MEDICINE AND SOCIETY

Debra Malina, Ph.D., *Editor*

From the Eyeball Test to the Algorithm — Quality of Life, Disability Status, and Clinical Decision Making in Surgery

Charles E. Binkley, M.D., Joel Michael Reynolds, Ph.D., and Andrew Shuman, M.D.

"Good surgeons know how to operate; better surgeons know when to operate. But only the wisest surgeons know when *not* to operate."

This adage rings true for all who decide when and whether to wield the scalpel. From early in medical training, the ability to make rapid decisions and snap judgments is inculcated in aspiring physicians, perhaps no more dramatically than in the field of surgery. Many seemingly objective decisions about whether to offer life-sustaining surgical interventions are actually rooted in subjectivity. Judgments about the relationship between quality of life (QoL) and various disabilities, whether those disabilities predate a surgical or medical intervention or result from it, are a case in point. Qualitative evidence concerning the relationship between QoL and a wide range of disabilities suggests that subjective judgments regarding other people's QoL are wrong more often than not[1,2] and that such judgments by medical practitioners in particular can be biased.[3,4]

To address this problem, it's helpful to understand the history and norms behind "the eyeball test" — intuitively sizing up a patient's physical appearance in order to estimate surgical risk — and the reasons why subjective judgments often fall prey to problematic ableist assumptions. Such understanding may allow us to create better tools for making high-stakes clinical decisions. Ensuring that physicians and health care organizations do not discriminate on the basis of disability requires careful consideration of the question of to whom surgery is offered and to whom it is denied.[5]

GAUGING QUALITY OF LIFE

QoL — or the presumed lack thereof — is commonly used as a justification in medical and surgical decision making, just as it's used in everything from the deliberations of a hospital ethics committee to those of the World Health Organization.[6] In many ways, an emphasis on QoL is understandable. With the rapid development of new life-sustaining technologies in the mid-20th century, it became possible to keep people "alive" in conditions that would have spelled death at any other point in human history. Neither the Hippocratic, nor the Maimonidean, nor any other professional oath helps a physician determine the point at which to cease interventions in light of the advances of modern medicine. The highly publicized Karen Ann Quinlan and Nancy Cruzan cases of the late 1970s and 1980s made it clear that supporting life regardless of its conditions ran afoul of the ethical principle of beneficence and perhaps other core ethical principles undergirding medical practice.[7]

That these two cases forced physicians to take into account a patient's QoL, rather than only life itself, in making treatment recommendations is significant. Just a few years before the Quinlan case, physicians had treated the severely burned patient Dax Cowart against his express and capacitated refusal of treatment.[8] Cowart's refusal was based in part on his judgment that if he survived, his QoL would be such that death was preferable. His physicians, on the other hand, believed it was their ethical duty to preserve his life regardless of his assessment of its quality. Cowart's case was one of the first to highlight the importance — ethically, not just medically — of letting the patient's own QoL valuation guide surgical decision making.

Just as physicians were beginning, in the 1980s, to consider QoL more explicitly in their treatment recommendations, the disability rights movement was building and disability studies programs were

being launched. Disability rights advocates argued that judgments about which conditions are compatible with a "livable life" — not to mention the higher standard of a good life — are not best made by medical experts alone.[9] Rather, they should be made in concert with people who have the relevant lived experience and in light of research rooted in such people's worldviews.

THE EYEBALL TEST

Guided by their desire to do good and avoid harm, surgeons have relied on the eyeball test to decide, from the vantage point of their own experience and intuition, whether a given patient will or will not benefit from surgery.[10] Although the intention of the eyeball test may be to predict the likelihood of major perioperative adverse events or death, the surgeon cannot help but also see the patient as a full person and make some assumptions about the patient's current QoL and about how the proposed operation might improve or diminish it. The eyeball test can easily harbor a range of implicit judgments and biases. For example, "frail elderly woman in a wheelchair" carries very different implications from "robust and fit well-dressed woman." Whether spelled out in the chart or merely mentally noted by the physician, such characterizations are powerful and will shape how an argument for or against surgery is crafted in the medical record. Subjective characterizations of functional status are even incorporated into performance scales — such as the Eastern Cooperative Oncology Group (ECOG) performance-status scale — used throughout oncology to gauge candidacy for cancer-directed interventions including surgery and chemotherapy.[11]

