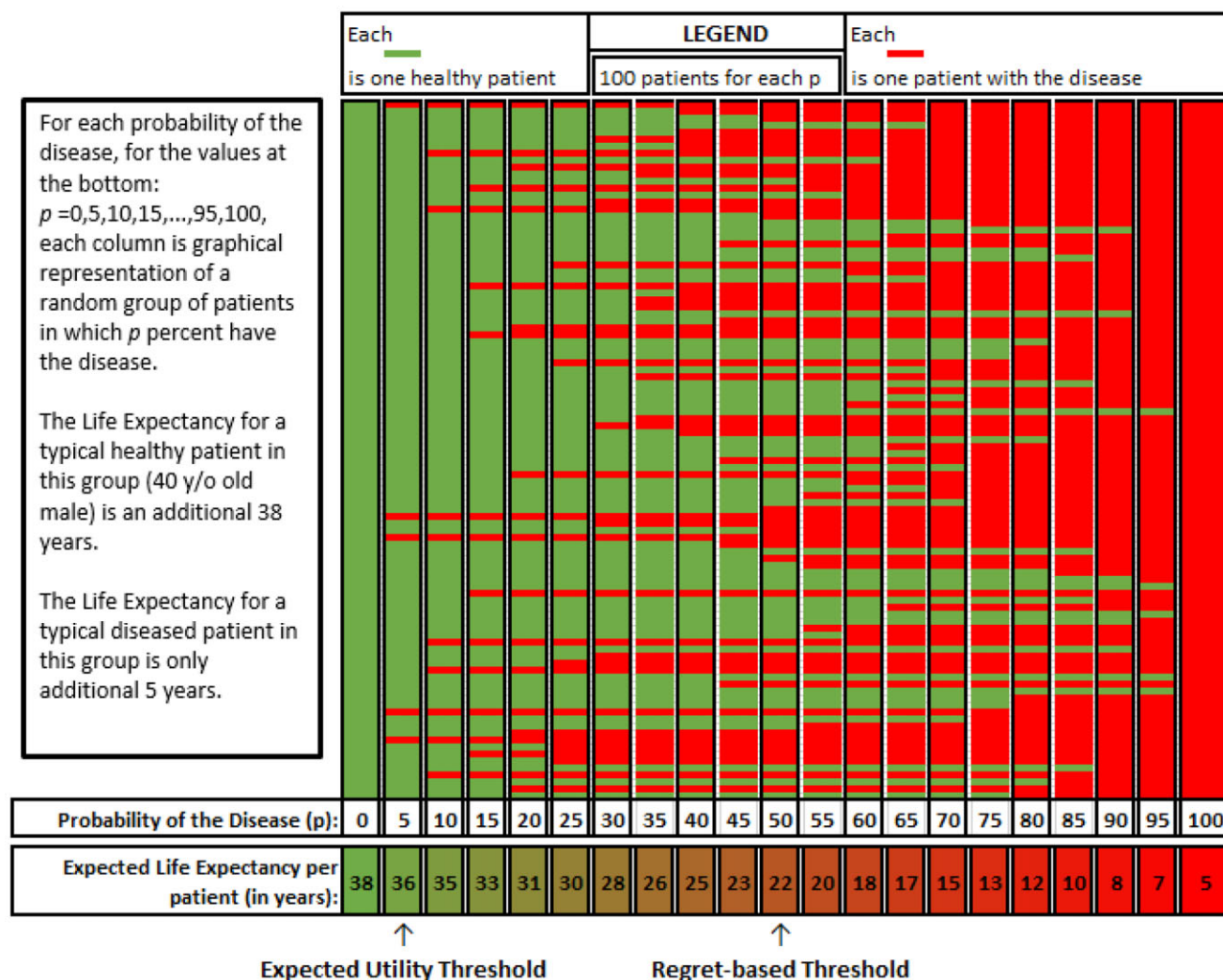
and 50% [11,13]. Despite the fact that expected utility is greatest when all patients above this low probability (2–6%) are treated, these doctors fear that many patients with probability of tuberculosis above 6% will be 'false positives' and thus receive unnecessary treatment (Fig. 1). Similar variation in treatment thresholds has been observed in decisions about pulmonary embolism, malaria, allogeneic transplant for acute leukaemia and H1N1 vaccine [8]. In some cases, such as decision to undergo transplant, physicians have only 'one shot' to make their decisions; in other settings such as management of chronic diseases, doctors may re-evaluate their decisions based on the patient's clinical features. In either case, the EUT theoretical threshold for action often dramatically differs from the physicians' personal thresholds.

