MEDICAL PAPERS

Medical Papers
Introduction

BY JEFFREY R. BOTKIN

Dr. Sara Goldkind and I had the privilege of working with 13 extraordinary medical students from across the U.S. and one from Germany during the 2018 FASPE Medical Fellowship Program. This highly select group proved to be insightful, passionate about ethics, and fully engaged with their colleagues and the historical learning environments offered by FASPE. The focus of the experience is, of course, the study of professional ethics. For medical students, this includes the study of ethics in the practice of medicine, biomedical research, and public health.

The historical context of Nazi Germany challenges us to understand how it was that physicians who were dedicated to health and healing and to the science of medicine could become committed designers and participants in mass murder on a massive scale. The primary focus of our attempts at understanding this were the perpetrators, with the notion that without an understanding of our own ethical concepts, strengths, and vulnerabilities, we too could be swept up by misguided philosophies to commit small and large atrocities.

As instructors, our task was to prepare a curriculum that addressed contemporary ethical issues in medicine with roots or relevance to Nazi medicine. Using primarily a case-based approach, the FASPE Medical program's seminars dealt with issues such as the nature of a profession, the role of race and heritage in healthcare, the value of human life, the dual obligations of physicians in war, the allocation of scarce resources, the emotions that arise in patient care, speaking truth to power, and the ethical conduct of biomedical research. In interdisciplinary sessions with the Seminary Fellows, we explored ethical issues at the end of life and the enduring question of why evil exists.

Once home, the Medical Fellows were invited to submit a short essay on a topic of their choosing. Not surprisingly, given the extraordinary quality of this group of Fellows, we received thoughtful essays across a broad range of topics from a consideration of how to approach parents who are skeptical of vaccines to an exploration of the role of disgust to the ethics of naming diseases after physicians in

Germany who were implicated in the Holocaust. Here we present three of the papers we received.

Kelly Schuering's beautifully written piece addresses the role of ethical reflection and moral communities in the training of physicians, a topic directly relevant to the FASPE enterprise. Melissa Lavoie's thoughtful essay addresses the concept of futility when managing patients with severe anorexia nervosa. Is it ever appropriate to transition such patients into palliative care rather than persist with treatments and hopes of recovery? And Amelia Haj adeptly explores the myriad ethical issues that will arise as medicine comes to rely ever more heavily on artificial intelligence for the purposes of diagnosing disease.

On behalf of Sara Goldkind and myself, it was a privilege to explore the complicated and important issues with these talented young scholars and future leaders.

Jeffrey R. Botkin, *MD, MPH, is Professor of Pediatrics and Medical Ethics at the University of Utah School of Medicine. In 2018, he co-led the FASPE Medical Fellowship Program with* ***Sara Goldkind,*** *MD, MA, Bioethics Consultant at Goldkind Consulting, LLC and former Senior Bioethicist at the Food and Drug Administration.*

The Moral Community of Medicine and Its Role in Medical Training

BY KELLY SCHUERING

Physicians are given a privileged position in society to serve people at their most vulnerable, openly discuss patients' intimate problems, and perform acts that are otherwise considered taboo. However, with this privilege comes societal expectations that the physician subscribe to higher ideals of the greater good and be dedicated to something other than self-interest: namely the welfare of the patient.[1] Medicine is thus an inherently moral enterprise, which—as medical ethicists Edmund Pellegrino and David C. Thomasma write in their book, *The Virtues in Medical Practice*—combined with physicians' shared moral purpose of "healing, helping, and caring in a special kind of human relationship," binds physicians together into a moral community.[2] Because of the trust placed in physicians, society is rightly interested in the character of physicians as reflected in the belief that "only a good person can be a good physician."[3]

However, medicine and the moral community of physicians are not incorruptible, as revealed, among other instances, by the role that physicians played in euthanasia and killing during the Holocaust, the infamous Tuskegee experiment, and the financial conflicts of interests that exist today in the provision of care. In fact, the statement cited above on what makes a good physician was made by Rudolf Ramm, a leading Nazi ethicist, who through his influential medical ethics textbook spread the Nazi ideals of the authoritarian physician, the unequal value of humans, racial purity, and an overarching emphasis on a particular notion of public health.[4] The history of physicians abusing their power points to the need for a continued emphasis on ethics within the medical moral community and the need to develop not only medical competence but also strong morals in medical trainees in order to prepare them for

[1] Edmund D. Pellegrino and David C. Thomasma, *The Virtues in Medical Practice* (New York: Oxford University Press, 1993), 32.
[2] Pellegrino and Thomasma, 3 and 21.
[3] Robert Lifton, *The Nazi Doctors: Medical Killing and the Psychology of Genocide* (New York: Harper Collins, 1986), 17.
[4] Florian Bruns and Tessa Chelouche, "Lectures on Inhumanity: Teaching Medical Ethics in German Medical Schools Under Nazism," *Annals of Internal Medicine* 166 (8) (April 18, 2017): 591-96. doi:10.7326/m16-2758.

their privileged role in society. Only by embracing its role as a moral community can medicine more clearly define its norms to prevent the moral drift that led to atrocities in the past, best meet the needs of current patients, and ensure an ethical future for medicine.

Promoting a more ethical medical community first requires agreeing upon what norms and virtues medicine should espouse. Throughout history, physicians have sought to formalize the ethical commitments of the profession through credos like the Hippocratic Oath. More recently, groups including the American Medical Association, American College of Surgeons, American Board of Internal Medicine, American College of Physicians, and the European Federation of Internal Medicine have each come out with ethics statements of their own.[5,6,7] Generally agreed upon principles, including beneficence, non-maleficence, autonomy, and justice, predominate in discussions of biomedical ethics today.[8] Numerous and varied ethical credos promote the primacy of patient concerns, medical competence, the just distribution of resources, respect for and collegiality with other medical professionals, cultural sensitivity, and full disclosure of conflicts of interests. Nevertheless, such codes of conduct and physicians' claims to moral grounding are not in and of themselves self-justifying. As Pellegrino and Thomasma write, they must be grounded in "something more fundamental, on a philosophy of medicine."[9] Such a philosophy is still based upon general moral principles that are themselves subject to independent evaluation. Deciding upon the norms of medicine, therefore, also requires agreement upon acceptable moral standards.

As contemporary Western society has become more pluralistic, it has become increasingly challenging to agree upon one moral code of conduct within medicine due to competing values and beliefs. Throughout history, many have appealed to religious doctrine or human reason to ground their beliefs.[10] Today, appeals are often made to other values, such as legality or efficiency, or in the case of medicine, the benefit to

---

[5] American Medical Association, "AMA Code of Medical Ethics," accessed December 6, 2018, https://www.ama-assn.org/delivering-care/ama-code-medical-ethics.
[6] "Statements on Principles," April 12, 2016, *American College of Surgeons,* accessed December 06, 2018, https://www.facs.org/about-acs/statements/stonprin.
[7] American Board of Internal Medicine Foundation, American College of Physicians-American Society of Internal Medicine Foundation, European Foundation of Internal Medicine, "Medical Professionalism in the New Millennium: A Physician Charter," *Annals of Internal Medicine* 136 (3) (February 05, 2002): 243-46, doi:10.7326/0003-4819-136-3-200202050-00012.
[8] Warren A. Kinghorn et al., "Viewpoint: Professionalism in Modern Medicine: Does the Emperor Have Any Clothes?" *Academic Medicine* 82 (1) (January 01, 2007): 40-45, doi:10.1097/01.acm.0000249911.79915.4d.
[9] Pellegrino and Thomasma, 42.
[10] Lauris Christopher Kaldjian, *Practicing Medicine and Ethics: Integrating Wisdom, Conscience and Goals of Care* (New York: Cambridge University Press, 2014), 105.

the patient.[11] However, the validity of these principles is still subject to evaluation, which inevitably occurs within the context of one's background and community. As a group of academic physicians outline in a 2011 article on professionalism in medical education, "virtues have to be generated within communities that articulate and sustain them within their histories, institutions, and with practical wisdom."[12] Alasdair MacIntyre, a leading moral and political philosopher, agrees, arguing further that moral traditions and moral reflection are inherently situated in a particular culture and history.[13] The rise of new ethical approaches, such as feminist ethics, has challenged prevailing moral norms and led societies to question more broadly who should define morals and the moral life.[14] Due to their privileged position in society, physicians have retained the ability to define their own ethical norms within the broader societal boundaries of the law as a distinct moral community.