The eyeball test is thus problematic in several ways. First, it relies on one observer's assessment of the patient's candidacy for surgery based on the patient's appearance at a specific point in time. Because this sort of assessment is most often used to justify decisions not to undertake surgery, there may be little to no room for shared decision making with the patient and the family. In fact, the patient may never know that surgery was a possibility because the surgeon dismissed it as not indicated and never offered it. Though surgical decisions made in the settings of case conferences or tumor boards, with input from other clinicians, may be less subject to bias than a unilateral decision made by one surgeon, they are still vulnerable to ableist biases that may go unrecognized. And though the eyeball test is just one factor in surgical decision making, it too often plays an oversized role and one that is especially prone to bias, inconsistency, and lack of interrater reliability, as empirical data have confirmed.[12,13]

THE ABLEIST CONFLATION OF DISABILITY WITH PAIN AND SUFFERING

An additional challenge of such decisions, specifically in surgical oncology, is the intertwining and nuanced nature of function and survival, such that surgeons also "eyeball" how a proposed operation might affect a patient's future QoL and function. Instead of refusing to operate on a patient on the basis of the patient's already existing functional disability, the surgeon refuses to operate on the basis of the functional disability that is likely to result from the operation. In the first instance, the patient's QoL, in the surgeon's eyes, is so poor that surgery is not worth undertaking. In the second, the patient will be left with such a poor QoL after surgery that performing it is deemed unwise. Tellingly, in a recent survey of practicing physicians in the United States, 82.4% of 714 physicians indicated their belief that people with significant disability (as defined by the study) have worse QoL than people without disabilities.[3] Yet this judgment directly conflicts with a large body of social science research spanning decades suggesting that people with significant disability, like those with less severe disability, experience QoL that is similar to that of people without disabilities.[1,14-17]

Consider a case in which extensive craniofacial resection of a particularly aggressive (but anatomically resectable) skull-base tumor would result in the loss of one or both eyes. Laying the groundwork for a long history of wrongly conflating disability with pain and suffering,[18] Aristotle considered blindness a privation or *steresis*, loosely translated as "the violent taking away of something."[19] In step with Aristotle, there is an old surgical belief that extensive craniofacial resections should be questioned when they would require sacrifice of both orbits or an only-seeing eye.[20] Of course, this rule of thumb does not mean that blind patients with cancer should not have surgery, but rather that the functional consequence of a massive operation in a patient with locore-

gionally advanced cancer must be guided by prognosis in addition to other variables, as consensus statements and guidelines have established.[21,22] But we must recognize that such recommendations themselves are historically biased, and we can and should question the mantras of our professional indoctrination.

The belief that extensive craniofacial resections resulting in binocular or monocular vision loss should be avoided is not purely a consequence of concerns about resectability or oncologic prognosis. It derives from the assumption that when patients become functionally monocular or blind, their QoL will decrease. But a large body of social science research on the relationship between disability and QoL refutes this assumption.

Misunderstanding the complex relationship between disability and QoL has serious and far-reaching implications.[23] As Iezzoni et al. note, "examples include failures to provide Pap tests to women with disability or to discuss contraception options because of incorrectly assuming they are neither sexually active nor at risk for unintended pregnancy… [other examples include] physicians incorrectly believing that all patients with spinal cord injury cannot feel pain below the level of their injury and therefore [refusing] to provide pain relief for procedures below that level (for example, topical anesthetic during skin biopsy of the lower leg), thus causing these patients sometimes excruciating pain."[3] What's more, some research suggests a link between such ableist assumptions and the prevalence of medical error affecting patients with disabilities.[24,25]

There is a big difference between looking at a tumor and judging, "I can't get that out" and looking at a patient living with a disability and judging, "It won't be worth it to you." To be sure, there are many considerations involved in the decision making leading up to such a judgment — how one defines success, manages uncertainty, and stratifies perioperative risk, among others. But given the complexity of the issues and the importance of outcomes in such decision making, the eyeball test is simply not up to the job, and patients' own views need to be part of that calculation.

