

JOURNAL OF HOSPITAL ETHICS

THE JOHN J. LYNCH, MD CENTER FOR ETHICS

ROUNDING WITH THE EDITOR

From the Birth of Contemporary Bioethics to the Ethics of Medical AI

Evan G. DeRenzo, PhD

FEATURES

Adverse Events in AI-Assisted Pediatric Surgery: Risk Mitigation, Responsibility, Disclosure, and Surgeon Support

Niall M. Jones, MD, FRCSI, FRSM, MCLinLead and Katrina A. Bramstedt, PhD, MA

AI Errors and Human Complacency: The Output Can Be Tragic

Timothé Ménard, PharmD, MSc, MA and Katrina A. Bramstedt, PhD, MA

BOOK REVIEW

Returning to Old Haunts: Complex Ethics Consultations 17 Years later

Bryan Pilkington, PhD



MedStar Health

JOURNAL OF HOSPITAL ETHICS

THE JOHN J. LYNCH, MD CENTER FOR ETHICS

EDITORIAL GROUP

EDITOR-IN-CHIEF

Evan G. DeRenzo, PhD
Ethicist & Editor, Emeritus
John J. Lynch, MD Center for Ethics

MANAGING EDITOR

Christian Carrozzo, PhD (c)
Educator & Editor
John J. Lynch, MD Center for Ethics
Department of Philosophy, George Mason University
Department of Philosophy, University at Albany
State University of New York

MEDICAL EDITOR

Jack A. Sava, MD
Chief, Trauma
Director, Gold Surgery Team
MedStar Washington Hospital Center
Associate Professor, School of Medicine,
Georgetown University Medical Center

NURSING EDITOR

Lucia Wocial, PhD
Assistant Director,
John J. Lynch, MD Center for Ethics

BOOK REVIEW EDITOR

Patrick Herron, DBe
Assistant Director, Bioethics Graduate Education
Montefiore Einstein Center for Bioethics
Associate Professor,
Albert Einstein College of Medicine

CONTENT EDITORS

Fred King, MSLS
Medical Librarian
MedStar Washington

S. Layla Heimlich, MSLIS
Medical Librarian
MedStar Washington

ADMINISTRATIVE ASSISTANT

Nikki Glover
Administrative Coordinator
John J. Lynch, MD Center for Ethics

EXECUTIVE DIRECTOR

Ben Krohmal, JD
Director, John J. Lynch, MD Center for Ethics
Associate Professor, School of Medicine,
Georgetown University Medical Center

EDITORIAL ADVISORY BOARD

Amanda Anderson, PhD, MPA, RN

Postdoctoral Fellow, United States Department of Veterans Affairs

Amanda Bush, MA, BSN, RN

Regional Nurse Wellbeing Specialist, Center for Wellbeing, MedStar Health, and Adjunct Faculty, Berkley School of Nursing, Georgetown University

Chee M. Chan, MD, MPH

Medical Director, Intermediate Care Unit, MedStar Washington Hospital Center

Maria Susana Ciruzzi, JD, MSc, PhD

Clinical Ethicist, The St. Jude Bioethics Program, St. Jude Children's Hospital

Olubukunola (O. Mary) Dwyer, JD, MBE

System Director, Clinical Ethics, Ohio Health

Laura K. Guidry-Grimes, PhD

Associate Bioethicist, Center for Bioethics, Cleveland Clinic, and Clinical Assistant Professor of Medicine, Lerner College of Medicine, Case Western Reserve University

Josh Hyatt, DSH, MBE, MHL

Director, Risk Education Management and Strategy, and Adjunct Professor, Massachusetts College of Pharmacy and Health Sciences

Jack Kilcullen, MD, JD, MPH

Physician, Medical Critical Care Services, INOVA Health

Eran Klein, MD, PhD

Associate Professor of Neurology, Oregon Health and Sciences University, and Affiliate Assistant Professor, Department of Philosophy, University of Washington

Tim Lahey, MD, MMSc

Director of Medical Ethics, Infectious Disease Attending, and Professor of Medicine, University of Vermont Medical Center

Jason Lesandrini, PhD

Assistant Vice President of Ethics, Advanced Care Planning and Spiritual Health, Wellstar Health System

Stephen Peterson, MD

Physician, Department of Psychiatry, MedStar Washington Hospital Center

Jack Schwartz, JD

Former Chief Council for Opinions and Advice, Maryland Attorney General's Office

Lindsay Semler, DNP, RN

Executive Director of Clinical Ethics, Brigham and Women's Hospital

Eric A. Singer, MD

Assistant Professor, Section of Urologic Cancer, Rutgers Cancer Institute of New Jersey, Robert Wood Johnson Medical School

Carol Taylor, RN, MSN, PhD

Senior Research Scholar, Kennedy Institute of Ethics, and Professor, Berkley School of Nursing, Georgetown University

Eli Weber, PhD, MA

Bioethics Director, Kaiser Permanente, San Bernardino, California



MedStar Health

© Copyright 2026 John J. Lynch, MD Center for Ethics, MedStar Washington Hospital Center. ISSN (Print) 1938-4955, ISSN (Online) 1938-4920. All Rights Reserved.

JOURNAL OF HOSPITAL ETHICS

THE JOHN J. LYNCH, MD CENTER FOR ETHICS

CONTENTS

CONTRIBUTORS 1

ROUNDING WITH THE EDITOR

From the Birth of Contemporary Bioethics to the Ethics of Medical AI
Evan G. DeRenzo, PhD 2

FEATURES

Adverse Events in AI-Assisted Pediatric Surgery: Risk Mitigation, Responsibility, Disclosure, and Surgeon Support
Niall M. Jones, MD, FRCSI, FRSM, MClinLead and Katrina A. Bramstedt, PhD, MA 5

AI Errors and Human Complacency: The Output Can Be Tragic
Timothé Ménard, PharmD, MSc, MA and Katrina A. Bramstedt, PhD, MA 11

BOOK REVIEW

Returning to Old Haunts: Complex Ethics Consultations 17 Years Later
Bryan Pilkington, PhD 13



MedStar Health

CONTRIBUTORS

Katrina A. Bramstedt, PhD, MA is a bioethicist who has worked across industry, academia, and clinical care. Formerly on staff at the Cleveland Clinic, she is currently Adjunct Professor at the Queensland University of Technology School of Medicine (Australia) and Global Head of Bioethics at Roche (based in Switzerland). She is the author of 125+ articles on various topics in bioethics, most prominently pertaining to clinical ethics consultation, organ donation and transplantation, and research ethics and integrity. Her personal websites are AskTheEthicist.com and TransplantEthics.com. She is currently pursuing a masters degree in biotechnology and law. txbioethics@yahoo.com

Niall M. Jones, MD, FRCSI, FRSM, MClinLead is a Pediatric Surgeon with 30 years clinical experience in teaching hospitals stretching from the west of Ireland to Hobart, Tasmania. A keen advocate for improving surgical outcomes and patient safety he is currently Adjunct Professor of Clinical Surgery at Khalifa University, Abu Dhabi. He combines clinical insight with operational and legal acumen to drive impactful healthcare strategies thereby contributing to system-wide improvement. He has just completed a Diploma in Project Management and is currently half-way through an LLM in Medical Law and Ethics with the University of London. He is excited by the emergence of AI in Surgery and is thrilled to collaborate with the field of Bioethics to explore the associated benefits and opportunities for patients and organizations. niallmj71@gmail.com

Timothé Ménard, PharmD, MA, MSC is the Global Head of Quality Excellence Digital at Roche, where they lead advanced analytics and automation initiatives for clinical trials and pharmacovigilance. With over 17 years of international experience across Switzerland, Singapore, and Germany, He also serves as a research ethicist supporting Global Bioethics at Roche. In this capacity, they provide subject matter expertise on genetics and data/AI ethics, specifically leading the implementation of "AI ethics design." As a co-founder of the IMPALA Consortium, they contribute to the industry-wide adoption of AI for R&D quality. Timothé Ménard holds an MSc in pharmaceutical medicine, an MA in ethics, and a PgDip in cancer genetics. Their work focuses on the critical intersection of digital innovation and ethical integrity in drug development. timothe.menard@roche.com

Bryan Pilkington, PhD, is (Full) Professor in the Department of Medical Sciences, at the Hackensack Meridian School of Medicine (Nutley, USA), where he teaches ethics. He is author of *The Medical Act: Conscientious Practice in a World of Dissent and Disagreement* (Springer, 2025) and over 40 peer-reviewed publications. His research in bioethics focuses on conscience, moral responsibility, and the practices of the health professions. His work in moral and political philosophy focuses on human dignity and moral disagreement. bryan.pilkington@hmsom.edu

ROUNDING WITH THE EDITOR

From the Birth of Contemporary Bioethics to the Ethics of Medical AI

Evan G. DeRenzo, PhD

Dear Readers,

Welcome to the first issue of our 12th volume. This is a special issue for us because it is also the 10th anniversary of John “Jack” Lynch MD’s death. It is a poignant time for me. Whenever I have been quoted about what it means to me to for Jack to die, what I believe I’ve always said was, “Everything I know about clinical ethics I learned from Jack.”

When I got to know Jack, I was still at the National Institutes of Health (NIH) and I met Jack when he was spearheading a collaborative effort with the DC Hospital Association to create a DC-wide clinical ethics community. (This was a successful effort of a few years until the grant money ran out). I really got to know Jack when I bumped into him at the NIH Clinical Center only to find him bereft at the impending death of his beloved first wife. At her funeral, I’ll never forget the sight of Jack being escorted down the length of a huge Catholic DC church, an adult child on each side, holding him by his arms. In the many years I worked with Jack, that was the only time I ever saw him cry. Fortunately, years later he remarried and died as a loving and happily remarried man.

