

# UCLA

## UCLA Previously Published Works

### Title

Science Without Conscience Is but the Ruin of the Soul: The Ethics of Big Data and Artificial Intelligence in Perioperative Medicine.

### Permalink

<https://escholarship.org/uc/item/5np9w7s5>

### Journal

Anesthesia & Analgesia, 130(5)

### Authors

Canales, Cecilia

Lee, Christine

Cannesson, Maxime

### Publication Date

2020-05-01

### DOI

10.1213/ANE.0000000000004728

Peer reviewed



# HHS Public Access

Author manuscript

*Anesth Analg.* Author manuscript; available in PMC 2020 September 26.

Published in final edited form as:

*Anesth Analg.* 2020 May ; 130(5): 1234–1243. doi:10.1213/ANE.0000000000004728.

## Science Without Conscience Is but the Ruin of the Soul: The Ethics of Big Data and Artificial Intelligence in Perioperative Medicine

Cecilia Canales, MD, MPH\*, Christine Lee, PhD†, Maxime Cannesson, MD, PhD\*

\*Department of Anesthesiology and Perioperative Medicine, University of California Los Angeles (UCLA) David Geffen School of Medicine, Los Angeles, California;

†Department of Biomedical Engineering, University of California, Irvine, Irvine, California.

### Abstract

Artificial intelligence–driven anesthesiology and perioperative care may just be around the corner. However, its promises of improved safety and patient outcomes can only become a reality if we take the time to examine its technical, ethical, and moral implications. The aim of perioperative medicine is to diagnose, treat, and prevent disease. As we introduce new interventions or devices, we must take care to do so with a conscience, keeping patient care as the main objective, and understanding that humanism is a core component of our practice. In our article, we outline key principles of artificial intelligence for the perioperative physician and explore limitations and ethical challenges in the field.

---

Science sans conscience n'est que ruine de l'âme [Science without conscience is but the ruin of the soul].

—From *Pantagruel* by Francois Rabelais

On March 17, 2018, 3 major news organizations published articles detailing how Cambridge Analytica harvested personal information of millions of Facebook profiles without consent and used it for political advertising.<sup>1–3</sup> On November 11, 2019, *The Wall Street Journal* reported on “Project Nightingale” that transferred identifiable health information of millions

---

Address correspondence to Maxime Cannesson, MD, PhD, Department of Anesthesiology and Perioperative Medicine, UCLA David Geffen School of Medicine, 757 Westwood Plaza, Los Angeles, CA 90095. mcannesson@mednet.ucla.edu.

**Name:** Cecilia Canales, MD, MPH.

**Contribution:** This author contributed to the development and writing of the manuscript.

**Name:** Christine Lee, PhD.

**Contribution:** This author contributed to the development and writing of the manuscript.

**Name:** Maxime Cannesson, MD, PhD.

**Contribution:** This author contributed to the development and writing of the manuscript.

**This manuscript was handled by:** Thomas M. Hemmerling, MSc, MD, DEAA.

**Name:** Cecilia Canales, MD, MPH.

**Conflicts of Interest:** None.

**Name:** Christine Lee, PhD.

**Conflicts of Interest:** C. Lee is an employee of Edwards Lifesciences.

**Name:** Maxime Cannesson, MD, PhD.

**Conflicts of Interest:** M. Cannesson is a consultant for Edwards Lifesciences and Masimo Corp and has funded research from Edwards Lifesciences and Masimo Corp. He is also the founder of Sironis, and he owns patents and receives royalties for closed-loop hemodynamic management technologies that have been licensed to Edwards Lifesciences.

of Americans to Google without the knowledge of patients or their providers.<sup>4,5</sup> These events have caused a shift in public understanding of what constitutes personal data, consent for how data are used, and who has access to data. It has made data rights and data privacy part of human rights.<sup>6</sup> The future of artificial intelligence in medicine—which relies on patient data—cannot be successful if it is not guided by ethical and humanistic principles that place patients and their privacy in the center.