Variation in decision making can be explained using the threshold model as a descriptive theory: physicians act as if they refer to *different thresholds*, which in turn could be due to different ways of cognitively assessing disease probability or the consequences of treatments, or of integrating treatment benefits and harms. This directs attention to the factors that may affect each of these processes, such as emotions, regret of omission versus commission, financial incentives, poor quality evidence, application of average data to individual patients and individual differences in subjective judgments of risk assessment and disease prevalence. Understanding cognitive mechanisms that affect action thresholds may have important policy implications and can complement approaches to understanding practice variation that focus on systemic factors [4].

In some cases, physicians act as if they have a higher treatment or testing threshold than an EUT analysis would prescribe, implying that they are more sensitive to harms than benefits. Countering their over-perception of harms or heightening their awareness of benefits would reduce underuse. In other situations, physicians use a lower threshold than analysis would recommend, consistent with over-awareness of the benefits or neglect of potential harms of false positives. Clarifying the true benefits, or 'advertising' the harms, could reduce overuse. Note, however, because there are few tests causing immediately obvious harms, and because the regulatory agencies only approve treatments with benefits outweighing its harms, the thresholds – the ones based on the 'gold' standard rationality of EUT – are predictably low for most decisions in contemporary practice. As a result, we may expect continuing motivation for overuse, which would counter the current efforts to eliminate waste.

Our threshold hypothesis is empirically testable. Different threshold models based on EUT [14], regret-based [15], hybrid-based (combination of regret and EUT model) [16,17] or dual processing model [15] have been tested in small studies using vignette-based scenarios. Although limited evidence seems to suggest that non-EUT threshold models describe physician behaviour better, larger surveys, and in particular studies of the relation between the stated and/or computed thresholds and actual practice patterns, are needed to address our hypothesis more definitively. We predict that EUT threshold models will be supplemented if not supplanted by non-EUT account of threshold decision making.

Because we stipulate that we cannot rely on normative EUT as a rational tool to reduce underuse or curb the waste, overuse and unnecessary treatment and testing, we suggest that we should re-assess rationality by taking into account both formal principles of rationality and human intuitions about good decisions [8]. Stanovich [18,19] argues that we frequently deviate from the



**Figure 1** How much diagnostic certainty is needed before treating a patient? The graph illustrates regret theoretical approach why many physicians require higher level of diagnostic certainty and do not treat according to expected utility theory (EUT). When treatment benefits outweigh harms by 19 times, according to EUT we should administer treatment at the threshold of 5%. However, most physicians require higher level of diagnostic certainty (50%) to avoid regret of treating many healthy patients even though treating at this low (5%) threshold is associated with much higher life expectancy (LE) than acting on non-EUT (regret based) threshold of 50%. That is, acting at low (5%) of the threshold leads to treatment of many more healthy (shown in green) than diseased patients (shown in red). The opposite holds when our decisions are regret driven: we treat many more patients who have disease even though such a strategy is associated with lower LE. (The graph is based on data for treatment of smear-negative tuberculosis and assuming that 40-year-old patient with treated tuberculosis (TB) has LE of 38 years vs. the patient with untreated tuberculosis (TB) who has LE of 5 years. It is important to note that a different theoretical framework may generate different results).