So many of us at the hospital were bereft when Jack died. A common sentiment at the time was that we had lost an icon. So many commented on Jack’s passion, his commitment to excellence in oncology, and more broadly, medical ethics. Jack didn’t need to present the fact that he was committed to medical ethics, he just exuded it. He was the perfect moral model any hospital could have ever had. He manifested unwavering integrity for over 50 years at MedStar Washington Hospital Center.

But Jack was no saint. He could be wicked funny with the straightest face. Jack could also be incendiary. For instance, I recall there being a physician who I knew Jack didn’t very much respect, although Jack himself would rarely mention how he felt. One day, however, Jack came into the Center for Ethics office, there were only a couple of us there, and announced, “I just want you to know that I’m going into a board meeting (a MedStar corporate board meeting) and I’m going to drop a bombshell. If I’m fired, so be it.”

“Well ok Jack” I thought, keeping my mouth shut for a change. “You go get ‘em.” Fortunately for everyone, even though he did drop a bombshell, he didn’t get fired. I can’t imagine Jack doing anything that could have ever gotten him fired. No matter how upset he might be, for over 50 years, Jack was the epitome of incorruptibility, despite being a very practically driven guy. You couldn’t help but respect him as perhaps the most ethically wise, medically compassionate, honorable, decent, and inclusive clinician in the hospital.

The kind of intensity with which Jack announced the dropping of that bombshell was only matched by Jack’s humility. A founding and intermittent chair of the hospital’s ethics committee, and intermittent Director of the Center for Ethics, both critical ethics bodies he created, Jack’s work ethic outmatched everyone’s.

Jack died in the early morning hours of a Monday. The Friday before, Jack called for an ethics subcommittee meeting within the Center. In those days that meant the Center director and me. He called for it because he wanted feedback on how he was handling a particular case. I found this amazing because in the 20 years I worked with him, before and after I arrived at the Hospital Center, I never knew him to make an error in judgement. He always intuitively understood what the right, or at least ethically best, thing to do was. His moral intuition was beyond reproach.

Given that this issue of JoHE, a publication to which Jack contributed and supported in so many ways, coincides with the 10th anniversary of his death, we dedicate this issue to him. With that, we now turn to the advent and integration of AI into medicine, especially surgery, a phenomenon that Jack might have found terrifying.

AI technology is everywhere. Some tout AI as the greatest boon to personkind. Others only see the perils AI may pose, such as taking over jobs that humans now perform and depend on for their own and their families’ livelihoods. Some have even expressed the concern that AI, in “thinking” faster than humans, will someday conquer the human race, ultimately turning humans into AI servants.

In the context of this issue, concerns land somewhere in the middle. That is at least what our

two feature articles, generated from similar viewpoints, appear to suggest. That is if, prior to reading these pieces, one looks at the authors' bios in our contributors' section, the reader will see that these quite thoughtful, technology-conservative (and separately submitted) papers were each independently generated from within a technology-developing company.

After having connected these two pieces intellectually, in my view their shared normativity lies within the warning to make certain to accept moral responsibility, i.e., don't pawn responsibility off onto the AI tool. Also, and in a related sense, to always assure that there is a responsible human in-the-loop. The other ethical norms of safety, expressed as safety-in-design, moral courage, transparency and accountability are ethical norms standardly attached to any medical and/or surgical situation. And no matter how we have attempted to democratize medicine, it remains that at the end of the day, it is the senior physician who is and must remain the senior responsible party.

To my mind and in my experience, what is considered democratization and teamwork might just as well be called, "treating all team members with common decency," has produced improvements in team politeness, nonetheless that behavioral progress does not change the fact that senior physicians must remain those responsible for patients when all is said and done. The ethical principle most specific to the hospital situation is having a human (or physician) in-the-loop. To imagine hospital care, especially pediatric surgery that involves an AI tool, without an alert, attentive, and supervising physician is truly inconceivable.

As each of these papers suggest, a fully present human/supervising clinician in-the-loop is likely able to address safety needs best. And in terms of medical error transparency, including the cascading implications for all the many difficult conversations and reports that may emanate from a medical error, is likely best modeled by a senior physician who was the human in-that-loop.

Another throughline between each of these features, to my mind, is that one of the most prominent reasons for having a person in-the-loop is to combat complacency. Safety and risk mitigation requires sharp observation of a patient's condition and attentiveness to risk of harm in the patient's environment. Complacency is perhaps the most terrifying and realistic risk of harm when AI is being used in the operating room.

As we become increasingly accustomed to AI in medical decision-making and performance, the likely risk of complacency rises. For those accus-

tomed to the organized chaos of a busy hospital, the growing dependence on AI-assisted surgeries, and perhaps just about anything else in the hospital, can likely produce complacency in oversight. The faster the AI assistant works and the greater frequency of successful outcomes, the more likely it is that a dangerous complacency will occur. This possibility could lead the way to potentially catastrophic errors in both clinical care, surgery, and research.

The similar warnings raised by each of these papers should give everyone in medicine pause. The ethical principles these papers raise, even more than perhaps other ethical frameworks that might be applied, provide critically important guardrails for the application of AI-assisted medical care, surgery, and medical research.

As noted clearly within this issue, the authors come out of the Roche Scientific Ethics Advisory Group, for which Katrina Bramstedt is the chair. From the information provided and easily available on the Web, the employed Roche authors are salaried by Roche and hold Roche stock. Now, rather than go on some rant about pharmaceutical industry employees and the potential conflicts of interests that ethics consultants who work for them may have, I'm going to be transparent and explain that in the early 2000s I served in a seemingly similar position as Dr. Bramstedt with Johnson and Johnson as an ethics consultant to their neuropsychiatry research division. Ultimately, I wrote a book chapter about what I saw as industry obligations and the implications of conflict of interests for ethics consultants.¹

Rather than suggest banning ethicists from consulting for industry, or suggesting that ethicists not work within an industry setting, I favor it. Perhaps an ethicist ought not work in industry long, for fear of developing golden handcuffs. Nonetheless, working with industry insiders grappling with ethical problems the insiders know they have, as evidenced in each of these thoughtful and thought-provoking pieces, having industry bring outside ethicists in for specific collaborations can be instructive for both sides.

Here we publish the output of two such collaborations that are quite instructive. They raise serious concerns and sound the alarm. No single paper can resolve all of the problems on its own. Nor can single papers fully cover the broad terrain these papers mention, let alone the breadth of implications these warnings imply. Nevertheless, the authors raise sirens that alert us to the development of such problems and forward ethical principles that can create conceptual boundaries that

serve to curb the most immediate and potentially catastrophic possibilities. They raise important ethics concerns and articulate ways in which these concerns can be addressed. We would love to hear your thoughts about the substance of these pieces and about the peripheral issue of how best to work with industry partners in thinking through these complex AI issues.

The last piece in this issue is our annual book review. This time, of the second edition of what was a very important book in the field of clinical ethics, edited by Ford and Dudzinski, *Complex Ethics Consultations. Volume 1: Cases That Haunt Us*.² In 2025, Ford and Dudzinski updated that volume and we publish here a book review by Bryan Pilkington of Hackensack Meridian School of Medicine in New Jersey. As Pilkington makes clear in his introduction: it is a challenge to bring such a work as the 1st edition forward 17 years in a field that has been maturing as quickly as that of clinical ethics consultation. Pilkington also makes clear that the editors have met that challenge skillfully.

Unless one is reading the new volume front to back, the review beautifully summarizes the second edition's overall structure and each chapter's contents in a way that allows the reader to decide how to tackle the 2nd volume's material - either straight through (a useful approach for those of us who read the first volume straight through 17 years ago) or by picking through the chapters in order of one's own interests.

As I write this issue of JoHE's *Rounding*, my own idiosyncratic interests have sent me quickly back to the 2nd edition's new section on organizational ethics and its first case titled, "It's not my responsibility." Given my focus on the two AI articles in this issue and considerations of how important it is to take responsibility relative to the application of AI-assisted technologies, I recommend this chapter. The case presented is a doozy. Although I could go on analyzing the issues and questions raised by this thorough book review, I shall stop here and let you contemplate for themselves the questions Pilkington raises so intriguingly.

Finally, we wish you much interest in your own contemplation of the work presented in this issue, the 10th Anniversary issue of Jack Lynch's death, our beloved and dearly missed institutional champion of all that is ethically excellent in medicine.

Sincerely,



Evan G. DeRenzo, PhD
Editor-in-Chief
Journal of Hospital Ethics
John J. Lynch, MD Center for Ethics
MedStar Health, Washington, DC

REFERENCES

1. DeRenzo, EG. The rightful goals of a corporation and the obligations of the pharmaceutical industry in a world of bioterrorism. In Jonathan Moreno, ed. *After the Terror: Medicine and Morality in a Time of Crisis*. MIT Press, 2012:149-165
2. Ford, PJ and DM Dudzinski, eds. *Complex Ethics Consultations. Volume 1: Cases That Haunt Us* / Edited by Paul J. Ford, Denise M. Dudzinski. Cambridge Medicine. Cambridge University Press, 2008.