It is clear that data analytics and artificial intelligence are transforming health care by making it possible to layer enormous and varied sources of data, such as the electronic health record, physiologic waveforms, “Omics” (genomics, proteomics, and metabolomics, etc), wearable sensors, and by creating solutions to complex health problems. Machine-learning and deep learning algorithms have already been shown to potentially predict disease states better than physicians for interpreting electrocardiograms<sup>7</sup> and screening mammograms for cancer,<sup>8</sup> and the Food and Drug Administration (FDA) has already approved a web-based artificial intelligence system that can detect exact locations of congenital heart anomalies in newborns in approximately 10 minutes.<sup>9</sup> In the health care of the future, autonomous systems may even be able to use such predictive models to automatically deliver care.<sup>9–12</sup> All of this is possible because the electronic health record has transformed how data are stored, accessed, and used, and because systems are becoming more interoperable.<sup>13–15</sup>

This special issue of *Anesthesia & Analgesia* on artificial intelligence in anesthesiology and perioperative and critical care medicine features articles from scientists and clinicians in the field of artificial intelligence who have worked on systems that are accelerating scientific discovery. All of the works presented in this thematic issue of the Journal rely on patient data. But “who owns the data that are driving these scientific discoveries and transforming health care?”

Although artificial intelligence has the potential to improve health outcomes, it also creates ethical challenges that must be identified and mitigated in a way that preserves patient safety, autonomy, and privacy. In other words, we can only leverage these upcoming technologies if we clearly understand their technical, ethical, and societal implications.

Our article aims to outline the importance of understanding not only the principles of artificial intelligence and machine learning, but also the limitations and ethical concerns to clinical anesthesiologists. In the first part, we aim to give a brief overview of potential artificial intelligence solutions while outlining some of the limitations clinical anesthesiologists should understand to effectively work with the technology. We then delve into the ethical concerns and challenges that have to be tackled before any benefits from these systems can be reasonably expected.

## **BUILDING TRUST BETWEEN ARTIFICIAL INTELLIGENCE AND ANESTHESIOLOGISTS**

Headlines in the media may leave perioperative physicians concerned that a robot might take over their job, and patients concerned that they will cease being seen as patients or humans

and will morph into numbers<sup>16</sup>; this existential threat in a dystopian future is unlikely to unfold, but there are potential threats and ethical concerns this advancing technology can introduce. Clinicians, scientists, patients, and policymakers must engage toward demystifying technology, discussing its opportunities and limitations, and have open discourse on the ethical concerns in which this technology can create or exacerbate health disparities. As artificial intelligence advances, clinicians will have to decide whether to ignore or take recommendations from algorithms. We may have to explain to our patients what the risks and limitations are based on the technology influencing our decision making. To feel comfortable using these new systems, we have to be educated on how they work,<sup>17</sup> and more clinicians need to be involved in their development.

### **The “Black Box” Effect**

Artificial intelligence and machine learning can be difficult to trust, because in some cases, the algorithms rely on intricate, difficult-to-understand mathematics to get from data input to the final result, resulting in a black box effect.<sup>18</sup> For example, in deep learning, algorithms have the ability to teach themselves new feature representations of input data at every—hidden—layer, but even their creators cannot always understand what those new features represent, resulting in difficulties to map out the machine decision-making process.<sup>19</sup> In anesthesiology, the concept of a black box is not new, and has been debated at length.<sup>20,21</sup> Should an anesthesiologist be able to explain the signal processing used to derive a pulse oximeter waveform, or even something as basic as how noninvasive blood pressure is measured?<sup>20</sup> How much technological understanding does an anesthesiologist need? In health care, trust in technology is vital because information it provides might have life and death implications.<sup>20</sup> For many of us, we may never fully understand the intricacies of artificial intelligence algorithms.<sup>22</sup> But just as we do not fully understand how anesthetics induce anesthesia, we may not need to understand how artificial intelligence algorithms create a given decision. Trust can be gained by understanding fundamentals of artificial intelligence, seeing it work successfully in real-life situations, and understanding its potential and limits.

### **Medical Education**

Recently, the American Medical Association (AMA) House of Delegates adopted policies in an effort to promote a greater understanding of artificial intelligence for patients, physicians, and other health care providers.<sup>23</sup> Integrating knowledge of artificial intelligence into curriculums of medical school and residency programs should help to safely and ethically introduce artificial intelligence into our daily practice. As artificial intelligence is introduced into our specialty, we must proactively engage in learning opportunities in the same way we have done to stay up to date with emerging therapeutics and evidence-based medicine. In the Table, we provide a nonexhaustive list of artificial intelligence key terms for the clinician.