Although it is important that physicians acknowledge their privilege in being able to establish their own ethical professional norms, they remain best suited to do so. Ideally, physicians share the same commitment to caring for patients and are best positioned to hold each other accountable through open, pluralistic dialogue. A community's influence on how moral questions are framed and the moral conclusions drawn by the individuals within it are usually tempered by a person's level of identification with that moral community.[15] Physicians, in particular, usually strongly identify with the medical profession. Furthermore, the collegial, but specialized nature of medicine creates both a strong sense of accountability to other practitioners and a respect for the individual contributions and choices made by colleagues. This recognition of fellow physicians as moral agents is critical to their continued engagement with each other in working to define and enforce ethical norms. As peers working toward the common goal of the good of the patient, physicians are best positioned to question each other's moral reasoning and actions through open dialogue. These conversations can only be successful when the full range of perspectives held by physicians are encouraged and expressed, while the influence of moral communities outside of medicine are also acknowledged. Physicians must be open to dialogue and to engagement with different ideas in order to forestall the defensiveness that is all too often the starting point of conversations between those with opposing viewpoints today. As medical ethicist and physician Lauris

---

[11] Farr A. Curlin et al., "Religion, Conscience, and Controversial Clinical Practices," *New England Journal of Medicine* 356 (18) (May 03, 2007): 593-600, doi:10.1056/nejmc070628.
[12] Timothy P. Daaleman, et al., "Rethinking Professionalism in Medical Education Through Formation," *Family Medicine* 53 (5) (May 2011): 326.
[13] Alasdair C. Macintyre, *After Virtue: A Study in Moral Theory* (Notre Dame, IN: University of Notre Dame Press, 1981).
[14] Margaret Urban Walker, *Moral Understandings: A Feminist Study in Ethics* (New York: Oxford University Press, 2008).
[15] Kaldjian, 113.

Kaldjian writes in his book, *Practicing Medicine and Ethics: Integrating Wisdom, Conscience, and Goals of Care,* "We should expect persons who make conscience-based claims to describe how their moral beliefs and commitments are related to their moral positions and choices."[16] Shifting the norm for dealing with ethical conflict towards ongoing conversation will lead to greater discernment and ultimately "the courage to be decisive without the pretense of being definitive."[17] Because ethics is situational and thus greatly influenced by the circumstances surrounding a dilemma, a willingness to refrain from definitiveness further contributes to a culture where dialogue is valued and encouraged.

Through open dialogue, individual medical practitioners are not only further formed themselves, but also contribute to the continued development of the wider medical moral community. As Kaldjian notes, "as the norms of society influence the development of an individual's conscience, so also the words and actions of individuals sustain, challenge, or weaken the ethos of their community."[18] The tension and debate fostered by conflicting views can provide balance and prevent one idea or one individual's claims from having a tyrannical hold on the community:

> Moral communities provide socially sanctioned norms that are internalized within an individual's conscience, creating an equilibrium between the moral values of the community and the moral values of the individuals. This places checks on individuals against the risk of moral self-delusion, even as it also allows individuals to challenge and sustain the reigning values of the community.[19]

Continued dialogue that acknowledges the principles outside of medicine upon which medical ethical beliefs and judgements are grounded is necessary to further individual formation and to refine the norms and values of the medical moral community. Only through this continual refinement can medicine reach its ultimate goal of achieving the best possible outcomes for patients.

Fostering open dialogue about ethical norms and conflicts requires having ethically-formed physicians who are willing to speak out and challenge prevailing views. Despite the ongoing struggle to ground medical ethics education in one particular approach, the fact that ethical training is incorporated into medical school curricula demonstrates that the medical field is generally committed to the ethical formation of

---

[16] Kaldjian, 104.
[17] Larry R. Churchill, "Conscience and Moral Tyranny," *Perspectives in Biology and Medicine* 58 (4) (2015): 526-34, doi:10.1353/pbm.2015.0037 533.
[18] Kaldjian, 114.
[19] Kaldjian, 121.

future physicians.[20] Discussions of ethics can all to easily become stalled in semantics, and as new discoveries and new technology rapidly appear, along with their attendant ethical dilemmas, a return to virtue ethics may actually better serve society. Virtue ethics focuses on the character of a person performing an action rather than their duties or the consequences of their actions. As Pellegrino and Thomasma astutely comment, "the character of the physician is an irreducible factor in the healing relationship. How he or she interprets the moral principles, selects the values that will predominate, and shapes self-interest will be more important than how the moral principles are formulated or described."[21] The medical community and medical educators, in particular, must therefore recognize their role in character formation in addition to instruction in ethics.

One common objection to the need for ethical education among medical students is that individuals are already morally formed by the time they matriculate. This is undoubtedly true. Students are members of various familial, cultural, and/or religious communities that have shaped their individual views and values.[22] Medical school admissions committees screen for qualities that are valued by the medical community like altruism, justice, and honesty. However, moral formation is a lifelong process; additional virtues can be fostered, and individuals must learn to apply old virtues to new situations. Students, in particular, are still malleable, especially with respect to the traits that will serve them as future physicians. As Pellegrino and Thomasma note, "What we can hope to teach in medical school are those virtues appropriate to medicine—those that make for a good physician judged in terms of the telos of medicine, that is, a right and good healing action for a particular patient."[23] Despite coming to medicine with preexisting moral beliefs, medical students can and should develop morally throughout their training in order to be better equipped to manage the unique ethical challenges posed by medicine.

Trainee education must include formal instruction in medical ethics, for as Pellegrino and Thomasma say, ethics is ultimately "the only discipline that can place moral constraint on self-interest."[24] Formal ethical curricula can help students recognize ethical issues or dilemmas, reflect critically on their own values, defend their reasoning for making certain choices, respond to challenging viewpoints, and foster self-criticism. However, formal training is not sufficient. At the time of the Holocaust, Germany had the most developed ethical instruction in medical schools worldwide.

---

[20] Joseph A. Carrese et al., "The Essential Role of Medical Ethics Education in Achieving Professionalism," *Academic Medicine* 90 (6) (2015): 744-52, doi:10.1097/acm.0000000000000715.
[21] Pellegrino and Thomasma, 29.
[22] Kinghorn et al., "Viewpoint: Professionalism in Modern Medicine," 41.
[23] Pellegrino and Thomasma, 176.
[24] Pellegrino and Thomasma, 38.