FEATURES

Adverse Events in AI-Assisted Pediatric Surgery: Risk Mitigation, Responsibility, Disclosure, and Surgeon Support

Niall M. Jones, MD, FRCSE, FRSM, MCLinLead and Katrina A. Bramstedt, PhD, MA

ABSTRACT: Adverse events (AE), including medical errors, are possible in any surgical procedure; however, AI-assisted pediatric surgery (AIPS) adds more complexity to the topic. This paper explores the inter-relationship between bioethics and Pediatric Surgery in the context of AIPS and the identification, reporting, and investigation of AEs/medical errors, connecting the ethics topics of disclosure, moral courage, responsibility, human-in-the loop, transparency, and accountability. As risk mitigation, we pose limits to the use of fully autonomous AIPS, as well as an AI-timeout amid a hospital governance system. Also, we emphasize that the term “Responsible AI” is a misnomer as the surgeon bears the ultimate responsibility for safe and effective performance, not the AI tool.

KEYWORDS: pediatrics, artificial intelligence, robotic surgical procedures, medical ethics, medical errors, risk management

Introduction

When the first truly international association of Pediatric surgeons, the British Association of Paediatric Surgeons, was formed in 1953, its guiding motto was ‘to set a standard; not to seek a monopoly.’¹ This was, perhaps, a prescient nod to Beauchamp and Childress’ guiding principles of bioethics² – beneficence, non-maleficence, autonomy, and justice. From paternalism to shared decision-making, with plenty of controversy in between – for the pediatric surgeon this means setting the ethical standards for matters such as intersex surgeries,³ bariatric surgery,⁴ duodenal atresia repair in Down syndrome,⁵ hypoplastic left heart syndrome surgeries,⁶ face transplantation,⁷ and the separation of conjoined twins.⁸ With similar normative grounding via the goals of medicine, promoting patient well-being, and harm prevention, surgeons have a launch point to be actively involved in setting standards for ethical AI-assisted surgeries.

In recent years, the use of artificial intelligence (AI) across various aspects of the patient encounter has arisen. The surgeons’ use of AI could occur at numerous steps in the patient pathway, including review of AI-created clinical summaries provided by referring physicians (data from presenting complaint, past medical and surgical history, family history, social history, medications, allergies, travel history), review of AI-generated notes from team meetings and family meetings, review of AI-assessed investigations (imaging, laboratory assays), and the use of AI-decision tools to help decide treatment options. Kim et al.⁹ describe their development of a robot that can undertake clipping of the cystic duct and artery while being ‘instructed’ by a language model – so-called

‘autonomous surgery’ - poisoning autonomous AI-performed surgery for the near future.

This paper analyzes risk and risk mitigation regarding the use of AI in the operating room (OR) as part of medical device software and tools/robots for AI-assisted pediatric surgery (AIPS). We explore the inter-relationship between bioethics and pediatric surgery in the context of AIPS and the identification, reporting, and investigation of AEs/medical errors, connecting the ethics topics of disclosure, moral courage, responsibility, human-in-the loop, transparency, accountability, as well as governance. We highlight the need for careful escalation of the autonomy of robots and commend the introduction of guidelines and training by the surgical colleges. Overall, the addition of AI in the OR creates both responsibilities and accountabilities for the surgeon.

AI in the OR: Is the Surgeon in-the-Loop?

AI can be used as an adjunctive tool (adjunctive to the brain and hands of the surgeon) or a replacement, but in the latter role, *how much, if any, of the surgeon’s clinical and ethical decision-making (cognitive work) should be replaced by AI?* If used as a *replacement*, there is the assumption that the AI tool performs *equal to or better than* the surgeon. When used as an adjunctive tool, there is a risk that in some situations surgeons could assume that the AI is functioning at replacement level performance. Such an assumption could lead to surgeons knowingly or unknowingly removing themselves from the human in-the-loop that the AI needs.

If an AI-assisted tool/robot fails to detect a treatable malignant lesion adjacent to the primary

surgical target and the surgeon is not in-the-loop at the time, the delay in acting on this incidental finding could potentially impact the patient's options for future treatment and their eventual clinical outcome. If an AI-assisted surgical tool/robot is operating autonomously and there is a malfunction that causes it to perform slower than is needed for patient safety (e.g., suction is needed to clear the field of blood for urgent hemostasis with clips, or to contain infected material such as pus or feces to prevent contamination of the field) there is a risk that the tool/robot might continue to perform at sub-optimal speed. Used as an adjunctive tool with a surgeon in-the-loop, the surgeon can intercept and take over for the tool with timely actions. In this way, having a surgeon in-the-loop is a responsibility with associated accountability.

AI and surgical technology can be exciting, but that excitement should be balanced by considerations pertaining to the skill of human surgeons, and the outcomes of their decision-making and technical performance when involving AI in the OR. Consider the following two fictional scenarios, analyzed from the perspectives of a pediatric surgeon and a clinical ethicist:

Scenario #1

A research project with approval from an institutional review board (IRB) involves the retrospective review of an AIPS intestinal atresia case series belonging to a surgeon colleague who ultimately decides to withdraw as a co-investigator, as the preliminary data analysis shows a 90% failure rate with the need for revision surgeries. The AI-assisted tool is a surgical robot that performs various activities, including cutting, retraction, port placement, photographic imaging, stapling, and suturing. The principal investigator respects the colleague's wishes, and also terminates the research project. However, the preliminary findings are discussed at the weekly unit meeting (once juniors/trainees have left the office), urging investigation and corrective action. The implicated colleague refuses to change their practice and continues performing AIPS for intestinal atresia. Communication breaks down with incriminations and recriminations, as well as accusations of bullying. Numerous questions arise: Should a surgeon with a 90% failure rate in a particular AIPS voluntarily stop offering that procedure? Should colleagues (other surgeons, nurses, anesthesiologists) or the hospital stop the surgeon from offering that AI-assisted procedure? What becomes of the whistleblower

who reports the situation and suggests corrective action? Should parents be prospectively informed of such failure rates involving this AI-assisted robot and surgeon? Should the use of the AIPS robot be broadly halted across all surgeries and an investigation performed to determine if the device or its software is malfunctioning?

A Pediatric Surgeon's View: Surgeons, hospitals, patients, and families all want surgical success. In the situation described, few surgeons would want to continue performing AIPS with such high failure rates. Most surgeons would address the situation by seeking advice from colleagues, attending workshops, discussing the AI technology with the manufacturer, or having a mentor attend cases. In scenario #1, accepting a 90% failure rate and carrying on is unethical and unprofessional due to misalignment with the goals of medicine, the need to promote patient benefit, and the need to avoid harm. With knowledge of the failure rate a colleague's priority is to advocate for the child. The hospital and leadership of the surgical department also have accountability in such situations due to their obligation of a duty of care toward their patients. This could include quarantining the AIPS robot and initiating a manufacturer investigation. Discussion of the findings can be done in a non-confrontational way and with supportive measures and corrective actions as part of a defined risk management plan (e.g., sharing cases with a colleague; restricting AIPS cases to colleagues with lower failure rates; halting AIPS until all implicated device failure modes are rectified; requiring additional AIPS training).

Moral shyness¹⁰ can propel more failed operations (and potentially an unidentified problem with the AI tool/robot), as well as more distressed children and families. A formal investigation with transparency (including verification of the surgeon's AI tool training and tool performance) can bring clarity and promotes patient-centricity. Access to the expertise of a bioethicist could provide a clear structure or framework by which to discuss the pertinent moral issues (e.g., beneficence, non-maleficence, autonomy, professionalism, responsibility, accountability, stewardship) sometimes lacking in surgical units.

Scenario #2

Surgeon A calls on-call *Surgeon B* to see a sick baby at 6 a.m. *Surgeon A* had operated on the baby as a planned AIPS inguinal hernia repair 3 weeks

prior. The operation was reported as ‘standard,’ but the baby developed signs of GI obstruction resulting in bleeding that necessitated an emergency laparotomy by *Surgeon A*. Some ‘bruised’ intestine is discovered but nothing more is done. The baby is slow to recover and eventually is discharged on the 10th day of admission (a second day discharge is standard). This baby is now in the emergency department experiencing bilious and ‘coffee ground’ vomiting, as well as hypovolemic shock. *Surgeon B* inserts an intra-osseous needle, resuscitates the baby, and performs a laparoscopy/laparotomy, finding 30cm of dead jejunum and 5cm of dead sigmoid colon. The sigmoid tissue was inadvertently incorporated in a suture during *Surgeon A*’s first AIPS operation, while the jejunal tissue became adherent to it, leading to interrupted blood supply with subsequent ischemia and necrosis. *Surgeon B* informs *A* of the findings immediately after the operation, but *A* denies any responsibility. *Surgeon B* informs the baby’s mother of the findings and the cause of the findings. The infant’s mother lodges an official complaint with the hospital, supported by *Surgeon B*. The hospital initiates disciplinary action against *Surgeon B* for speaking negatively about a colleague, the institution, and for other unsubstantiated unprofessional behaviors. Three months later, the official investigation by the local health authority is still in progress and the hospital offers *Surgeon B* the choice of resignation or immediate termination. *Surgeon A* retains surgical privileges at the hospital with no corrective action. No manufacturer investigation of the AI robot is performed.

A Pediatric Surgeon’s View: With or without AI, every surgeon has the potential to make an error – what is ethically important is disclosure, accountability, and corrective action. There was an opportunity during the laparotomy on the night of the original AIPS operation for *Surgeon A* to explore the abdomen fully, identify what was causing the bruising seen on the bowel wall (ischemia developing from misplaced sutures), and fix the problem. The baby would have recovered, albeit slower than expected. Instead, nothing therapeutic was done during this operation. In spite of a prolonged recovery, no further surgical consideration was given to the baby until emergency re-admission in a state of extremis. Also, the AI-assisted robot was not investigated for its possible role in the suturing error.