### **Human Physicians as Part of Artificial Intelligence Teams**

Artificial intelligence is meant to be a complementary technology to enhance the performance of physicians, with the goal to improve outcome, streamline processes, and increase the availability of resources to a growing population. By reducing the physical or computational strength needed to perform a task, artificial intelligence can improve the

capacity of physicians to diagnose and treat disease, as well as better predict outcome. Health care and medicine are much more than the application of rule-based or other more complex algorithms.<sup>24</sup> Integral to health care and the art of medicine are empathetic, ethical, and conscientious care.<sup>24,25</sup> While artificial intelligence and predictive tools can help physicians' accuracy, human intuition is what allows for nuance in medicine when a clinical situation does not fit the standard practice and alternative management is necessary. Edge cases are a prime example of the need for physician oversight of artificial intelligence. Simply put, edge cases are when applications of artificial intelligence encounter a scenario in which the system does not perform as expected. In many cases, this is because the algorithm has never seen such a case before. Training of such complex algorithms in artificial intelligence typically depends on large development data sets that are expected to represent the population. However, if certain cases are very rare or never encountered during training, the algorithm is unlikely to perform as expected. While artificial intelligence and machine learning excel at navigating and computing large amounts of data, humans are able to navigate seamlessly through complex situations they may never have encountered previously. Machine learning can often struggle to translate the complexity and nuance of the real world coded into simple rules. Even as algorithms learn, inevitably, the algorithm will reach a scenario, or edge case, that it simply cannot navigate. A 16-year-old teenager can enroll in a driver education course, have 20 hours of behind-the-wheel training, perform superiorly, and navigate through complex real-world scenarios with much more proficiency than the most robust autonomous driving vehicle available today. Pilots can override automated systems and manually fly or land a plane; physicians must have the ability to override the automated systems when deemed necessary. The goal is to have artificially intelligent teams consisting of automated systems and physicians who will deliver better care to more patients.

## ETHICS OF ARTIFICIAL INTELLIGENCE IN HEALTH CARE

Ethical implications of artificial intelligence can be divided into the following categories: data stewardship, bias, implementation of artificial intelligence solutions, and societal implications (Figure).

### From Data Ownership to Data Stewardship

Companies, hospitals, universities, and researchers have come to understand the value of data. The ability to collect, access, mine, and share data brings up the idea of data stewardship: it is the responsibility to govern and protect the data, determine who can access or share it and how, ensure regulatory practices are met, and, above all, ensure patient privacy. Data stewards must be certain that data are not only compiled, accessible, and retrievable, but also that data are available for researchers. Thus, these researchers can develop new technology through collaborations and data exchange. Data stewards ensure that patient data stay private, and the use of data is compliant with regulatory agencies.

The concept of data stewardship brings up more ethical questions:

1. Who owns health care data?
2. Do patients have the right to opt in or out of data sets?

3. How do health care organization protect patient privacy while maintaining a collaborative environment?
4. Who has access to data and how?

Tech giants (Google, Apple, Facebook, Amazon [GAFA], and others) now see the combination of artificial intelligence and health care as an emerging market and an opportunity for growth. These companies know more and more about our personal health data: they know when a woman is due for her period, they track how much sleep or exercise we get, the quality of the air we breathe, our weight, some of our vital signs, and they are creating personalized health tabs to send reminders or advertisements.<sup>26</sup> They are making contracts with hospitals to have access to hospital data<sup>4</sup>; they are developing or acquiring wearable devices<sup>26</sup> in a commercial effort to develop and implement artificial intelligence. As such, consumers are demanding protections for their data privacy and security. Recently, federated learning has been introduced in an effort to reduce privacy risks.<sup>27</sup> Federated learning is an alternative to the current approach of training machine-learning models at a central location with centralized data. It allows for algorithms to be trained in a decentralizing manner: in doing so, data never leave a device, increasing privacy and data security. One group has already tested the use of federal learning on medical data for the task of predicting in-hospital mortality in intensive care unit (ICU) patients from the Medical Information Mart for Intensive Care (MIMIC) III database.<sup>28,29</sup> In this study, the data were split up into “different hospitals” (aka, devices) to simulate the real-life use of federated learning on large, aggregate medical data sets. While performance of the models trained with federated learning was slightly lower than that of the current standard centralized approach, the study demonstrates the potential of federated learning as a solution for using clinical data while preserving patient privacy.

Patient data should not necessarily be sold to the highest bidder; even with anonymous data, it can be relatively easy to identify individuals, eroding privacy and trust.<sup>30,31</sup> Just because a company has the ability to pay large sums of money for data does not mean it should have access to the data. Data stewards need to weigh what benefits the patient, institution, and scientific community as a whole will get by accessing or sharing the data. Likewise, a company that does not have millions of dollars to purchase data should not be disqualified from having access to data that may lead to scientific discoveries. Similarly, institutions should not parse out data and quantify the “value” of data based on how many patients are included in a data set. The data of a single patient can total terabytes, and the data of thousands of patients may only total megabytes, depending on the source and resolution of the data. The volume, velocity, variety, and veracity of data used for algorithm development make it so that artificial intelligence is not able to lend itself to the clinical trial model in which you charge on a per-patient/subject basis.