Physicians with strong ties to the Nazi party taught the curricula which emphasized the physician as a "health leader" responsible for the betterment of society through his contribution to racial purification.[25] This history again points to the need for pluralism within the moral medical community because, as psychiatrist Warren Kinghorn and his co-researchers discuss in their paper on developing trainee professionalism:

> The risk of a student being influenced by particular moral traditions that did not undergird commonly held professional virtues such as altruism, justice, and honesty, would be mitigated by the ability of particular traditions to come into conversation with one another and by the deep inculcation in the virtues that many, if not most, other students would experience.[26]

Though necessary for laying a foundation from which to approach ethical challenges, a formal ethics curriculum is far from sufficient to instill moral values in future physicians as evidenced by the Nazi regime.

Medical education should recognize and embrace the roles of narrative, emotion, humility, focused reflection, and maintaining engagement with outside moral communities in furthering the ethical formation of medical trainees. The humanities and narratives, in particular, highlight the interweaving of motivations, actions, and consequences.[27] Narratives help to refine one's ethics, for as ontology philosopher Annemarie Mol argues in her book *The Logic of Care: Health and the Problem of Patient Choice*, "conflicting stories tend to enrich each other. And while adding up arguments leads to a conclusion, adding on stories is more likely to be a way of raising ever more questions."[28] Through the use of narrative, individuals can come to better appreciate the shades of gray that color many ethical dilemmas today. Medicine must also recognize that emotions have a role to play in ethical formation. As psychologist Sidney Callahan argues in her chapter in *Medicine and the Ethics of Care*, "persons cannot engage in effective moral deliberation without activating their subjective responses of emotion, intuition, and imagination, all of which operate within the implicit assumptions and narratives of their communities."[29] Beyond acknowledging emotions within the medical moral community, physicians can foster the emotion of

---

[25] Bruns and Chelouche, "Lectures on Inhumanity," 593.
[26] Kinghorn et al., "Viewpoint: Professionalism in Modern Medicine," 44.
[27] Kaldjian, 111.
[28] Annemarie Mol, *The Logic of Care: Health and the Problem of Patient Choice* (London: Routledge, 2011), 88.
[29] Sidney Callahan, "The Psychology of Emotion and the Ethics of Care," in *Medicine and the Ethics of Care,* edited by Diana Fritz Cates and Paul Lauritzen (Washington D.C.: Georgetown University Press, 2002), 144.

care. Caring increases moral functioning by helping the physician to gather more information about people, the environment, and the relationships between the two to better inform ethical decisions.[30] However, engaging with various narratives and emotions is difficult without humility and an openness to change. Approaching ethical dilemmas with humility primes one, as Kaldjian says, to "learn from the experiences of others, acknowledge prejudices, and modify judgments in the light of more information or better insights."[31] If medicine, and in particular medical educators, were to more highly value narrative, emotion, and humility, it would help trainees to explore a wider range of viewpoints and new ideas that inevitably contribute to moral formation.

Encouraging trainees to reflect on their values and experiences can also further ethical formation. For many students, managing increasing patient loads and staying up to date with rapidly expanding medical knowledge limits time for reflection. Only with reflection, are students able to process the cognitive dissonance they may feel in seeing the discrepancies between the idealized medicine they have studied and the day-to-day realities of the wards. Without the chance to re-examine their values and the reasoning underlying their beliefs, people often change their beliefs in response to cognitive dissonance if they feel limited in the ability to change their behaviors.[32] Medical schools can further students' ethical formation by setting aside "dedicated time for trainees to debrief emotionally challenging situations, such as the death of a patient, and for reflection on positive and negative experiences and the implicit messages that contribute to the formation of professional identity," write several academic physicians in a recent position paper by the American College of Physicians on the medical learning environment and professionalism.[33] By building this time into the curriculum, medical educators can instill in their students a sense of the importance of reflection, which will continue to benefit them as they progress in their careers as physicians.

Finally, maintaining meaningful relationships with moral communities beyond the medical community is important in helping medical trainees gauge how they are changing over the course of their studies and for providing a meaningful outlet for processing and reflection. Inhabiting roles outside of medicine also increases empathy; students can recognize parts of themselves and the people they care about in their

---

[30] Callahan, "The Psychology of Emotion and the Ethics of Care," 153.
[31] Kaldjian, 112.
[32] Christopher R. Browning, "Revisiting the Holocaust Perpetrators. Why Did They Kill," The Raul Hilberg Memorial Lecture, University of Vermont, Burlington, VT, October 17, 2011.
[33] Lisa Soleymani Lehmann, Lois Snyder Sulmasy, and Sanjay Desai, "Hidden Curricula, Ethics, and Professionalism: Optimizing Clinical Learning Environments in Becoming and Being a Physician: A Position Paper of the American College of Physicians," *Annals of Internal Medicine* 168 (7) (April 3, 2018): 506-508, doi:10.7326/m17-2058, Appendix.

patients and thus be reminded of their common humanity. By providing additional points of view, these communities outside of medicine contribute to the pluralism that is so critical for continued ethical formation. These communities can also help students develop the critical virtues of temperance and practical wisdom. As Pellegrino writes in his article "Professionalism, Profession, and the Virtues of the Good Physician," "The virtue-based physician is also compelled on the basis of virtues of parent and husband to be faithful to his commitments to his family, friends, and society. He would, however, recognize what the limits of legitimate self-interest are and when that set of interests should be set aside in the interests of his patient or vice versa."[34] Having multiple roles to fill also allows a physician to find meaning in life outside medicine and may thereby reduce the sense of risk involved in speaking out about ethical considerations. Relationships beyond medicine remind physicians of the ultimate aim of medicine and of their shared humanity with patients.

In an opinion piece for *Academic Medicine,* physician Warren Kinghorn and his coauthors write that in order to develop a meaningful professionalism, ethics education should:

> … embody an open pluralism, giving voice to diverse moral communities, encouraging critical self-exploration and discussion about the truth claims of these communities, and, if possible, facilitating the integration of students' professional development with their ongoing participation in these communities.[35]

Even when starting from an ideal, open environment, it can be a struggle for students to integrate their beliefs and actions. Only by approaching moral formation in multiple ways, including through encouraging open dialogue that values narrative, emotion, and humility, focused reflection, and participation in moral communities outside medicine, can medicine hope to achieve moral formation in trainees that actively impacts their day-to-day practice.

However, all of these efforts to improve moral formation will still have a limited impact if students do not have the opportunity to see their mentors actively apply ethical principles and if they have no chance to practice applying what they have learned themselves. Trainees are often more formed by what they actually observe while practicing medicine in clinics and on the wards than by what they are taught in the classroom. Students' desires to perform well and be liked make them particularly susceptible to being influenced by their superiors. Studies like the Milgram

---

[34] Edmund D. Pellegrino, "Professionalism, Profession, and the Virtues of the Good Physician," *The Mount Sinai Journal of Medicine* 69 (6) (November 2002): 378-84, 382.
[35] Kinghorn et al., "Viewpoint: Professionalism in Modern Medicine," 40.

experiments have shown that conformity is a powerful force, especially within a hierarchical system.[36] Because students seek to identify with their superiors, they are particularly susceptible to following unethical instructions without pausing for further evaluation. Even when students know that what they are observing or are asked to do is morally wrong, "under the pressure to conform in order to succeed and not to imperil a career, even a virtuous student will need extra courage to resist," note Pellegrino and Thomasma.[37] Especially in environments in which individuals know their performance is being assessed, human beings will tend to judge themselves by group norms and how others view them rather than against their own individual principles and standards. Basing one's self-esteem on external evaluation has been shown to lead to a realignment of morals.[38] Furthermore, many students have limited prior clinical experience and may assume whatever they see modeled is the norm. In addition to witnessing various clinical approaches, students observe the words used, policies, practices, and institutional structures in clinical settings, and thus medical faculty members inevitably influence the ethical and moral development of their students by example alone. All too often students observe contradictions between what they have been taught about ethics and professionalism and what they observe in clinical practice. The negative effects of this "hidden curriculum" was recognized as early as 1994, when one study reported that 62 percent of medical students felt some of their ethical principles had eroded during the course of their training.[39] Recognizing and correcting the negative hidden curriculum observed by impressionable medical trainees is thus a critical step to forming ethical physicians.