rationality of EUT because our species engages in reflective processing, which takes contextualization, symbolic values and higher order preferences into account, none of which are typically built into the EUT models. Miles *et al.* [20] contend that decision making has to rely on the concepts of person-centred health care that insist on social and humanistic ideals of care for the sick. Taking these views into account can further illustrate how context and symbolic complexity may violate standard normative principles of rationality. For example, one way to curb over escalating health care costs is to make decisions about coverage for health services using cost-effectiveness analysis (CEA) [21], a poster child of EUT. CEA is typically conducted from a particular society

perspective. According to blank prescription of CEA, it would be irrational to cover services such as allogeneic transplant for a refugee or an undocumented immigrant. Yet, these services are often covered because symbolic utility plays a crucial role in human rationality: it would be counter to our humanistic values if we turn back on other people who are as every bit humans as we are [18,19]. When our policies clash with the way we uniquely experience this *person* who is in front of us, the group interests become less important [22]. After all, who is to say that one day we, or our dear ones, cannot find ourselves in the similar predicaments where we could only be saved by reasoning based on humanistic principles? Or, imagine the effect of contextual factors

on our choices, which are typically not used in standard EUT analyses. According to EUT, our choice should be context independent: if we prefer  $x$  over  $y$ , we should not choose  $y$  if the choice is presented as  $x$  versus  $y$  versus  $z$ . Stanovich [18] illustrates this situation with this compelling example. Imagine that you at a party where you saw one apple in a bowl. The apple is your favourite fruit, and you certainly prefer it over nothing. However, 'taking the last apple in the bowl when I am in public' [18] may not be socially acceptable, so you decide not help yourself with it. That is, you preferred 'nothing' ( $y$ ) over the apple ( $x$ ). A few minutes later, the host puts a pear ( $z$ ) in the bowl. You face the same choice again: apple ( $x$ ) versus nothing ( $y$ ) versus pear ( $z$ ). This time, however, you pick up your apple – the context has entirely affected your choice contrary to the standard normative theory [18]. This example is not of theoretical importance only; the context dramatically affect people's choices [23] and even the results of clinical research [23].

Paying attention to these uniquely human provenance of ideas requires re-definition of traditional rationality based solely on normative EUT, supplementing it with a descriptive acknowledgement of human reflective rationality, striving to integrate all aspects relevant to the assessment and weighing of the benefits and harms of medical practice – medical, humanistic and socio-economic – within a coherent reasoning system. No formula or algorithm can substitute for this multidimensional reflective rationality, which philosopher John Rawls [24] called 'reflective equilibrium/considered judgment', but we may improve decision making if we ask physicians (1) to explicitly state their *diagnostic or treatment threshold* and (2) to *reflect on their choice and consider the appropriateness of our intuitions and emotions* to our chosen decision threshold [18,19]. In time, this approach may result in further building the consensus on 'rational decision making' in medicine and, we suspect, would improve the current large and unsatisfactory variation in health care.

## Acknowledgement

Supported in part by DoD grant (no. W81 XWH 09-2-0175; Djulbegovic).