As what are considered patient data change and grow,<sup>32</sup> there are players who want to capitalize on the information.<sup>30,33</sup> There is a marked difference between responsible companies and researchers collaborating to further diagnostic and treatment modalities, and taking the concept of precision medicine and corrupting it into precision marketing in an effort to sell a specific product or goods. Clinical anesthesiologists need to take an active and proactive approach to mitigate possible ethical conflicts between patient care and the

generated data. Our data (as either a health care provider or a patient) may be used to shape the medicine of the future. Engaging early in the process is the only way to have a voice and influence the future of medicine.

### **Algorithmic Bias in Artificial Intelligence Can Lead to Inequitable Care**

It is well understood that humans are subject to cognitive errors or biases, which can lead to incorrect diagnoses, treatments, or both.<sup>34</sup> Artificial Intelligence has been regarded as a solution to human error because it can eliminate the human aspect of error and improve patient safety. However, there is a fundamental problem of bias in artificial intelligence: algorithms are built and developed by humans. Implicit or explicit bias is integrated into the code written by humans. Algorithms are being developed and validated on data in health care systems in which the current practice may already be inequitable, and therefore, bias in the health system may be mirrored or exacerbated.

**Data Source.**—Predictive algorithms are only as good as the data on which they are based. Data scientists and engineers know that a robust and varied data set is integral to designing an algorithm for real-world clinical scenarios. The largest source of bias is flawed data sampling, which overrepresents or underrepresents training data by having a lack of equal representation of patients with diverse ethnic or ancestral backgrounds, socioeconomic status, diagnoses, or treatment modalities. Lack of diversity in genomic databases is impeding advances in precision medicine initiatives, with far fewer representation from African, Asian, and Latin American ancestral populations when compared to European populations, which can make it difficult to identify gene–disease relationships and limit translation into diverse patient populations.<sup>35</sup> In the United Kingdom BioBank, the nearly 500,000 participants tend to be healthier, leaner, smoke less, and are less likely to have cancer or heart or kidney disease than the country as whole.<sup>36</sup> Demographically, 94.5% of the participants are ethnically white,<sup>36</sup> leaving many of the Pakistani, Bangladeshi, black, mixed, and other ethnic minorities concerned that the genomic advances from the BioBank will lead to genetic testing or treatment for health conditions that may be applicable only to white populations, while other groups may be offered incorrect testing or no testing at all since sample size alone cannot overcome sampling issues.<sup>37,38</sup>

There is no one BioBank or database that can answer all the health questions in the world today. Algorithms based on isolated, nonrepresentative databases cannot be applied to a population as a whole. To combat this lack of diversity, it is essential to collaborate and make concerted efforts to be more inclusive of underrepresented populations. Scientific articles detailing advances in the field should report demographic information to help understand to what populations their discoveries and algorithms can be applied.

**Model Development.**—Developers have their own view on health, the world, or science: their implicit or explicit racial, ethnic, or social biases are being coded into the technology.<sup>39</sup> There are many examples in technology development that show obvious bias. Joy Buolamwini, a researcher at the Massachusetts Institute of Technology, and others<sup>40</sup> have shown that facial recognition systems have bias in the databases used to create and test their algorithms—which are often composed of light-skinned faces. Also, the programmers are

not representative of minority populations, resulting in systems with limited ability to recognize or distinguish dark-skinned faces.<sup>41</sup> In medicine, similar algorithms are being used by dermatologists to distinguish between benign and malignant skin lesions.<sup>8</sup> Although accurate with fair skin color, the disease can manifest differently in darker skin types, and without inclusive representation, the algorithms tend to underperform with darker shades.<sup>42,43</sup> Yet, these systems and technologies have been deployed without enough oversight or audits of the system and have implications on people of color.

Recently, an algorithm used to manage the health care of millions of Americans was shown to have significant biases against African American patients. The algorithm systematically risk stratified African American patients who were sicker than their Caucasian counterparts as requiring less care.<sup>44</sup> According to the authors, the bias occurred because the model developers used health costs as a proxy for health needs, not considering that in the United States, less money is spent on the health of African American patients.