Few clinical faculty members explicitly acknowledge being responsible for their students' ethical formation. According to an American College of Physicians (ACP) position paper, in order to promote greater professionalism, "what is taught in the classroom must be reinforced and enhanced by what is practiced at the bedside," especially as, "the intensity of medical training is a cultural immersion in which values are often communicated and adopted without adequate reflection and critique."[40] For example, the emphasis on efficiency and dangerous levels of pride in professional skill and knowledge often conflicts with the primacy of what is best for

---

[36] Stephen D. Reicher, S. Alexander Haslam, and Joanne R. Smith, "Working Toward the Experimenter: Reconceptualizing Obedience Within the Milgram Paradigm as Identification-Based Followership," *Perspectives on Psychological Science* 7 (4) (June 29, 2012): 315-25, doi:10.1177/1745691612448482.
[37] Pellegrino and Thomasma, 177.
[38] Browning, "Revisiting the Holocaust Perpetrators. Why Did They Kill."
[39] Chris Feudtner, Dimitri A. Christakis, and Nicholas A. Christakis, "Do Clinical Clerks Suffer Ethical Erosion? Students' Perceptions of Their Ethical Environment and Personal Development," *Academic Medicine* 69 (8) (1994), doi:10.1097/00001888-199408000-00017.
[40] Lehmann, Sulmasy, and Desai, "Hidden Curricula, Ethics, and Professionalism," 507.

the patient.[41] How students see these conflicts resolved affects their sense of medical norms and inevitably affects their moral formation, especially when time for reflection is not embedded into their training. Alternately, witnessing strong professionalism and ethical behavior can be just as influential. The ACP paper continues, "If they participate in a culture of compassion, curiosity, respect, and empathy, they are more likely to adopt these virtues, especially when they are identified as explicit values and expectations."[42] Physicians must take responsibility for modeling professionalism to better align the hidden curriculum with formal instruction in order to strengthen the medical moral community and foster the ethical development of trainees.

Ethical formation occurs within a community that provides guidelines and feedback; the medical moral community is no exception. Daaleman and his co-authors describe the ideal formation of a physician as the "ongoing integration of an individual who grows in self-awareness, with a group of companions who share both their interior and outwardly lived experiences as they participate in the common mission of the community."[43] This statement succinctly expresses the reflection, action, and shared mission that must all coexist to define and advance sustainable ethics within a community. The medical community needs to more fully embrace its role as a moral community in order to best serve patients and train the next generation of ethical providers. Such efforts would strengthen not only trainees, but current physicians as well. Daalemen, et al. write:

> Formation—medical learners, faculty mentors discovering who they are becoming as they move along together in lives of service within the environment of the medical home—offers a way to awaken, enrich, and sustain the virtues of both emerging and established physicians, and their capacities of caring, for the long haul.[44]

This formation is critical to all in the moral medical community in order to ensure that medicine is able to confront current and future ethical challenges. As Pellegrino and Thomasma note:

> Internalization of the right and the good through training and disposition will not only ensure application of the meaning of moral rules to life circumstances but will also lead to refinements of the moral principles and even to new moral theories that will try to resolve the new issues of the day.[45]

---

[41] Bruno Bettelheim, Afterword to *Auschwitz: A Doctor's Eyewitness Account* by Robert Jay Lifton (New York, NY: Arcade Publishing, 1960).
[42] Lehmann, Sulmasy, and Desai, "Hidden Curricula, Ethics, and Professionalism," Appendix.
[43] Daaleman et al., "Rethinking Professionalism in Medical Education Through Formation," 327.
[44] Daaleman et al., "Rethinking Professionalism in Medical Education Through Formation," 328.
[45] Pellegrino and Thomasma, 29.

Both trainees and current physicians can contribute to and learn from a stronger moral community within medicine for the good of current and future patients.

Only by banding together and embracing their role as a moral community can physicians and medical trainees more clearly define their ethical norms, advance professionalism in the day-to-day practice of medicine to better serve patients, and prepare leaders in health care to confront new ethical issues that will inevitably arise in the future. As Pellegrino and Thomasma conclude, physicians must:

> …for the first time in medical history establish themselves as a true moral community, as a group of persons dedicated to something other than their own self-interest, as a group that recognizes its responsibility to support the ethical members of its company, to repel or reject those who are not faithful to the ethical bonds that unite the community, and to advocate the cause of the sick, even when society and politics militate against it. These duties flow not only from the characteristics of the healing community, … but also from the qualities of professional commitment.[46]

This is medicine's main hope to resist the small compromises that subtly and incrementally have led to gross human rights violations in the past. Physicians and medical trainees must respond to the faith that society has placed in them by embracing a stronger emphasis on ethics to best care for patients both now and in the future.

Kelly Schuering *is a student at Vanderbilt University School of Medicine. She will graduate with a medical degree and a certificate in biomedical ethics in May of 2019. In June, she will start a residency in primary care and population health at Brigham and Women's Hospital.*

[46] Pellegrino and Thomasma, 39.

Palliative Care and the Concept of Futility in Severe Anorexia Nervosa

BY MELISSA LAVOIE

Anorexia nervosa has the highest mortality rate of any mental illness.[1] A diagnosis of anorexia requires several features: restricted energy intake leading to low body weight, poor insight or perceptions surrounding body weight or shape, and either a fear of fatness or behavior that interferes with weight gain.[2] Malnutrition in patients with extreme anorexia affects nearly every organ system, causing fractures, liver failure, and cardiac arrhythmias. Severely malnourished patients with anorexia require emergency treatment to prevent sudden death from cardiac complications and to normalize life-threatening electrolyte disturbances. Individuals with anorexia are nearly six times more likely to die than their peers without the disorder.[3]

Case reports have described patients with severe, unremitting anorexia nervosa who were withdrawn from involuntary treatment and transitioned to palliative care.[4,5,6] Often appealing to medical notions of futility, proponents of this approach have argued that the prognosis is poor for patients with severe anorexia who have already undergone multiple treatment attempts. They have also asserted that some patients with severe anorexia nervosa have the capacity to decide whether to receive life-sustaining care, and that involuntary treatment may violate autonomy.
One report describes a 30-year-old woman, identified as Ms. A.[7] Ms. A presented to an eating disorder clinic at a height of 5'4", a weight of 64 pounds and a body mass index

[1] Frédérique Smink, Daphne Van Hoeken, and Hans W. Hoek, "Epidemiology of Eating Disorders: Incidence, Prevalence, and Mortality Rates," *Current Psychiatry Reports* 14 (4) (2012): 406-14.
[2] American Psychiatric Association, *Diagnostic and Statistical Manual of Mental Disorders* (Arlington, VA: American Psychiatric Publishing, 2013).
[3] Smink, Van Hoeken, and Hoek, "Epidemiology of Eating Disorders."
[4] Amy Lopez, Joel Yager, and Robert Feinstein, "Medical Futility and Psychiatry: Palliative Care and Hospice Care as a Last Resort in the Treatment of Refractory Anorexia Nervosa," *International Journal of Eating Disorders* 43 (4) (2010): 372-377.
[5] Robin Mackenzie, "Ms. X: A Promising New View of Anorexia Nervosa, Futility, and End-of-Life Decisions in a Very Recent English Case," *American Journal of Bioethics* 15 (7) (2015): 57-58.
[6] Joseph O'Neill, Tony Crowther, and Gwyneth Sampson, "A Case Study: Anorexia Nervosa. Palliative Care of Terminal Psychiatric Disease," *American Journal of Hospice and Palliative Medicine* 11 (6) (1994): 36-38.
[7] Lopez, Yager, and Feinstein, "Medical Futility and Psychiatry."