Surgeon A was made aware of the second laparotomy and findings, providing another opportunity to accept that an error had occurred during

the original operation and to disclose the matter to the baby’s parents. *Surgeon B* was obliged to report findings to the baby’s parents, including the cause of the findings, and to support the official complaint to protect future patients. The presence of a bioethicist during a team discussion of the case could potentially ease interpersonal friction and tension, returning the focus to the patient. As the AI-assisted robot could have contributed to the misplaced sutures, an investigation of the device as part of the reportable adverse event was also warranted but not performed, which is problematic.

An Ethicist’s View: The role of a hospital ethicist is to dive deeply into each case, to understand the clinical situation as well as the contextual features, the patient’s quality of life (now and future expectations), and the patient and family preferences for treatment.¹¹ Typical pediatric surgery issues for an ethicist include futility, in-utero fetal surgery, organ transplant¹² and re-transplant candidacy, parent-to-child living organ donation, pediatric deceased donation,¹³ adolescent bariatric surgery, and reconstructive plastic surgery subsequent to child abuse. Other topics include the creation of and donations from savior siblings,¹⁴ and posthumous sperm retrieval for in vitro fertilization.¹⁵ The advent of AI in surgery raises hopes for equal or improved outcomes; however, AI is not magic, it is not human, and it lacks decisional skills and a moral compass.

The use of the term “responsible AI” is popular in academic, research, and media arenas. Words matter, especially when AI literacy is in its infancy. While “responsible AI” is meant to engender uptake and trust of AI, it is a misnomer since AI as a technology cannot bear responsibility (unless, for instance, AI becomes a citizen).¹⁶ In the setting of AIPS, the surgeon using the AI bears primary responsibility and accountability, with the AI-developer also holding a weighty responsibility and accountability in terms of tool validation for safety and performance. Lyons et al. express this succinctly: “the problem with the terms such as responsible AI (and AI ethics more generally) is that such terms (if misused) can promote an expectation that the onus for ethical behavior rests with the AI, which may create unrealistic expectations for AI technologies as they exist today.”¹⁷

In the two scenarios posed, the following are key considerations for the ethicist: risk (the level of possibility of patient harm), harm (injury/damage/loss), and safety (protection from harm). Further, the absence of a legal requirement to re-

port significant risk and harm is immaterial to an ethical duty to report. A surgeon's moral courage to self-report poor performance can be impacted by many things, including workplace culture and role-power differential (trainee/attending/hospital executive). For workplace colleagues, the same variables exist, as well as whether there are legal protections for whistleblowers. The value of a manufacturer's investigation of a device is ethically significant as the analysis can potentially detect problems with components and raw materials, as well as assembly and software/programming. It may be that an adverse event or error in surgery is related to use of the AI-assisted tool.

Risk Mitigation and AIPS

Knudsen et al.¹⁸ have identified 5 levels of automation in robotic surgery and reports that the autonomy of robots engaged in surgery has reached Level 3 (conditional automation). At Level 3, surgical robots are performing some surgical steps without human input or humans placing the device within the surgical site. With the current state of AI governance and a culture that often misunderstands AI ethics, we argue that Level 4 (high automation) and Level 5 (complete automation) should be off limits, for now. Progress in these areas could be rapid, thereby ethically permitting advancement to patient use. According to Knake¹⁹ the future is now for AI in pediatrics as AI is everywhere across the specialty. Nonetheless, at this time, a human [i.e., surgeon] in-the-loop remains essential and aligns to the step-level autonomy of AI proposed by Kim et al.⁹ Having a surgeon in-the-loop should be viewed as an ethical responsibility that promotes patient well-being, potentially reduces harm, and could mitigate a mindset prone to offshoring accountability to the robot itself.

While there is a time and place for surgical speed, there should also be concern of the need for a timeout/pause.²⁰ We pose that this timeout period customarily used to prevent performing the wrong procedure, performing it on the wrong patient, wrong site errors, could also be used to cognitively reflect on the use of AI in the procedure (e.g., indication for use, adjunctive vs replacement) so as to reinforce the necessity of the human in-the-loop as appropriate to the regulatory labeling of the AI-assisted tool and the surgeon's experience and skill. Additionally, having a human in-the-loop is a reminder that while AI can give linguistic apologies in the setting of errors (when prompted), the surgeon remains the moral agent

responsible for disclosing and apologizing in such situations.²¹

Institutional responsibility and accountability means that problems that have been identified with AIPS cannot be ignored. The use of AI-assisted surgical tools, including robots, requires surgeons and institutions to re-evaluate their surgical consent process and documentation practices to ensure transparency to patients regarding the use of such devices. Also, regulatory labeling may require these devices to be placed amid a system of planned maintenance and calibration. Hospitals should block device use until surgeon training has been successfully completed and documented. Institutional policies and procedures that involve risk management and quality are needed as these incorporate strategies for device quarantine and return for manufacturer investigation, as well as formalized plans that include both action and follow-up to ensure compliance and risk prevention.

Mandatory Notifications and Whistleblowing by Surgeons

Mandatory notifications are enshrined in law such that the law itself describes the type of behaviors that risk patient safety and must be reported to the medical board and the protections for the reporting surgeon, as well as the surgeon being reported. Depending on jurisdiction, these behaviors generally include the following:

- I. Practicing with an impairment and thus placing the public at risk of substantial harm.
- II. Practicing while intoxicated by alcohol or drugs.
- III. Practicing in a way that significantly departs from accepted professional standards such that the public is at risk of harm.
- IV. Engaging in sexual misconduct in connection with their practice.

From an ethical perspective, reporting these behaviors aligns with harm prevention. In some jurisdictions, failure to make a mandatory notification (a legal requirement) can result in a penalty for the witnessing surgeon (e.g., registration caution from a medical board). Failing to keep a surgeon in-the-loop when ethically appropriate (and when required by regulatory labeling of the AI-assisted system) could be viewed by some medical boards as a significant departure from accepted medical standards (putting patients at risk of harm).

Whistleblowing is another form of notification in cases wherein the behavior may or may not be subject to a legal requirement to report, but the risk of patient/societal harm is present. Depending on the jurisdiction, there may be protections for whistleblowers, such as anonymous reporting, retaliation prohibitions, and even financial rewards for reporting unnecessary surgeries, billing for services not rendered, or other fraudulent practices that lead to false claims being submitted to government healthcare programs.²²

No matter the setting (OR, ED, inpatient, outpatient), a hesitance to report can be a matter with which ethicists can help. Such guidance is advisory in nature and aims to provide the surgeon with an understanding of the problem/dilemma from an ethical lens, as well as distinguish what are ethically appropriate and ethically inappropriate choices. Ethicists can also help to identify and relieve moral distress in situations involving adverse events, as well as matters related to personal, religious values,²³ abusive patients, futility,²⁴ and relationships with peers, including bullying and witnessing poor job performance. The chair of a surgical program should always be supportive of colleagues in times of moral distress and in relation to reporting.

Surgical professional societies have an important role in relation to these concerns. In 2025, the Royal Australasian College of Surgeons announced its intent to establish an advisory group on artificial intelligence in surgery in order to help guide their leadership in terms of surgical standards, professionalism, and education in the context of AI and machine learning. This includes setting ethical guidelines that promote safe care.²⁵ The American College of Surgeons offers an online course (Artificial Intelligence and Machine Learning: Transforming Surgical Practice and Education)²⁶ that includes content pertaining to the ethics of AI-assisted surgery. Medical schools, surgical fellowships, and surgical ethics fellowships should incorporate AI ethics topics into their training curricula, evolving the curricula as developing AI technologies creates the potential for new ethical dilemmas.

Conclusion

AI ethics should remain in the sphere of normativity and human responsibility, avoiding a solely compliance-based legal framing. In the pediatric surgical setting, ethical AI means using AI in ways that promote patient benefit and minimize harm,

with surgeons holding responsibility and accountability for AI use amid a robust institutional governance plan. In the setting of AI harm and risk of harm, promoting moral courage, preventing reporting hesitancy, as well as supporting the surgical reporter and the reported²⁷ are essential in order to facilitate a culture of safety and patient-centricity. A patient-centric lens creates an environment for ethical guardrails that shape AI-assisted surgical practice.

Disclosures: Niall M. Jones is a paid member of the F. Hoffmann-La Roche AG Scientific Ethics Advisory Committee. Katrina A. Bramstedt is an employee of F. Hoffmann-La Roche AG, owns stock in F. Hoffmann-La Roche AG, and is chair of the F. Hoffmann-La Roche AG Scientific Ethics Advisory Committee.