Algorithm development and implementation has the capacity to take inequalities that exist in health care around the world and amplify them.<sup>45</sup> To correct these types of programming biases, there must be inclusivity in the groups that are developing the models that are affecting patient care. Companies, universities, and principle investigators working in the field must realize that a diverse and inclusive workforce offers diversity of experiences, ethnicities, and backgrounds, which are an asset to help create stronger, more robust algorithms to eventually improve patient outcomes. Differences in experiences allow for alternative views of the world, and when a program is based on a set of rules created by the programmer, having a singular point of view might limit the application to a health care system composed of patients that are diverse in their disease processes, access to health care and resources, and ultimately their genetics. Diverse data sets are necessary but not sufficient to overcome all bias in health care. Data can be representative of unequal health access globally, cultural norms, and resources currently available; therefore, inclusivity in the programming and model development is another layer in addressing bias in artificial intelligence.

**Algorithm Testing and Validation.**—Just as bias can be introduced in the development of an algorithm, it can also be introduced when testing the algorithm. Creation of machine-learning algorithms and artificial intelligence is much more than simply writing code. Developers help to model a view of the world through the set of rules they create. Testing the algorithms in different and diverse environments helps the algorithms to become more inclusive and more predictive for a wider range of patient population. Understanding where these systems fail helps to understand possible biases in the code and also identifies the limits in which the algorithm is viable. Testing in diverse environments, with diverse data, is essential to identifying bias within the system.

**Real-World Application of Artificial Intelligence.**—Real-world application of artificial intelligence systems is both an opportunity to introduce new bias and identify bias. Bias in clinical culture, training, and how we manage or treat patients is important to consider with widespread introduction of artificial intelligence systems. The heterogeneity or homogeneity of any given condition can affect how an artificial intelligence system



diagnoses or treats a disease. For example, although breast cancer is a problem globally, in sub-Saharan Africa, women tend to be diagnosed with breast cancer younger and their disease tends to be more advanced with poorer prognosis than women in high-income countries. Artificial intelligence diagnostic tools that have been trained on mammograms of high-income countries are of low benefit for women in sub-Saharan Africa because younger women have increased breast density and mammograms are not appropriate for detecting tumors at advanced stages.<sup>46</sup> In addition, how one institution treats or manages breast cancer in the United States may be different from a medical center in Europe, Asia, or sub-Saharan Africa.

Further, clinical workflows are optimized differently in different institutions, there is variety from clinician to clinician, and there is even bias on how data are recorded. For example, one institution might annotate blood gases at the time of results versus at the time of blood draw. Or one electronic health record may allow for nonbinary gender, while another allows only for binary. Clinical care and workflows in the operating room for something as routine as total intravenous anesthesia (TIVA) techniques vary greatly from country to country and institution to institution. In Europe, Asia, and South America, target-controlled infusion (TCI) uses patient characteristics such as age, sex, or creatinine clearance, along with pharmacokinetic principles, to achieve a desired drug concentration, while in the United States, this technology never received regulatory approval, and TIVA is achieved through pumps that only take weight into account.<sup>47</sup> Testing for bias in an application can be costly and time-consuming, requiring specialized code to detect what the inflection points of the algorithm are. But this testing is necessary to ensure that bias does not preclude high-quality patient care.

### Artificial Intelligence Solutions

Implementation of an artificially intelligent system raises ethical questions about safety, whether or not patients are informed or not, and who is responsible for the technology and liable for failure. Artificial intelligence and machine learning in software considered as a Medical Device is an evolving field for regulatory agencies. The FDA and International Medical Device Regulators Forum have proposed regulatory frameworks for modifications to artificial intelligence and machine learning–based software as medical devices to ensure that safe and effective technologies reach the patient.<sup>48</sup> But the reality is that navigating the field is a moving target and regulators are trying to catch up with the science following that is evolving at Moore's law: an observation and projection originally applied to the density of circuits doubled about every 2 years, while the cost of producing the circuits halved.<sup>49</sup>

Informed consent for the use of artificial intelligence raises additional ethical concerns. When we look at traditional devices or treatment modalities, we do not go through specific consents for the use of one anesthesia machine versus another, one monitoring device versus another, but we do consent for special procedures such as arterial line or central line placement. Our surgical colleagues consent for robotic versus laparoscopic versus open procedures.