(BMI) of 10.9.[8] She was diagnosed with anorexia nervosa at age 19 and had undergone multiple prior unsuccessful treatment attempts. She was admitted to the hospital from the clinic, and after medical stabilization, she refused to continue treatment. Her team concluded that her disorder was unlikely to respond to further treatment. Rather than pursue involuntary treatment, the team, in consultation with the hospital ethics committee, decided to transition her to a palliative care approach. She was told that she would not be asked to participate in any treatment she did not want, and she was discharged from the hospital at a weight of 85 pounds and a BMI of 14.6. At home, Ms. A received weekly visits from a palliative care nurse. She died several months later in inpatient hospice due to complications from her anorexia.

Had Ms. A not entered palliative care and instead received treatment in an inpatient eating disorder program, she would have likely been placed on a weight gain protocol. As part of that protocol, programs typically require patients to eat a prescribed amount of food at meals. If patients fail to receive adequate calorie intake through meals, they may eventually receive nutrition through a nasogastric tube.

I argue here that the concept of futility, which is used in medicine to assess the potential benefit of treatment, is ethically unsubstantiated as applied to even the most severely ill patients with anorexia. Physicians have a moral obligation to pursue involuntary treatment for patients with life-threatening anorexia nervosa who refuse care. I first examine the concept of futility in general. I then describe several challenges with applying the concept of futility to anorexia nervosa. Lastly, I discuss the risks of extending the notion of futility even to specific cases where the treatment team might feel the patient has the capacity to refuse treatment and the prognosis is poor. Here I also refer to statements made by Nazi physicians and even Adolf Hitler himself, which serve as sobering reminders of the danger of deeming psychiatric treatment futile for particular patients. Nazi physicians claimed to be able to determine with complete certainty which mentally ill patients were "incurable," and these claims masked a deep-seated disgust towards patients whose lives they saw as worthless.

The medical literature has coalesced around three broad categories of medical futility: physiological, quantitative, and qualitative.[9] Physiologically futile interventions have no chance of being effective. The use of antibiotics (which can only treat bacterial infections) for viral infections is an example of physiological futility. Quantitative futility applies to cases where the *likelihood* that an intervention will benefit a

[8] A BMI between 18.5 and 25 is considered healthy, and a BMI below 15 is considered extreme anorexia, the most severe subtype of the disorder.
[9] Thaddeus Pope, "Futility," in *Guidance for Healthcare Ethics Committees,* edited by D. Micah Hester and Toby Schonfeld (Cambridge, UK: Cambridge University Press, 2012): 88-97.

patient is exceedingly poor, whereas qualitative futility applies to those situations where the *quality* of the benefit the patient will receive from a particular intervention is exceedingly poor.

Of these categories, bioethicist and psychiatrist Cynthia Geppert argues that only qualitative futility could conceivably apply to severe anorexia nervosa.[10] Malnourishment is not physically irreversible and even severely starved patients remain physiologically able to respond to refeeding. Nor does quantitative futility apply. Geppert refers to a 2012 study which compared 41 anorexia patients with the most severe levels of malnutrition and an average BMI of 10.1 to 443 less malnourished patients with less severe anorexia nervosa.[11] The authors found that compared to the less malnourished patients, seven percent of the patients with severe malnutrition died over a six-year period (versus 1.2 percent in the less severe group) and 41 percent recovered (versus 62 percent in the less severe group). While the prognosis for patients with the most severe form of anorexia is grimmer than for those with more mild manifestations of the disorder, a disease that over two fifths of patients recover from within six years falls far short of any standard that has been used in medical care for quantitative futility.

Qualitative futility, on the other hand, applies when treatment is unlikely to provide an adequate quality of life or the harms of treatment outweigh the benefits. Some physicians have argued that qualitative futility may apply to certain cases of anorexia where the patient feels her quality of life is so poor that continuing life-sustaining treatment is not worthwhile.[12]

Anorexia is marked by poor insight and judgment relating to weight, body image, and eating. It therefore becomes ethically problematic to assume or conclude that patients suffering from the illness retain intact decision-making capabilities with regards to anorexia treatment. One of the criteria given for anorexia in *the Diagnostic and Statistical Manual of Mental Disorders 5* (DSM-5) (the standard manual used by mental health professionals to diagnose and assess mental illness) entails "disturbance in the way one's body weight or shape is experienced, undue influence of body shape and weight on self-evaluation, or persistent lack of recognition of the seriousness of the current low body weight."[13] Patients with anorexia may be

---

[10] Cynthia Geppert, "Futility in Chronic Anorexia Nervosa: A Concept Whose Time Has Not Yet Come," *American Journal of Bioethics* 15 (7) (2015): 34-43.
[11] Daniel Rigaud, Isabelle Tallonneau, Marie-Claude Brindisi, and Bruno Verges, "Prognosis in 41 Severely Malnourished Anorexia Nervosa Patients," *Clinical Nutrition* 31 (5) (2012): 693-698.
[12] Joel Yager, "The Futility of Arguing About Medical Futility in Anorexia Nervosa: The Question Is How Would You Handle Highly Specific Circumstances?" *American Journal of Bioethics* 15 (7) (2015): 47-50.
[13] American Psychiatric Association, *Diagnostic and Statistical Manual of Mental Disorders.*

convinced they are overweight when in fact they are severely underweight. Moreover, they may not appreciate the serious medical complications associated with low body weight. While patients with anorexia may retain the capacity to make other types of medical decisions, their insight and judgment surrounding weight is often so impaired that it renders them incapable of adequately assessing the risks and benefits of eating disorder treatment.

Furthermore, research has suggested that malnutrition itself inhibits decision-making capacity and worsens symptoms of anorexia. The Minnesota Starvation Study from 1945-46 remains one of the most detailed investigations into the impact of malnutrition on psychological functioning. As the previously healthy participants in the program lost up to a quarter of their body weight on a calorie-restricted diet, they exhibited behavior patterns often seen in anorexia. They became preoccupied with food, engaged in complex food rituals, and withdrew socially.[14] They also developed misperceptions relating to body size, including overestimating the width of their own faces.[15] Importantly, these abnormalities tended to resolve with weight normalization.

The Minnesota Starvation Study illustrates that starvation impairs thinking, and that for severely malnourished patients weight restoration is a critical step in improving insight. While allowing patients to pursue a palliative care approach may appear to be more consistent with the value of autonomy, paradoxically, refeeding enables patients to regain their decision-making capacity.

Some of the patients described in the literature who received palliative care for chronic anorexia nervosa do not appear to meet the criteria for the capacity to refuse eating disorder treatment. Ms. A, the patient who received palliative care for her eating disorder, refused to sign a "Do Not Resuscitate" order because she did not believe she would die. Formal assessments of medical decision-making capacity typically examine the patient's understanding of her condition and the risks and benefits of various treatment options. Given that Ms. A did not believe her eating disorder would lead to her death, she likely did not appreciate the severity of her condition and the harms associated with treatment refusal. This raises questions about her capacity to refuse care.