REFERENCES

1. Spitz L. The history of paediatric surgery in the United Kingdom and the influence of the national health service on its development. *J Pediatr Surg*. 2012;47(1):29-35.
2. Beauchamp TL, Childress JF. *Principles of Biomedical Ethics*. Oxford University Press, 8th edition; 2019.
3. Clune-Taylor C. Arguments for a ban on pediatric intersex surgery: a dis/analogy with Jehovah witness blood transfusion. *Bioethics*. 2024; 38(5):460-468. doi:10.1111/bioe.13280.
4. Moore JM, Glover JJ, Jackson BM, Coughlin CR 2nd, Kelsey MM, Inge TH, Boles RE. Development and application of an ethical framework for pediatric metabolic and bariatric surgery evaluation. *Surg Obes Relat Dis*. 2021; 17(2):425-433. doi:10.1016/j.soard.2020.09.014.
5. Fost N. "The Hopkins Mongol Case": the dawn of the bioethics movement. *Pediatrics*. 2020;146(Suppl 1):S3-S8. doi:10.1542/peds.2019-3692.
6. Edwin F, Edwin AK, Palacios-Macedo A, Mamorare H, Yao NA. Management of hypoplastic left heart syndrome in low-resource settings and the ethics of decision-making. *World J Pediatr Congenit Heart Surg*. 2022;13(5):609-614. doi:10.1177/21501351221103511.
7. Dorante MI, Tasigiorgos S, Pomahac B. Pediatric face transplantation: a viable option? *Curr Opin Organ Transplant*. 2018;23(5):598-604. doi:10.1097/MOT.0000000000000572.
8. Woo YT. Evaluation of the ethical controversies in separation surgery of conjoined twins: a review of the literature. *Cureus*. 2025;17(1):e77382. doi:10.7759/cureus.77382.
9. Kim JW, Chen J-T, Hansen P, Shi LX, Goldenberg A, Schmidgall S, Scheikl PM, Deguet A, White BM, Tsai DR,

- Cha RJ, Jopling J, Finn C, Krieger A. SRT-H: a hierarchical framework for autonomous surgery via language-conditioned imitation learning. *Sci Robot*. 2025;10:eadt5254.
10. Antiel RM, Blinman TA, Rentea RM, Gonzalez KW, Knott EM, Juang D, Oyetunji T, Holcomb GW, Angelos P, Lantos JD. When a surgical colleague makes an error. *Pediatr atr*. 2016;137(3):e20153828.
 11. Anderson EE. What we talk about when we talk about goals. *Virtual Mentor*. 2007;9(6):407-9. doi:10.1001/virtualmentor.2007.9.6.fred1-0706.
 12. Carrion L, Bramstedt KA. Exploring the ethical complexity of pediatric organ transplant candidates and COVID-19 vaccination: tensions between autonomy and beneficence, children and parents. *Pediatr Transplant*. 2023; 27(1):e14408. doi:10.1111/ptr.14408.
 13. Martin DE, Nakagawa TA, Siebelink MJ, Bramstedt KA, et al. Pediatric deceased donation - a report of the Transplantation Society meeting in Geneva. *Transplantation*. 2015;99(7):1403-9. doi:10.1097/TP.0000000000000758.
 14. Cherkassky L. The interfamilial principle and the harvest festival. *Eur J Health Law*. 2016; 23(1):61-79. doi:10.1163/15718093-12341379.
 15. Ovics SO, Baram S, Nothman S, Weiss A, Beck-Fruchter R. Perimortem and postmortem sperm acquisition: review of clinical data. *J Assist Reprod Genet* 2022;39:977–986. doi:10.1007/s10815-022-02427-x.
 16. O'Sullivan S, Nevejans N, Allen C, Blyth A, Leonard S, Pagallo U, Holzinger K, Holzinger A, Sajid MI, Ashrafian H. Legal, regulatory , and ethical frameworks for development of standards in artificial intelligence (AI) and autonomous robotic surgery. *Int J Med Robotics Comput Assist Surg*. 2019; 15:e1968. doi:10.1002/rcs.1968.
 17. Lyons JB, Hobbs K, Rogers S, Clouse SH. Responsible (use of) AI. *Front Neuroergon*. 2023;4:1201777. doi:10.3389/fnrgo.2023.1201777.
 18. Knudsen JE, Ghaffar U, Ma R, Hung AJ. Clinical applications of artificial intelligence in robotic surgery. *J Robotic Surg*. 2024;18:102. doi:10.1007/s11701-024-01867-0.
 19. Knake LA. Artificial intelligence in pediatrics: the future is now. *Pediatr Research*. 2023;93:445-446. doi:10.1038/s41390-022-01972-6.
 20. Jones NM. Tune-in and time-out: toward surgeon-led prevention of 'never' events. *J Pat Saf*. 2019;15(4):e36-e39.
 21. Magnus PD, Buccella A, D'Cruz J. Chatbot apologies: Beyond bullsh*t. *AI and Ethics*. 5(5):5517-5525. doi.org/10.1007/s43681-025-00800-x.
 22. 18 U.S. Code § 287 - False, fictitious or fraudulent claims
 23. Millis MA, Vitous CA, Ferguson C, Van Wieren I, Kalata S, Shen MR, MacEachern M, Suwanabol PA. To feel or not to feel: a scoping review and mixed-methods meta-synthesis of moral distress among surgeons. *Ann Palliat Med*. 2023;12(2):376-389. doi:10.21037/apm-22-916.
 24. Zimmermann CJ, Taylor LJ, Tucholka JL, Buffington A, Brasel K, Arnold R, Cooper Z, Schwarze ML. The association between factors promoting nonbeneficial surgery and moral distress: a national survey of surgeons. *Ann Surg*. 2022;276(1):94-100. doi:10.1097/SLA.0000000000004554.
 25. Royal Australasian College of Surgeons. *Expressions of interest: Advisory Group on Artificial Intelligence in Surgery*. 26 August 2025. <https://www.surgeons.org/News/News/Expressions-of-interest---Advisory-Group-on-Artificial-Intelligence-in-Surgery> Accessed 19 November 2025.
 26. American College of Surgeons. *Artificial Intelligence and Machine Learning: Transforming Surgical Practice and Education*. Undated. <https://www.facs.org/for-medical-professionals/education/programs/artificial-intelligence-and-machine-learning-transforming-surgical-practice-and-education/>. Accessed 19 November 2025.
 27. Berman L, Rialon KL, Mueller CM, Ottosen M, Weintraub A, Coakley B, Brandt ML, Heiss K. Supporting recovery after adverse events: an essential component of surgeon well-being. *J Pediatr Surg*. 2021;56(5):833-838. doi:10.1016/j.jpedsurg.2020.12.031.

AI Errors and Human Complacency: The Output Can Be Tragic

Timothé Ménard, PharmD, MSc, MA and Katrina A. Bramstedt, PhD, MA

ABSTRACT: There's no argument about the speed of AI – it's fast. And sometimes, in research and healthcare, fast is wanted and needed. But fast should not compromise the integrity of the AI outputs. Research performing organizations and technology developers are promoting the use of AI tools including Large Language Models (LLMs) as an aid to the research process. Clinicians are using such tools to help inform their clinical decision-making. In research and healthcare, some errors, if not detected, could be irreversible or catastrophic. Exploring two real-life examples, alarm is raised about a potentially deadly combination: AI errors + human complacency. The role of operationalized AI Ethics Principles (safety-by-design, transparency, human-in-the-loop, and accountability) is emphasized as a form of risk mitigation.

KEYWORDS: artificial intelligence, medical ethics, medical errors, risk management, accountability, research

Introduction

One of the selling points of artificial intelligence (AI) is its speed. In the settings of healthcare research and clinical care, however, this speed risks transforming AI from an adjunctive tool to an unmonitored replacement, with the potential of undetected errors as collateral damage. Consider the general example of taking a shortcut while on a journey from A to B. The shortcut fails if the resultant journey is on a low-quality, unlit, unpaved road and your car lacks 4-wheel drive, potentially slowing travel time and damaging the vehicle. AI in research and healthcare can be a timesaver, generating valuable insights; however, it can also be potentially harmful and therefore requires risk mitigation.

Research organizations and technology developers are promoting the use of AI tools including Large Language Models (LLMs) as an aid to the research process. Clinicians are using such tools to inform their clinical decision-making. While not the focus of this manuscript, patients are using AI tools to educate themselves about the diagnosis, treatment, and prevention of illness. Using two real-life case examples of AI in the medical domain, the importance of the ethical principles of safety-by-design, transparency, human in-the-loop, and accountability are explored.¹ These principles, grounded in the professional ethics of some organizations and industries, are commonly agreed upon as key AI ethical principles linked to trust-building.^{2,3} Yet, these principles are idle words if stated as mere platitudes and not operationalized in AI development and use. Applying these ethical principles adds reflectivity to the innovation process and its outputs, and helps mitigate risk. Applying them during use adds a layer of safety for clinicians and their patients. Furthermore, while AI might be able to offer a linguistic apology as an automated response for an error,⁴ the error, if not detected by the clinician/user could be harmful,

and that harm irreversible, or catastrophic.

Examples

A: Sepsis is a life-threatening overreaction of the immune system during an infection, leading to the need for hospitalization and urgent treatment. An AI-assisted algorithm intended for the prediction of sepsis evidenced a high false-negative rate when deployed in U.S. hospitals. Specifically, the model missed approximately two-thirds of actual sepsis cases.⁵

B: Post-operative acute intracranial hemorrhage is bleeding after a cranial surgery that requires surgical evacuation. A problem with false-positives has been reported whereby AI prioritization software falsely flags imaging findings as post-surgical acute intracranial hemorrhage when patients did not actually evidence such.^{6,7}

AI errors are problematic, but AI errors in the setting of human complacency have the potential to result in harm that could even be irreversible or catastrophic (patient death). In fact, the online Merriam-Webster Dictionary definition of complacency (“self-satisfaction especially when accompanied by unawareness of actual dangers or deficiencies”) includes the following example to help explain the definition: “When it comes to safety, complacency can be dangerous.”⁸ In Example A, Wong et al.⁵ discovered that an AI-assisted algorithm intended for the prediction of sepsis evidenced a high false-negative rate upon use in the hospital setting. Notably, approximately two-thirds of actual sepsis cases were missed by the model. Undiagnosed and ultimately untreated treated sepsis can quickly result in death. A safety-by-design approach is needed for high-risk AI tools such as those for sepsis prediction. This approach proactively considers patient safety by identifying po-

tential hazards, evaluating the model against those hazards, and targeting weaknesses in the model in order to mitigate clinical risk. While a safety-by-design approach has already been identified for sepsis treatment, accurate diagnosis is the first step.⁹ A human in-the-loop providing physician oversight ensures that the AI output is not accepted at face value and that additional diagnostic measures are employed against a presentation of symptoms that point toward sepsis (e.g., tachycardia, hypotension, confusion, purpura fulminans.)