IMPROVING ON THE EYEBALL TEST

We are not advocating that surgical interventions should be offered indiscriminately. Rather, we believe that a patient's candidacy for a proposed treatment should be based on an objective assessment of the likely outcome of the treatment and the value that the patient, rather than the physician, places on that outcome, rather than on a flawed intuitive assessment. Medical practice should be guided not by how QoL is judged looking in from the outside, but by how it feels from the inside. There are two crucial sources for this inside view: how individual patients assess and value their QoL and how their assessment aligns with the best social science evidence available on the conditions, transitions, and outcomes at issue, in relation to relevant groups.

Distinguishing between a decision not to operate based on a surgeon's risk calculation, which is part of good surgical judgment, and a decision not to operate based on the surgeon's ableist assumptions about a patient's QoL requires a razor sharper than most surgeons' scalpels. It is therefore our professional duty to present the patient with information that is as accurate and objective as possible, and it is up to the patient to determine what value the outcome holds, how the proportionality of risk is gauged, and what QoL is acceptable or unacceptable.[10] One approach to making these determinations involves a best case–worst case framework designed to allow surgeons to reframe decisions more holistically rather than focusing purely on operative details.[26]

Surgical calculators have been introduced for predicting more accurately the risks and benefits of operations,[27] but these tools generally cannot be personalized for individual patients.[28] Using predictive health data and utilizing machine learning and artificial intelligence systems, it is possible to be more objective and accurate in predicting the likelihood of a desired outcome for a particular patient.[29,30] What is required is the creation of data sets drawn from patient populations that are diverse — in terms of not just race and gender, but also disability, functional status, and perceived QoL. It is also imperative that algorithms not be biased by assumptions about what a patient may or may not believe to be an acceptable QoL.[31] For example, training an algorithm using a data set from a fully sighted population that largely believes vision loss is an unacceptable surgical outcome would bake into the system the very biases that these systems have the potential to overcome.

As the surgeon–patient relationship continues to evolve and patients' values are more consistently prioritized in surgical decision making, it will be imperative that surgeons provide patients with the best and most personalized prediction of likely treatment outcomes. In light of the critical nature of this information, it must be based on more than one physician's subjective assessment.

Disclosure forms provided by the authors are available at NEJM.org.

From the Department of Bioethics, Hackensack Meridian Health, Edison, and the Department of Surgery, Hackensack Meridian School of Medicine, Nutley — both in NJ (C.E.B.), and Markkula Center for Applied Ethics at Santa Clara University, Santa Clara, CA (C.E.B.); the Department of Philosophy and the Kennedy Institute of Ethics, Georgetown University, Washington, DC (J.M.R.); the Hastings Center, Garrison, and the Greenwall Foundation — both in New York (J.M.R.); and the Center for Bioethics and Social Sciences in Medicine and the Department of Otolaryngology — Head and Neck Surgery, University of Michigan Medical School, Ann Arbor (A.S.).