What appears to be treatment unresponsiveness may also represent poor access to high-quality, evidence-based care. Many levels of care and approaches to anorexia treatment exist, and numerous programs utilize approaches supported by little or no

[14] Leah Kalm and Richard Semba, "They Starved So That Others Be Better Fed: Remembering Ancel Keys and the Minnesota Experiment," *The Journal of Nutrition* 135 (6) (2005): 1347-1352.
[15] Leah Kalm and Richard Semba,"They Starved So That Others Be Better Fed."

evidence.[16] So, while a patient who has not recovered after numerous treatment programs may seem treatment-resistant, she might also simply have yet to receive the most appropriate care. This is different from fields such as oncology, where large clinical trials have evaluated the efficacy of treatments, and the protocols for chemotherapy are more similar across institutions. A woman with a history of anorexia writes in *The American Journal of Bioethics*:

> In my case it was more than 12 years before I began to recover, and I experienced numerous failed interventions … including behavior modification (a reward/punishment system), Freudian, group, individual, occupational, and cognitive behavioral therapy. The successful treatment program was far more comprehensive than any I had previously experienced."[17]

With such a wide range of treatment approaches in use and variable access to quality care, it is unlikely that any patient has exhausted all the options available.

Given the severity of impaired thinking in patients with anorexia, it is important not to accept as fact statements from patients about the hopelessness of treatment. Patients with anorexia often claim that their prognosis is poor and that future treatment would be useless. A study of physician aid in dying in the Netherlands found that a patient's feeling of hopelessness was one of the most important criteria affecting a physician's perception of the appropriateness of euthanasia.[18] Christopher Williams, a researcher at the University of Leeds, writes in *The BMJ,* "When low weight levels are reached, staff may accept at face value statements that previous treatments have been ineffective, or assume that no other treatments will be effective in the future, and therefore cease active treatment for a distressed patient who is not improving."[19] It is crucial for clinicians to understand that a patient's expression of hopelessness may not be rooted in clinical reality.

Joel Yager, a psychiatrist and eating disorder specialist at the University of Colorado, has suggested that there are certain cases in which palliative care is appropriate for

---

[16] Erica Goode, "Centers to Treat Eating Disorders Are Growing, and Raising Concerns," *The New York Times,* March 14, 2016, https://www.nytimes.com/2016/03/15/health/eating-disorders-anorexia-bulimia-treatment-centers.html.
[17] Cushla Mckinney, "Is Resistance (N)ever Futile? A Response to 'Futility in Chronic Anorexia Nervosa: A Concept Whose Time Has Not yet Come' by Cynthia Geppert," *American Journal of Bioethics* 15 (7) (2015): 53-54.
[18] Alan Ogilvie and SG Potts, "Assisted Suicide for Depression: The Slippery Slope in Action?" *The BMJ* 309(6953) (1994): 492-493.
[19] Christopher Williams, "Does Palliative Care Have a Role in Treatment of Anorexia Nervosa? We Should Strive to Keep Patients Alive," *The BMJ* 317 (7152) (1998): 195-196.

patients with anorexia. He writes that he remains "highly suspect of arguments that (even if everyone could agree on a definition) considerations of 'futility' might never ever be suitable for a given patient."[20] Yager describes severely malnourished patients who expressed an enduring wish to discontinue treatment and consistently asserted that continued treatment was certain to do more harm than good. They had failed to recover after years of therapy and numerous inpatient admissions for anorexia. Yager asserts that these patients were able to separate their beliefs about quality of life and the costs and benefits of treatment from thoughts induced by the eating disorder itself.

I argue above that the impaired insight and judgment caused by anorexia, the wide range of approaches to anorexia treatment and poor access to care, and the challenges of distinguishing treatment-unresponsiveness from patient hopelessness make it unethical to discontinue life-sustaining care in most if not all eating disorder patients. However, even if a patient meets Yager's description of someone with an enduring opposition to treatment and a clear understanding of its risks and benefits, making these decisions remains morally problematic.

There is no evidence, expert consensus, or clinical guideline that would enable psychiatrists to determine whether or not a patient is exceedingly unlikely to recover from severe anorexia.[21] This makes seemingly objective prognostic assessments vulnerable to value judgments about the types of lives that are worth living. These kinds of value judgments played a central role in claims by Nazi physicians that certain conditions were incurable.

Alfred Hoche and Karl Binding's 1920 book *The Permission to Destroy Life Unworthy of Life* illustrates the tendency of Nazi physicians to invoke absolute prognostic certainty when sending patients to be murdered. Hoche, a Nazi psychiatrist, described a concept of "mental death" in which individuals with mental illness were considered "human ballast."[22] About these patients he wrote, "the physician has no doubt about the hundred-percent certainty of correct selection" and "proven scientific criteria" to establish the "*impossibility of improvement* of a mentally dead person."[23] He expressed a completely unsubstantiated claim about the physician's ability to assess a patient's prognosis with regards to mental illness, a claim that would be unjustified even today. Entangled with this "medical hubris," as psychiatrist and

[20] Joel Yager, "The Futility of Arguing about Medical Futility in Anorexia Nervosa."
[21] Cynthia Geppert, "Futility in Chronic Anorexia Nervosa."
[22] Robert Lifton, *The Nazi Doctors: Medical Killing and the Psychology of Genocide* (New York: Harper Collins, 1986), 47.
[23] Lifton, 47.

author Robert Jay Lifton describes it, was the Nazi biomedical worldview that killing mentally ill individuals was necessary for the health of the *Volk*.[24]

Hitler invoked Hoche's belief in the "impossibility of improvement" in his "Führer Decree" of 1939, in which he wrote: "Reich Leader Bouhler and Dr. Brandt are charged with the responsibility for expanding the authority of physicians … to the end that patients considered incurable according to the best available human judgment of their state of health, can be granted a mercy death."[25] While not expressing the same degree of medical hubris as Hoche, Hitler asserted that some patients are incurable. Yet, as with Hoche, this belief is inseparable from his vision for strengthening the *Volk*. Both men made disingenuous claims of medical authority to promote eugenic policies while masking their disgust towards those whose lives they considered worthless.

As Hoche and Hitler's words demonstrate, the concept of qualitative futility carries with it an intrinsic danger: assertions of futility may easily stem from judgments about which lives have value. In the case of psychiatric illness, this risk looms particularly large. Qualitative futility requires consideration of a patient's quality of life, an inherently subjective task that is difficult to separate from value judgments. Patients with mental illness face continued stigma and marginalization. That there is no evidence base that could inform a conclusion of qualitative futility in anorexia makes these conclusions even more morally problematic. While anorexia is a life-threatening condition, most patients survive and many recover. Even those who do not recover still lead lives worth living. Physicians caring for patients with severe anorexia should proceed from that standpoint and engage their patients in life-saving treatment with the goal of eventual recovery.

Melissa Lavoie *is a student at Johns Hopkins University School of Medicine. She will graduate with a medical degree in May of 2019 and will start a residency in psychiatry at Johns Hopkins in June.*

[24] Lifton, 47.
[25] Lifton, 63.