Example B highlights that since post-operative acute intracranial hemorrhage can be potentially life-threatening (i.e., intracranial pressure rises leading to brain compression, herniation, coma, and death), an AI-assisted tool that could rapidly and correctly identify this situation would be valuable. False-positives, however, can create anxiety for the patient, family, and medical team, while potentially triggering unnecessary imaging studies with associated temporal and financial costs. Confirming a post-operative acute intracranial hemorrhage might indeed involve imaging studies, however, these are standardized, simple tests. As such, this is another example of a high-risk AI-assisted tool that would benefit from the proactive safety-by-design approach. Considering the vital role of a human in-the-loop reminds us that these imaging tools are not a replacement for a physician's eyes, but merely adjunctive to a physician's own clinical judgement. Consistent with the American College of Radiologists, physicians have a vital role in the supervision of AI imaging technology in order to ensure patient safety.¹⁰

In both examples, physician users require transparency from developers about the presence of AI in the tool, the limitations of the tool's performance, and the role of a physician in-the-loop. This transparency is needed even when there may be black box elements of a decisional support system that are not understood, because transparency helps to inform physician users of their responsibilities when using the tool (ultimately linking to matters of ethical accountability).

The cases presented are only a sampling of the reality that AI deployment and use should leverage AI ethical principles. Other interesting cases include machine learning models for COVID-19 prognostication that lacked potential clinical use due to the problem of garbage-in-garbage-out, reminding developers that robust AI outputs require robust AI inputs.¹¹ Also, there is the matter of the identification of ChatGPT as an unreliable research assistant when it is used for literature

searches due to its creation of fake citations.¹² Notably, there are numerous open/free and commercial LLM platforms being promoted as rapid tools for research citation assistance (i.e., finding references on specific topics, performing scoping literature reviews, collating found references into bibliographies). LLM developers should disclose to users information about the tool's indications for use (scope and limits), validation and reliability, and the need for output verification. Research organizations should be satisfied with the tools' performance before promoting it for researcher use. Researchers using such LLMs must also be in-the-loop, disclosing their use of the LLM, and checking for output veracity.

The Legal Lens: Negligence and Liability

Returning to Merriam Webster ("When it comes to safety, complacency can be dangerous"),⁸ the satisfaction of using an AI tool as a timesaver can be quickly diminished if preventable harm results from the absence of the human in-the-loop. One's comfort zone using AI should not be so comfortable such that the human may be missing when they are needed (e.g., a labelled requirement of the tool). From a legal lens, courts can make a determination of negligence: failure to behave with a reasonable level of care when the failure to take action results in harm to another person (omission). Considering chatbots in the healthcare context, the legal case of *Moffatt v Air Canada* is a reminder that deployers may be liable in negligence for AI-generated content or advice from their chatbots.¹³ If a physician or patient acts on faulty information (AI output), courts can determine that the deployer (physician, health system) is held responsible for negligence.¹⁴ If patient death results, there is also the potential for criminal law applicability depending on regional jurisdiction (e.g., gross negligence manslaughter, corporate manslaughter).¹⁵ The legal context includes laws and regulations, but in the absence of these (AI technology is evolving faster than the legal and regulatory landscape] the ethical duty of care and professional ethics remain. Similarly, even if legal accountability is dispersed across developers, deployers, and users, the ethical duty of care and professional ethics is important to help keep patient-centricity in focus.

Overconfidence and Poor Data/AI Literacy

Overconfidence, coupled with poor data and AI literacy, creates a significant risk. An emerging challenge is a phenomenon not yet extensively discussed in the literature: because AI tools are so easy to use, sometimes appearing almost 'magical,' users can develop an overconfidence in them as a simple solution for nearly any problem. This issue is magnified when users lack a fundamental understanding of the technology's limitations, predisposing them to automation bias.¹⁶ This creates a significant risk as (generative) AI systems are rolled out widely, and existing AI literacy programs may be insufficient to counter it. Henderson briefly discusses overconfidence in the context of automation bias in medical AI applications.¹⁷ Keeping a human in-the-loop to check AI outputs can mitigate the risk of blindly accepting them, but only if that human has sufficient literacy to critically evaluate the AI's suggestions. To reduce the risk of harm, AI-assisted systems designed for use by patients need to consider these issues of data and AI literacy, as well as user overconfidence.

There is already concern about the overuse of AI potentially causing failure to develop critical thinking, as well as problems with physicians having the ability to recognize biased AI.¹⁸ Additionally, a recent study by Waldock et al. explored the ability of 148 final year UK medical students' ability to identify LLM errors in clinical vignettes.¹⁹ They found that only 61% (median) identified the LLM errors. This means that those assumed to possess higher data and AI literacy (compared to patients) are not spotting AI errors.

Conclusion

The arguments presented are a wake-up call for users of AI. Indeed, there are benefits of using AI in research and healthcare; however, the risks must be acknowledged as well as the need to warn against deployer and user complacency. Faster research and healthcare isn't better if it is error-filled.²⁰ Additionally, the concept of LLMs acting with intentionally harmful behaviors is important for reflection.²¹

Disclosures: At the time this article was written, Timothé Ménard and Katrina A. Bramstedt were employed by Roche and owned F. Hoffmann-La Roche Ag stock.

REFERENCES

1. Ménard T, Bramstedt KA. Developing a Set of AI Ethics Principles to Shape Ethical Behavior in Drug Development. *Ther Innov Regul Sci*. 2025; 59:399–402
2. IFPMA Artificial Intelligence Principles (2022). <https://www.ifpma.org/publications/ifpma-artificial-intelligence-principles/> Accessed 17 Nov 2025
3. ISO/IEC TR 24028:2020(en) Information technology — artificial intelligence — overview of trustworthiness in artificial intelligence. <https://www.iso.org/standard/77608.html>. Accessed 17 Nov 2025
4. Magnus P, Buccella A, D'Cruz J. Chatbot Apologies: Beyond Bullshit. *AI Ethics*. 2025; 5(5):5517-5525.
5. Wong A, Otles E, Donnelly JP, et al. External Validation of a Widely Implemented Proprietary Sepsis Prediction Model in Hospitalized Patients. *JAMA Intern Med*. 2021;181(8):1065–1070
6. Cardoso O, Adly M, Hamade M, Saigal K, Saigal G. False Positives in Artificial Intelligence Prioritization Software for Intracranial Hemorrhage Identification in the Postoperative Period: A Report of Two Cases. *Cureus*. 2023;15(8):e44215. <https://doi.org/10.7759/cureus.44215>.
7. Seyam M, et al. Utilization of Artificial Intelligence-based Intracranial Hemorrhage Detection on Emergent Noncontrast CT Images in Clinical Workflow. *Radiol Artif Intell*. 2022;4(2):e210168. <https://doi.org/10.1148/ryai.210168>.
8. "Complacency." *Merriam-Webster.com Dictionary*, Merriam-Webster, <https://www.merriam-webster.com/dictionary/complacency>. Accessed 17 Nov 2025
9. Jia Y, et al. Safety-driven design of machine learning for sepsis treatment. *J Biomed Inform*. 2021;117:103762. <https://doi.org/10.1016/j.jbi.2021.103762>.
10. Brady AP, et al. Developing, purchasing, implementing and monitoring AI tools in radiology: Practical considerations. A multi-society statement from the ACR, CAR, ESR, RANZCR & RSNA. *J Med Imaging Radiat Oncol*. 2024;68(1):7-26. <https://doi.org/10.1111/1754-9485.13612>.
11. Roberts M, Driggs D, Thorpe M. et al. Common Pitfalls and Recommendations for Using Machine Learning to Detect and Prognosticate for COVID-19 Using Chest Radiographs and CT scans. *Nat Mach Intell*. 2021;3:199–217. <https://doi.org/10.1038/s42256-021-00307-0>.
12. Merz JF. ChatGPT Just Makes Stuff Up: A Conversation on a Controversial Topic. *Hastings Bioethics Forum*. <https://www.thehastingscenter.org/chatgpt-just-makes-stuff-up-a-conversation-on-a-controversial-topic/> Accessed 17 Nov 2025
13. Moffatt v Air Canada BCCRT 149, 2024. <https://www.mccarthy.ca/en/insights/blogs/techlex/moffatt-v-air-canada-misrepresentation-ai-chatbot>
14. P9_TA(2020)0276 Civil liability regime for artificial intelligence. European Parliament resolution of 20 October 2020 with recommendations to the Commission on a civil

liability regime for artificial intelligence (2020/2014(INL)) (2021/C 404/05). <https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:52020IP0276>

15. Smith H. Artificial Intelligence for Clinical Decision-Making: Gross Negligence Manslaughter and Corporate Manslaughter. *New Bioeth.* 2024;30(3):228-242. <https://doi.org/10.1080/20502877.2024.2416862>

16. Goddard K, Roudsari A, Wyatt JC. Automation bias: a systematic review of frequency, effect mediators, and mitigators. *J Am Med Inform Assoc.* 2012;19(1):121-7. <https://doi.org/10.1136/amiajnl-2011-000089>

17. Arnold MH. Teasing out Artificial Intelligence in Medicine: An Ethical Critique of Artificial Intelligence and Machine Learning in Medicine. *J Bioeth Inq.* 2021;18(1):121-139. <https://doi.org/10.1007/s11673-020-10080-1>

18. Abdulnour RE, Gin B, Boscardin CK. Educational Strategies for Clinical Supervision of Artificial Intelligence Use. *N Engl J Med.* 2025;393(8):786-797. <https://doi.org/10.1056/nejmra2503232>.