1. Albrecht GL, Devlieger PJ. The disability paradox: high quality of life against all odds. Soc Sci Med 1999;48:977-88.
2. Campbell SM, Nyholm S, Walter JK. Disability and the goods of life. J Med Philos 2021;46:704-28.
3. Iezzoni LI, Rao SR, Ressalam J, et al. Physicians' perceptions of people with disability and their health care. Health Aff (Millwood) 2021;40:297-306.
4. Iezzoni LI, Rao SR, Ressalam J, et al. US physicians' knowledge about the Americans with Disabilities Act and accommodation of patients with disability. Health Aff (Millwood) 2022;41:96-104.
5. Keune JD. Disability and the contemporary surgical gestalt. Disabil Stud Q 2018;38.
6. World Health Organization. WHOQOL: measuring quality of life. March 1, 2012 (https://www.who.int/tools/whoqol).
7. Wolf SM, Berlinger N, Jennings B. Forty years of work on end-of-life care — from patients' rights to systemic reform. N Engl J Med 2015;372:678-82.
8. Cowart D, Burt R. Confronting death: who chooses, who controls? A dialogue between Dax Cowart and Robert Burt. Hastings Cent Rep 1998;28:14-24.
9. Evans JH. A critical history of bioethics. In: Reynolds JM, Wieseler C, eds. The disability bioethics reader. New York: Routledge, 2022:41-9.
10. Welke KF. The eyeball test: can the blind leading the blind see better than the statistician? Circ Cardiovasc Qual Outcomes 2014;7:11-2.
11. ECOG-ACRIN Cancer Research Group. ECOG Performance Status Scale. 2022 (https://ecog-acrin.org/resources/ecog-performance-status).
12. Yanagawa B, Graham MM, Afilalo J, Hassan A, Arora RC. Frailty as a risk predictor in cardiac surgery: beyond the eyeball test. J Thorac Cardiovasc Surg 2018;156(1):172-176.e2.
13. Jain R, Duval S, Adabag S. How accurate is the eyeball test? A comparison of physician's subjective assessment versus statistical methods in estimating mortality risk after cardiac surgery. Circ Cardiovasc Qual Outcomes 2014;7:151-6.
14. Edwards DJ, Sakellariou D, Anstey S. Barriers to, and facilitators of, access to cancer services and experiences of cancer care for adults with a physical disability: a mixed methods systematic review. Disabil Health J 2020;13:100844.
15. Iezzoni L, O'Day BL. More than ramps: a guide to improving health care quality and access for people with disabilities. New York: Oxford University Press, 2006.
16. Campbell SM, Stramondo JA. The complicated relationship of disability and well-being. Kennedy Inst Ethics J 2017;27:151-84.
17. Amundson R. Disability, ideology, and quality of life: a bias in biomedical ethics. In: Wasserman D, Bickenbach J, Wachbroit R, eds. Quality of life and human difference: genetic testing, health care, and disability. New York: Cambridge University Press, 2005:101-24.
18. Reynolds JM. The life worth living: disability, pain, and morality. Minneapolis: University of Minnesota Press, 2022.
19. Aristotle. Metaphysics. Ross WD, trans. Oxford: Clarendon Press, 1924.
20. Conley J. The risk to the orbit in head and neck cancer. Laryngoscope 1985;95:515-22.
21. Neel GS, Nagel TH, Hoxworth JM, Lal D. Management of orbital involvement in sinonasal and ventral skull base malignancies. Otolaryngol Clin North Am 2017;50:347-64.
22. Yousem DM, Gad K, Tufano RP. Resectability issues with head and neck cancer. AJNR Am J Neuroradiol 2006;27:2024-36.
23. Reynolds JM, Binkley CE, Shuman A. The complex relationship between disability discrimination and frailty scores. Am J Bioeth 2021;21:74-6.
24. Janz HL. Ableism: the undiagnosed malady afflicting medicine. CMAJ 2019;191(17):E478-E479.
25. Peña-Guzmán DM, Reynolds JM. The harm of ableism: medical error and epistemic injustice. Kennedy Inst Ethics J 2019;29:205-42.
26. Taylor LJ, Nabozny MJ, Steffens NM, et al. A framework to improve surgeon communication in high-stakes surgical decisions: best case/worst case. JAMA Surg 2017;152:531-8.
27. Clark DE, Fitzgerald TL, Dibbins AW. Procedure-based postoperative risk prediction using NSQIP data. J Surg Res 2018;221:322-7.
28. Hyde LZ, Valizadeh N, Al-Mazrou AM, Kiran RP. ACS-NSQIP risk calculator predicts cohort but not individual risk of complication following colorectal resection. Am J Surg 2019;218:131-5.
29. Balch J, Upchurch GR Jr, Bihorac A, Loftus TJ. Bridging the artificial intelligence valley of death in surgical decision-making. Surgery 2021;169:746-8.
30. Loftus TJ, Tighe PJ, Filiberto AC, et al. Artificial intelligence and surgical decision-making. JAMA Surg 2020;155:148-58.
31. Binkley CE, Kemp DS, Braud Scully B. Should we rely on AI to help avoid bias in patient selection for major surgery? AMA J Ethics 2022;24(8):E773-E780.

DOI: 10.1056/NEJMms2207408