Artificial Intelligence and Physician Conscience

BY AMELIA HAJ

In viewing the crimes of Nazi-era physicians, it is easy to find ourselves passing judgment. It is difficult to imagine any but the most depraved doctors willingly participating in murdering the disabled, experimenting on and torturing innocents, and supporting a regime that worked tirelessly to exterminate entire swaths of the population. No authority in this day and age, we think, could possibly compel us to set aside our morals so effortlessly. In the decades since the 1930s and 1940s, the practice of medicine has changed dramatically: advances in medical knowledge and in standards of patient-physician interaction have both empowered patients to be active participants in their care and have also created expectations of near perfection in physicians' diagnostic and treatment abilities. In our ongoing effort to further improve our standards of care, artificial intelligence (AI) software is being developed that can in some cases outperform physicians, and a future in which doctors' actions are guided by computerized algorithms no longer seems fantastical. Here, I argue that modern day physicians face a far more insidious, well-intentioned authoritative threat than did their Nazi-era counterparts: the gradual incorporation of AI into medical diagnostics and decision-making, a shift that will require physicians to carefully examine the role of their own consciences in their daily work to a greater degree than they already do.

Artificial Intelligence and Medicine

At its most basic, a medical diagnosis carried out by a physician could be said to rely on algorithms of a sort: a patient's symptom(s) or complaint triggers a series of questions, each leading the clinician down a path to a diagnosis and subsequent treatment. Years of study go into developing an understanding of these algorithms and learning when to apply them. Then, with experience, physicians can go on to gain a deeper understanding of the population they treat and can begin to account for less tangible factors, such as minor, seemingly irrelevant details in a patient's history, subtle changes in body language, or an unusual blip on a scan. It is these intangibles, this "gut sense" attained after seeing hundreds of cases, that make physicians more

than simply mouthpieces for medical textbooks. Anyone with enough coding skills to write a series of if/else statements could develop a tool that would roughly approximate the stepwise process used by a clinician to determine the cause of a symptom such as chest pain, for instance. Only in recent years, however, has software engineering become sophisticated enough to begin to dynamically adapt to new information much as the human brain would.

Many definitions of artificial intelligence (AI) exist, but all converge around the use of software that can mimic human decision-making. Simpler forms of AI usually rely on convoluted decision trees, but machine learning, a more sophisticated subtype of AI, incorporates the ability to learn by recognizing patterns in datasets. Deep learning is a subcategory of machine learning that trains artificial neural networks on large datasets, so that these networks then emerge with the ability to look at and interpret documents and images, much as humans can learn to recognize patterns without a conscious awareness of having learned discrete facts. Incredibly, and somewhat troublingly, a neural network cannot provide clear step-by-step rationales for its decisions. It may be able to accurately identify photos of chickens, say, and you could tease apart the factors it weighed to determine what it considered most important in identifying a particular bird as a chicken, but it could not provide you with a protocol for improving your own ability to distinguish chickens from parakeets.

Artificial intelligence is already seamlessly integrated into our lives. Individualized Netflix suggestions, tailored Google search results, and optimized driving routes on Google Maps are all driven by AI. Smart assistants such as Siri and Alexa use AI to understand our verbalized demands, and dating sites such as Hinge are now exploring the use of machine learning to improve their suggestions. Far from being a futuristic prospect, our actions and beliefs are already guided and shaped by decisions made by computer algorithms and neural networks.

Researchers have, of course, already begun to train machine learning software on medical datasets. A group of dermatologists published a letter in the journal *Nature* in 2017 describing their use of a deep neural network to classify photographs of skin lesions as either benign or malignant.[1] On average, the neural network outperformed dermatologists, suggesting that the technology has the capacity to very quickly and accurately reach levels of mastery in skills that humans take years to learn to a lesser degree of accuracy. Some have gone so far as to assert that certain tasks, especially those requiring image analysis, will soon be taken over entirely by deep learning

[1] Andre Esteva, et al., "Dermatologist-Level Classification of Skin Cancer with Deep Neural Networks," *Nature* 542 (7639) (2017): 115–18, doi:10.1038/nature21056.

algorithms, as physician and author Siddhartha Mukherjee describes in his 2017 piece on the subject for *The New Yorker.*[2]

There are many other avenues for artificial intelligence to become integrated into medical practice. The broad movement towards using electronic health records (EHRs) creates an obvious opening for the use of natural language processing software, which can "understand" human-written language, to analyze and form recommendations based on patterns in patients' written medical records. Indeed, researchers have already explored the use of deep neural networks for earlier prediction of diseases and for predicting events relevant to hospital performance metrics, such as hospital readmissions. Some have suggested using AI technologies to assist patients with dementia, and others make the case for developing AI tools that can approximate the decisions incapacitated patients would have made had they still possessed all their faculties in order to ease the difficulty of relying on surrogate decision makers. Given the integration of AI into other areas of our lives and the advances in its use in medical contexts, it is reasonable to expect that we will increasingly see its use for medical purposes in the coming years.

Ethical Concerns

Many of the ethical considerations surrounding the use of AI in medicine center on the largely "black box" nature of machine learning software. Without knowing how decisions are being made, we run the risk that AI tools will unwittingly perpetuate and amplify human biases based on the datasets that are entered by humans into the software. One of the most oft-cited examples of this is the use of risk-assessment tools by criminal justice systems to predict the likelihood that a person convicted of a crime will reoffend. In an investigation conducted by the news site ProPublica, these algorithms were nearly twice as likely to incorrectly predict that black defendants would reoffend than white defendants.[3] Remarkably, the algorithms did not use direct data about the races of the individual defendants, but rather relied on defendants' answers to a series of questions, demonstrating that bias can still emerge even in scenarios that appear to be deliberately neutral. It is not difficult to imagine machine learning software intended to predict the best individualized treatment plan for a particular condition inadvertently exacerbating healthcare disparities along racial or health literacy lines.

---

[2] Siddhartha Mukherjee, "A.I. Versus M.D.," *The New Yorker*, June 19, 2017, www.newyorker.com/magazine/2017/04/03/ai-versus-md.
[3] Julia Angwin, et al., "Machine Bias," *ProPublica*, May 23, 2016, www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing.

Sometimes, algorithms can simply misinterpret the information they are given, with potentially devastating consequences. In one striking example, a machine learning tool was trained to predict the risk of a patient dying of pneumonia and offer a recommendation as to whether the patient should be treated as an inpatient or whether he/she could safely be treated as an outpatient.[4] Paradoxically, the tool determined that pneumonia patients who also suffered from asthma had a considerably lower risk of death and recommended that they be treated as outpatients. In reality, such patients are generally at a much higher risk of death and more consistently receive ICU levels of care, and as a result ultimately have better outcomes. This led the algorithm to the erroneous (and circular) conclusion that because this subset of pneumonia patients generally had good outcomes, they could be treated safely as outpatients. The researchers in this study noted that because they used an "intelligible" machine learning tool to predict patient outcomes, they were able to identify the software's reasoning; had the researchers used a neural network, it would have likely come to the same conclusions, but in a more opaque fashion, making it much more difficult for the researchers to identify the source of the error.

The competing motivations of AI software creators, buyers, and users are also an important consideration.[5] Even though we generally trust them to have their patients' best interests in mind, physicians are not, of course, flawlessly unbiased actors; employers' expectations of profits and physicians' fears of malpractice lawsuits can often lead to unnecessary testing and overtreatment of conditions. In theory, AI could be taught to neutralize motivations that threaten to compromise patient care, but in a profit-driven healthcare system, it is reasonable to assume that the software employed may reflect motives beyond simply improved patient care.