19. Waldock WJ, et al. Which curriculum components do medical students find most helpful for evaluating AI outputs? *BMC Med Educ.* 2025;25(1):195. <https://doi.org/10.1186/s12909-025-06735-5>

20. Karpowicz MP. On the Fundamental Impossibility of Hallucination Control in Large Language Models. <https://arxiv.org/pdf/2506.06382>

21. Lynch A, et al. Agentic Misalignment: How LLMs could be Insider Threats. Anthropic Res. <https://www.anthropic.com/research/agentic-misalignment>

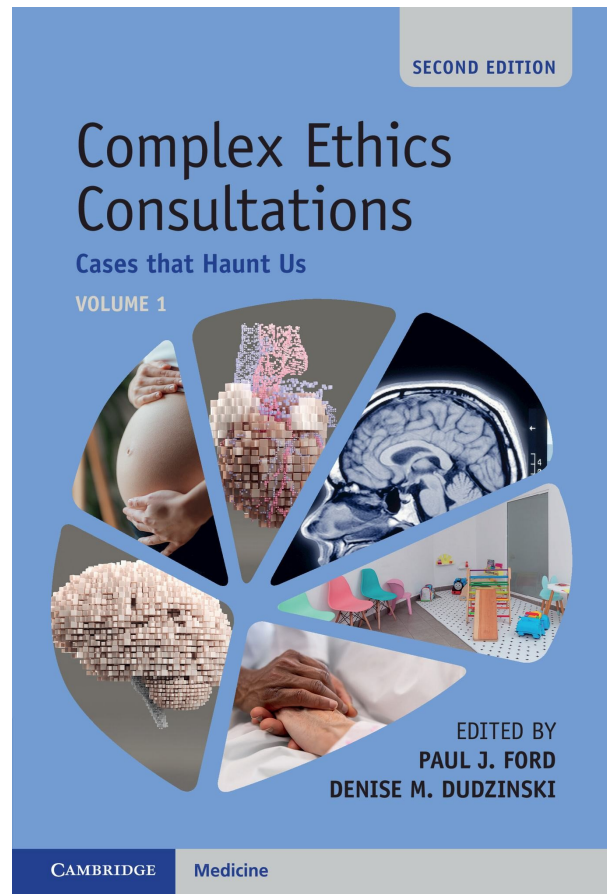
BOOK REVIEW

Returning to Old Haunts: Complex Ethics Consultations 17 Years Later Bryan Pilkington, PhD

Ford, Paul J., and Denise M. Dudzinski, eds. *Complex Ethics Consultations. Volume 1: Cases That Haunt Us*. Second edition. Cambridge University Press; 2nd Edition. June 5, 2025. 281 p. \$32.99 ISBN: 978-1009400954

Seventeen years have passed since the publication of the first edition of *Complex Ethics Consultations. Volume 1: Cases That Haunt Us*.¹ As the editors note in their introduction to the 2nd edition, those cases, and that essential descriptor, have made their way into the bioethics canon and parlance, highlighting the importance of the first edition's contribution to the literature and, indeed, the field. This situation presents an interesting challenge for a sequel; more so than that which standardly faces the second edition of any book. A standard second edition of a volume might add new content or house-revised versions of previously published work, bringing the volume up-to-date, so to speak, with the current landscape. It is a process that can be executed repeatedly, as bioethicists have seen with some influential texts, such as Beauchamp and Childress's *Principles of Biomedical Ethics*, the ninth edition of which was published earlier this year.² Yet, considering the powerful cases and lasting narratives included in the first edition of *Cases*, the challenge of a second edition would not be met through mere updates. To move away from those original cases would be to lose something substantial (both for those who study bioethics, generally, and also for those either practicing or training in clinical ethics consultation), for it was those cases and their discussion that helped to normalize what it is that clinical ethics consultants may very well endure during the course of their work.

Much has happened in bioethics and in the practice of clinical ethics consultation within the last 17 years. Moves toward professionalization have increased, typified by the advent of the HEC-C designation. Artificial intelligence is impacting the practice of medicine and, in turn, what it is that clinical ethics consultants must address - not to mention other advances in technology, changes in laws, cultural and societal shifts, health systems practice revisions, and an increasingly extensive bioethics literature. The challenge was in place: how to revise such an impactful set of case discus-



© 2025 Cambridge University Press

ussions, bringing that work forward almost two decades without losing the power of the individual narratives, case analyses, and professional reflections housed in the original edition. The editors have met that challenge and done so deftly with a second edition that not only adds new voices and insights, but retains the essential cases and discussions from the first.

Seven parts comprise the second edition. Each of these includes four cases from the first edition with new commentary. By structuring the second edition in this way, the editors have both retained

the valuable content of the first edition and updated the volume through the lens of seven new contributors.

Part I, "Starting at the Beginning: Prenatal and Neonatal Issues," focuses on ethical dilemmas at the beginning of life. Its structure includes case narratives, professional reflections, outcome summaries, questions for discussion, references, and of course a description of the haunting aspects of determining quality of life in neonatal intensive care units (NICUs). These challenging aspects include the pain experienced by a dying infant, difficult choices that both parents and clinicians face, and ethical questions surrounding maternal-fetal surgeries. New to the second edition is a commentary by Lucia Wocial including her thoughts on each of the four case discussions.

Part II, "The Most Vulnerable of Us: Pediatrics," moves from the beginning of life to its early stages, with four pediatric cases housed in chapters following the same structure as the rest of the book. These chapters address the emotional impact of a child's death, the challenges of communication between medical staff and a patient's family, the importance of including the patient's own voice, and the lasting effects of failing to communicate with parents. Part II's new offering is a commentary on these pediatric cases by Nneka Sederstrom.

Part III, "Diversity of Desires and Limits of Liberty: Psychiatric and Psychological Issues," engages ethical considerations associated with mental health through cases that touch on challenging patient-staff dynamics, the use of "Ulysses contracts" or advance directives for psychiatric care, difficulties that can arise in assessing a psychiatric patient's needs, as well as a patient's refusal to eat. The novel contribution to Part III is a commentary on these four cases by Maya Scott.

Part IV, "Withholding Therapy with a Twist," explores nuanced situations where the decision to withhold therapy is not straightforward. Cases discussed in this section include reflection on conflicts between a surgeon and an ethicist, the moral ambiguity of a good outcome that was unanticipated, and the debate over what counts as a "terminal condition." Crystal Brown's commentary brings these cases, originally published in 2008, up to date.

Part V, "The Unspeakable/Unassailable: Religious and Cultural Beliefs," offers four cases in which healthcare intersects with diverse religious and cultural values. These cases touch on adolescent pregnancy while respecting cultural norms and confidentiality, ethical considerations of a liv-

ing organ donation between church members, Islamic beliefs regarding futility and death, and understanding suffering as an expression of God's will. Mahwish U. Ahmad's commentary on these articles is the new addition here.

The penultimate section, Part VI, "Human Guinea Pigs and Miracles: Clinical Innovations and Unorthodox Treatment," takes up ethical issues related to novel and unconventional medical interventions. It includes discussions of a patient's request for amputation, conflicts between surrogate decision-makers regarding the use of herbal therapies, the use of "destination therapy" as a last resort, and finally, the credibility of altruistic organ donation. Offering a new commentary for this section is Kaarkuzhali B. Krishnamurthy.

The final section of the volume, "Part VII: The Big Picture: Organizational Issues," broadens the volume's focus to systemic and institutional level ethical considerations. Cases explore the diffusion of responsibility among staff, the dilemma of whether to disclose an intra-operative exposure to disease, pressures related to patient discharge, and an institution's response to a larger organizational ethics problem. It concludes with a commentary by Ruchika Mishra.

The editors close the volume by highlighting a few substantive conclusions, including the importance of possessing the virtue of humility for clinical ethics consultants, and outlining a set of helpful educational activities and resources. The latter are of great potential benefit and include ways to use the text to study particular themes within bioethics (e.g., legal issues), as well as suggested activities for seminars and the like. The aims of the book – the education and professional development of clinical ethics consultants – is reaffirmed in its conclusion, which further frames what is outlined as a valuable resource to those preparing to sit for the Healthcare Ethics Consultation-Certified examination.