## References

- Wennberg, J. E. (2010) *Tracking Medicine. A Researcher's Quest to Understand Health Care*. New York: Oxford University Press.
- Djulbegovic, B. & Guyatt, G. H. (2014) Evidence-based practice is not synonymous with delivery of uniform health care. *JAMA: The Journal of the American Medical Association*, 312, 1293–1294.
- Wennberg, J. (2011) Time to tackle unwarranted variations in practice. *BMJ (Clinical Research Ed.)*, 342 (26), 687–690.
- Djulbegovic, B., Beckstead, J. & Nash, D. B. (2014) Human judgment and health care policy. *Population Health Management*, 17 (3), 139–140.
- Institute of Medicine (2013) *Variation in Health Care Spending: Target Decision Making, Not Geography*. Washington, DC: The National Academies Press.
- Pauker, S. G. & Kassirer, J. P. (1975) Therapeutic decision making: a cost benefit analysis. *The New England Journal of Medicine*, 293, 229–234.
- Pauker, S. G. & Kassirer, J. (1980) The threshold approach to clinical decision making. *The New England Journal of Medicine*, 302, 1109–1117.
- Djulbegovic, B., van den Ende, J., Hamm, R. M., *et al.* (2015) When is rational to order a diagnostic test, or prescribe treatment: the threshold model as an explanation of practice variation. *European Journal of Clinical Investigation*, 45 (5), 485–493.
- U.S. Preventive Services Task Force (2008) Screening for colorectal cancer: U.S. Preventive Services Task Force recommendation statement screening for colorectal cancer. *Annals of Internal Medicine*, 149 (9), 627–637.
- Gigerenzer, G., Hertwig, R. & Pachur, T. (eds) (2011) *Heuristics. The Foundation of Adaptive Behavior*. New York: Oxford University Press.
- Basinga, P., Moreira, J., Bisoffi, Z., Bisig, B. & Van den Ende, J. (2007) Why are clinicians reluctant to treat smear-negative tuberculosis? An inquiry about treatment thresholds in Rwanda. *Medical Decision Making: An International Journal of the Society for Medical Decision Making*, 27 (1), 53–60.
- Kopelman, R. I., Wong, J. B. & Pauker, S. G. (1999) A little math helps the medicine go down. *New England Journal of Medicine*, 341 (6), 435–439.
- Sreeramareddy, C., Rahman, M., Harsha Kumar, H., *et al.* (2014) Intuitive weights of harm for therapeutic decision making in smear-negative pulmonary tuberculosis: an interview study of physicians in India, Pakistan and Bangladesh. *BMC Medical Informatics and Decision Making*, 14 (1), 67.
- Felder, S. & Mayrhofer, T. (2014) Risk preferences: consequences for test and treatment thresholds and optimal cutoffs. *Medical Decision Making: An International Journal of the Society for Medical Decision Making*, 34 (1), 33–41.
- Djulbegovic, B., Elqayam, S., Reljic, T., *et al.* (2014) How do physicians decide to treat: an empirical evaluation of the threshold model. *BMC Medical Informatics and Decision Making*, 14 (1), 47.
- Moreira, J., Alarcon, F., Bisoffi, Z., *et al.* (2008) Tuberculous meningitis: does lowering the treatment threshold result in many more treated patients? *Tropical Medicine and International Health*, 13 (1), 68–75.
- Moreira, J., Bisig, B., Muwawenimana, P., *et al.* (2009) Weighing harm in therapeutic decisions of smear-negative pulmonary tuberculosis. *Medical Decision Making: An International Journal of the Society for Medical Decision Making*, 29 (3), 380–390.
- Stanovich, K. E. (2013) Why humans are (sometimes) less rational than other animals: cognitive complexity and the axioms of rational choice. *Thinking & Reasoning*, 19 (1), 1–26.
- Stanovich, K. E. (2011) *Rationality and the Reflective Mind*. Oxford: Oxford University Press.
- Miles, A., Asbridge, J. E. & Caballero, F. (2015) Towards a person-centered medical education: challenges and imperatives. *Educación Médica*, 16 (1), 25–33.
- Neumann, P. J., Cohen, J. T. & Weinstein, M. C. (2014) Updating cost-effectiveness – the curious resilience of the \$50,000-per-QALY threshold. *New England Journal of Medicine*, 371 (9), 796–797.
- Slovic, P. (2010) *The Feeling of Risk. New Perspectives on Risk Perception*. New York, NY: Eeatchscan.
- Ariely, D. (2008) *Predictably Irrational*. New York: HarperCollins Publishers.
- Rawls, J. (1999) *A Theory of Justice. Revised Edition*. Cambridge, MA: Harvard University Press.