Conscience and Authority in Medical Decision-Making

A future in which AI software governs a physician's every move is not on the horizon quite yet. The idea of such an extreme, however, offers an opportunity to explore how physicians should respond in situations in which their authority is supplanted by a supposedly superior, yet still fallible, authority. To develop a framework for how physicians can navigate this hypothetical new world, we must first establish how to determine whether an authority is legitimate and what this means in a medical

[4] Rich Caruana, et al., "Intelligible Models for HealthCare: Predicting Pneumonia Risk and Hospital 30-Day Readmission," *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining - KDD '15,* Sydney, Australia, August 10-13, 2015, doi:10.1145/2783258.2788613.
[5] Danton S. Char, et al., "Implementing Machine Learning in Health Care—Addressing Ethical Challenges," *New England Journal of Medicine* 378 (11) (2018): 981–983, doi:10.1056/NEJMp1714229.

context and then identify how an individual physician's conscience operates in normal medical settings.

The concept of authority is typically defined within political and legal contexts, but it is relevant to attempt to apply these definitions to medicine, which is notoriously hierarchical. The word "authority" can be used in a theoretical sense to describe someone who is an expert, and it can be used in a practical sense to describe a person or entity that has the right and power to guide and control a group of people. There are different views on how to determine whether a practical authority is legitimate, but for the purposes of this discussion, I will define a legitimate authority as one that is "right, justified, [and] supported by good reasons"[6] and that carries some requirement that it be obeyed, even though obedience may be inconsistently enforced.

In an ideal 21st-century patient-doctor relationship, the physician acts as a theoretical authority, but the practical authority is distributed across the patient, the physician, and a set of collectively-held expectations that the standard of care for the patient's situation will be available and offered. No longer is the physician the paternalistic authority figure from a bygone era. Today, physicians are obligated to make available the standard of care, as determined by the broader medical community, while still ensuring that the patient is an active participant in his or her care decisions. The legitimacy of the physician's and healthcare system's authority is granted by ongoing trust from the patient that both the physician and the system are committed to making decisions in the best interest of each patient; conflicting interests, such as the desire to increase profits, deliver a blow to the credibility of the physician and healthcare system.

When a physician perceives that his or her professional and moral obligations diverge, the role of the physician's conscience comes into play. Conscience is a nuanced concept that can be defined in a multitude of ways, but for our purposes here, I will use one of the definitions offered by the *Stanford Encyclopedia of Philosophy*: conscience can be considered an internal "sense of duty" that compels us to "act according to moral principles or beliefs we already possess."[7] An individual's conscience does not always correspond to perceived objective morality and does not need to originate in any particular source; it can arise from a religious background, one's own moral code, or an intuitive sense of right and wrong. A simple example of a conscience-triggering conflict would be a case in which a patient requires a lifesaving abortion, but the

[6] Michael Lacewing, "Authority and Legitimacy," (London: Routledge, 2018), cw.routledge.com/textbooks/alevelphilosophy/data/AS/WhyShouldIBeGoverned/Authorityandlegitimacy.pdf.
[7] "Conscience," *Stanford Encyclopedia of Philosophy,* last modified December 2, 2016, plato.stanford.edu/entries/conscience/#ConsMotiActMora.

physician treating the patient believes that abortion is murder and does not want to be complicit by performing or assisting with the abortion.[8] In this scenario, the physician's conscience serves as a check of sorts on the actions other practical authorities (the patient and medical establishment) are expecting him or her to perform. Essentially, the exercise of one's conscience can provide a means of overriding an authority in situations where a physician perceives the authority's judgment to be in error. A survey of physicians found that 42 percent believed that physicians should never be expected to do something that conflicts with their conscience, reflecting how important it is to many physicians to maintain their individual moral integrity.[9] The American Medical Association recognizes that doctors are human beings who cannot be expected to divorce their personal beliefs from their actions at work. It therefore offers guidelines for how doctors can maintain their personal moral integrity while also ensuring that their patients receive the appropriate standard of care.[10] If a physician's conscience can be said to act as a check on the overreach of other authorities, then the recommendations put forth by the AMA close the loop by acting as a check on the unfettered exercise of a physician's conscience—which could itself lead to overreach and diminish the quality of patient care—and thereby also ensure that the physician's authority remains legitimate.

Physician Conscience and Artificial Intelligence

Subjects in Stanley Milgram's well-known electric shock experiments demonstrated the human propensity for following authority without question, in part offering a possible explanation for why physicians in Nazi Germany were able to seemingly unquestioningly commit atrocities. In both situations, the actions being committed—"euthanizing" disabled children, on the one hand, for example, and delivering (fake) fatal doses of electricity, on the other—were objectively bad, despite the fact that those ordering them often presented such actions as means to positive ends. Taken to its logical conclusion, a healthcare system driven by AI would be one in which a new source of authority is introduced—potentially to the exclusion of the current three sources of authority outlined above, since, as we have seen, deep neural networks have the capacity to outperform humans at complex tasks. In this scenario, the authority would likely be intended as a genuine improvement upon a healthcare system that relies on the efforts of well-meaning yet error-prone humans. Yet the

[8] Martha S. Swartz, "'Conscience Clauses' or 'Unconscionable Clauses': Personal Beliefs Versus Professional Responsibilities," *Yale Journal of Health Policy, Law, and Ethics* 6 (2) (2006): 269–350.
[9] Ryan Lawrence et al., "Physicians' Beliefs About Conscience in Medicine: A National Survey," *Academic Medicine* 84 (9) (2009): 1276–82.
[10] American Medical Association, "Code of Medical Ethics Opinion 1.1.7," *Ethics: Physician Exercise of Conscience,* accessed September 28, 2018, www.ama-assn.org/delivering-care/physician-exercise-conscience.

potential pitfalls seem nearly as varied as the humans the algorithms could one day replace, and the potential for error just as grave or worse.

The growing use of AI in medicine will undoubtedly create situations in which a physician disagrees with a computer's decision. Perhaps a computer algorithm deems a woman's very severe abdominal pain to be insufficiently suggestive of ovarian torsion and recommends outpatient treatment for constipation. Perhaps it decides that a white man's MELD (Model for End-Stage Liver Disease) score of 34 is more deserving of a liver transplant than a black man's MELD score of 35. Perhaps it performs a behind-the-scenes cost-effectiveness analysis and determines that a suicidal teenager's inpatient psychiatric care should not be covered by insurance. In such situations, the appeal to a physician's conscience may be expanded from an expression of personal morality to a tool for patient advocacy. The potential risks to the physician in speaking up, however, remain, making such situations especially challenging. In an era of defensive medicine, it may feel safer to defer to the algorithms—it may seem preferable, for instance, that a malpractice lawsuit land in the lap of the software manufacturer who failed to correctly fine-tune the algorithms to distinguish between constipation and ovarian torsion than in the lap of a physician who chose to override the computer's suggestion to give the patient Miralax with a decision to perform emergency surgery.

A careful approach to implementing AI in medicine will require an ongoing dialogue among physicians, healthcare administrators, and software developers regarding the legitimacy of AI's authority. Is it "right, justified, [and] supported by good reasons"? Will we hold onto the expectation that physicians deliver a certain standard of care, or will that standard be relaxed if the algorithm says it should be? Will practical authority one day reside only in AI, or will AI simply have a seat at the table, along with the other sources of medical authority? Certainly, we will need to be mindful of historical precedent and create an environment where conscientious human decision-making is explicitly allowed and encouraged. A road paved with good intentions will not lead to hell if it is engineered with an understanding of the potential risks along the way and traveled on with caution.

Amelia Haj *is a student at the University of Wisconsin School of Medicine and Public Health. She will graduate with an MD-PhD degree in 2021.*