For those who practice or are interested in clinical ethics consultation, this volume is a must read. Bioethicists familiar with the first edition will find much to reflect on in the commentaries that were added in the second, and bioethicists unfamiliar with the first edition will benefit from both the original case discussions, as well as the new commentaries. As the editors note, the original edition is a core resource suggested by the American Society of Bioethics and Humanities (ASBH) for those studying for the HEC-C exam, and the new edition will likely find its way to that list, as well. The book should also be of interest to advanced undergraduates and graduate students in

the humanities as well as the medical sciences interested in the ethically normative aspects of clinical medicine. The case presentations make for excellent conversation starters for classroom discussion, and the chapters are structured well for anyone employing debate pedagogy in a course. The book may also be of interest for those in academic healthcare institutions, such as medical schools and training hospitals that have space in their curricula for more complex ethics conversations having already built an appropriate ethics foundation. In fact, the category of potentially interested parties may be wider than the intended audience, depending on how teachers and students use the text.

Of course, no book review is complete without asking some critical questions of the text. Placing aside debates and questions about what might occur in each of the very interesting cases offered in both editions, the cases are too numerous to assess individually. Choosing a select few seems neither fitting nor fair for a review of the whole text. Focusing instead on what is new in the second edition, a careful reader might ask about the questions that provide structure to each new commentary. As noted, the commentaries all follow the same structure; each of the commentators addresses four specific questions as they reflect on the four cases in their respective sections. First, what are the most powerful lessons that remain relevant in today's practice of ethics consultation? Second, how might the consultations be carried out differently today? Third, how would you appraise the presence or absence of equity, diversity, and inclusion and their impact on the quality of the ethical analyses and reflections? Fourth, how should ethics consultation better address affective components in clinical ethics practice in the future?

While these are excellent questions in terms of bridging the last 17 years, one might consider whether other questions could have been added to the list or even prioritized over these four. Are there lenses additional to those of diversity, equity, and inclusion from which current consultants would benefit, as well as those thinking and teaching about the practice of ethics consultation? For instance: How has the professionalization of ethics consultation impacted cases like this? What could have been done to improve patient-centricity in this set of cases? What resources are (or should be) available to consultants as they engage with this type of case, given what bioethicists know now? Consideration of these additional kinds of questions would add value. In response, it might be noted that there are only so many questions a

commentator can reasonably address in a single essay, and such questions ought to be general enough to be appropriate in response to a diversity of topics housed in each section.

This, however, leads to additional, critical considerations about the diversity of perspectives offered. Since what is new in each of these sections is the commentary of only one individual, one might wonder what others might have to say about those cases. Granted, there are only four cases in each section, having at least two commentaries on each set of cases could enhance the diversity of assessments and thus further strengthen the legitimacy of the project as one intending to bring the volume up to date. In response to this additional consideration, one might argue that sufficient diversity already exists once the reader takes into consideration the additional perspectives offered by the new commentaries themselves within the second volume. This is an important reply, especially given the openness and vulnerability expressed in the discussions of each case. However, since what the second edition brings to the reader is, in part, traversing a chasm of more than a few years, additional perspectives would be welcome, as they bring with them additional normative lenses at work in bioethics which might be of use to readers. Consider, for example, that though the principles of autonomy, nonmaleficence, and justice appear sufficiently in terms of their inclusion in the volume's index, other ethical concepts relevant to clinical consultation, such as dignity and the dignity of risk, as well as to professions and professionalization, such as self-regulation, do not.^{3,4}

A final critical question that must be asked is why the commentary format was chosen as a proper conduit by which to introduce new content for a second edition. Why not simply add new cases? This review opened with praise for certain aspects of this approach, but a careful reader might push for a larger sampling of cases from the current landscape of clinical ethics. If it is indeed the cases, the narratives, and the reflections, that do the normative heavy lifting of the text, why think commentaries (as opposed to more cases) would carry more weight? The response to that question is undoubtedly a "both/and." The second edition, published last summer, was followed by the publication, late last year, of a second volume of cases. That new volume, *Complex Ethics Consultations: Cases That Haunt Us in a New Era*,⁵ is outside the scope of this review, but if it is similar to the second edition of the first volume, it will likely prove to be an extremely helpful, instructive, and, more

importantly, humanizing tool. Toward the latter, reflection on the text's content will raise a host of important questions on a range of topics in need of answers - from the proper role of professionalization in clinical ethics, to the fittingness of various normative lenses for clinical ethics, to practicing allyship with those who aid patients, as well as their care teams and families in navigating many of the murky and difficult decisions associated with caring for those we love.

REFERENCES

1. Ford, Paul J., and Denise M. Dudzinski, eds. *Complex Ethics Consultations. Volume 1: Cases That Haunt Us / Edited by Paul J. Ford, Denise M. Dudzinski*. Cambridge Medicine. Cambridge University Press, 2008.
2. Beauchamp, Tom L., and James F. Childress. *Principles of Biomedical Ethics*. Eighth edition. Oxford University Press, 2019.
3. Kondrat, Andy. "Clinical Ethics Consultation and the Reframing of Risk." *Perspectives in Biology and Medicine* 65, no. 2 (2022): 207–12. <https://doi.org/10.1353/pbm.2022.0016>.
4. Buchanan, Allen. "Is There a Medical Profession in the House?" In *Justice And Health Care*, by Allen Buchanan. Oxford University Press. New York, NY, 2009. <https://doi.org/10.1093/oso/9780195394061.003.0009>.
5. Dudzinski, Denise M., Kaarkuzhali Babu Krishnamurthy, and Paul J. Ford, eds. *Complex Ethics Consultations: Cases That Haunt Us in a New Era*. Cambridge: Cambridge University Press, 2026.

The *Journal of Hospital Ethics* (JoHE) is a peer-reviewed academic journal devoted to interdisciplinary discourse associated with moral complexities and ethical conflicts as they relate both practically and theoretically to the provision of healthcare within hospitals and hospital systems. JoHE's annual volume consists of 3 issues, published in May, September, and December. JoHE is the official proceedings publication for the International Conference on Clinical Ethics and Consultation (<https://www.clinical-ethics.org>), and a partner journal of the American Society for Bioethics & Humanities (<https://asbh.org/resources/journals>). JoHE is published by the John J. Lynch, MD Center for Ethics, MedStar Health, Washington, DC.



EDITORIAL POLICIES AND GUIDELINES

JoHE adheres to the editorial standards and authorship criteria recommended by the International Committee of Medical Journal Editors (ICMJE) and is produced with guidance from the World Association of Medical Editors (WAME). We welcome submissions of completed manuscripts as well as abstract paper proposals by authors with academic or professional expertise in relevant fields. To access the full run of JoHE volumes as well as complete information regarding JoHE's editorial policies and guidelines (Manuscript Submission, Peer-Review Process, Conflicts of Interest, Data Sharing, Policy on the Use of Artificial Intelligence, Errata and Corrigenda, Editorial Disclaimer, Advertising Policy, Indexing, Open Access, Copyright, and ISSN), please scan the QR code or visit: www.medstarwashington.org Click on *About MedStar Washington, The John J. Lynch, MD Center*

for *Ethics*, then *Journal of Hospital Ethics*. Or simply search for JoHE's webpage directly via your preferred search engine. Please send all inquiries to johe@medstar.net.

Manuscript Submission

Please limit manuscript submissions to a range of 4000-6000 words, and abstract paper proposals to 250 words. All manuscripts should be anonymized for peer review and include an abstract of no more than 250 words. Abstracts for original studies (empirical work) should be organized in the following manner: Background, Methods, Results, Conclusions. In addition, submissions should include 3-5 keywords for greater discoverability in search engines, databases, and other indexing systems. Manuscripts should be anonymized for review and include a separate cover letter with the title of the article, name of each author, contact information, institutional affiliation(s), and a statement declaring any conflicts of interest. All submissions should adhere to the American Medical Association Manual of Style (11th Edition). Whereas book reviews are solicited by the Editorial Group, all other submissions to the journal will be reviewed for publication with the understanding that they are not under consideration, accepted for publication, or published elsewhere. To facilitate the review process, all manuscripts should be submitted as MS Word documents and should adhere to JoHE's complete set of editorial policies and guidelines as stated on this page at the time of submission. Please use JoHE's Manuscript Submission Checklist (available online) when submitting a manuscript for review. All manuscript submissions should be sent to Christian Carrozzo christian.carrozzo@medstar.net.

Editorial Disclaimer

The opinions and perspectives offered in JoHE reflect only the views of the authors and editors, and do not represent the views of others within the MedStar Health system or other institutions or organizations with which any of the members of the JoHE Editorial Advisory Board are affiliated.

Open Access, Copyright, and ISSN

JoHE is an open access journal. Given JoHE's institutional funding, its status as "Diamond" open access means no article processing fees for authors nor subscription fees for readers. Submitting authors whose work is accepted for publication in JoHE will never be asked for payment of any kind. As an open access, peer-reviewed journal, JoHE publishes its content under a version of the Creative Commons license that does not allow the material to be used for commercial purposes and prohibits that any modifications/derivations of the material be distributed (CC BY-NC-ND 4.0: <https://creativecommons.org/licenses/by-nc-nd/4.0/>). Open distribution is strictly limited to the published material as it stands without any related commercial charges or content modifications of any kind. Authors of all material accepted for publication will be required to assign copyright to the publisher. The following ISSN assignments have been made under the auspice of the Library of Congress (OCLC: 136972691): JoHE (Print) ISSN 1938-4955, JoHE (Online) ISSN 1938-4920 © Copyright 2012-2025 The John J. Lynch, MD Center for Ethics. All rights reserved.





MedStar Health

John J. Lynch, MD Center for Ethics
MedStar Washington Hospital Center
110 Irving Street, NW
EB 3108
Washington, DC 20010

Non-Profit
Organization

U.S. Postage
PAID

Washington Hospital
Center