**Should We Consent Patients for the Use of Artificial Intelligence?**—Our patients are becoming older and sicker and are undergoing riskier and more extensive procedures, yet the expectation is for improved patient outcomes. Artificial intelligence may allow us to risk-stratify patients further and determine which patients are likely to have poor outcomes. If algorithms support physicians in their decision making, who is responsible for artificial intelligence, especially when things go wrong? As the Swiss Cheese Model of failure within complex organizations in fields such as medicine exemplifies,<sup>50</sup> assigning responsibility or liability can be difficult even without the additional layer of artificial intelligence. Recently, Schiff and Borenstein<sup>51</sup> outlined key players that could be ethically responsible for a medical error in an artificially intelligent team: coders, medical device companies, physicians and other health care professionals, hospitals and health care entities, regulators, medical schools, and even insurance companies. Yet, legally, it may be difficult to assign liability, intent, or causation, very human principles of the law, and apply them to artificial intelligence, especially when there is lack of transparency in a system.<sup>52</sup> Currently, most artificially intelligent solutions are focused on clinician support tools instead of replacing clinical judgment<sup>53</sup> which implies responsibility rests with the clinician. Clinicians will be left with the decision to follow or not follow a recommendation from an algorithm or device. However, as standards of care evolve and change to integrate the technology, clinicians are likely to face automation bias, a tendency to trust a machine more than they may trust themselves. Liability for medical errors typically falls under negligence, organizational liability, and product liability.<sup>54</sup> As artificial intelligence technology becomes more autonomous, it will be difficult to foresee how the system will react, may cause unforeseen consequences, and will introduce new risks for patients. Current legal doctrine assigns liability of negligence based on what a reasonable physician would do under similar circumstances.<sup>54</sup> This standard will probably evolve as new techniques and technologies are introduced into clinical care.

### Implications for Society

We are taking data from millions of patients, looking at health records, at patients' habits, and ingesting varied sources of data, creating learning algorithms to predict who will develop diabetes, heart disease, and other diseases. Will insurance companies be able to use these predictive algorithms to deny or alter benefits? Will there be protections against projected diseases that have yet to or may never develop? Policymakers need to proactively work with the health care industry to protect patient rights, to protect patient privacy, and to protect patients from discrimination for disease they have not yet developed.

In the United States, we already have a disparity in terms of access and outcomes for patients in different communities.<sup>55–58</sup> We need to ensure that artificial intelligence is applied equitably across communities and in health care systems that will reach all patient populations, not further limit it. We must ensure that the technologies are not disproportionately tested in one community versus another. Systems can fail, learning algorithms can encounter cases not previously seen, and errors can occur. Therefore, constant monitoring of proper functioning of artificial technologies in an evolving population is a key objective.

## SUMMARY

Artificial intelligence–driven health care may be just around the corner. Artificial intelligence promises improved safety and patient outcomes. This can only occur if we take the time to examine the technical, ethical, and moral implications of artificial intelligence. When introducing artificial intelligence, we must take care to do so with a conscience, keeping the patient first, understanding that humanism is a core component of our practice, while being open to innovation.

## Funding:

This work is supported by the National Institutes of Health (R01 HL144692).

## GLOSSARY

<b>AMA</b>	American Medical Association
<b>EU</b>	European Union
<b>FDA</b>	Food and Drug Administration
<b>GAFA</b>	Google, Apple, Facebook, Amazon
<b>ICU</b>	intensive care unit
<b>MIMIC-III</b>	Medical Information Mart for Intensive Care
<b>TCI</b>	target-controlled infusion
<b>TIVA</b>	total intravenous anesthesia

## REFERENCES

1. Rosenberg M, Confessore N, Cadwalladr C. How Trump Consultants Exploited the Facebook Data of Millions. *The New York Times*. Available at: <https://www.nytimes.com/2018/03/17/us/politics/cambridge-analytica-trump-campaign.html>. Accessed December 22, 2019.
2. Cadwalladr C, Graham-Harrison E. Revealed: 50 million Facebook profiles harvested for Cambridge Analytica in major data breach. *The Guardian*. Available at: <https://www.theguardian.com/news/2018/mar/17/cambridge-analytica-facebook-influence-us-election>. Accessed December 20, 2019.
3. Seetharaman D Facebook Suspends Data Firm that Helped Trump Campaign. *The Wall Street Journal*. Available at: <https://www.wsj.com/articles/facebook-suspends-cambridge-analytica-for-failing-to-delete-user-data-1521260051>. Accessed December 17, 2019.
4. Copeland R. Google's 'Project Nightingale' Gather Personal Health Data of Millions of Americans. *The Wall Street Journal*. Available at: <https://www.wsj.com/articles/google-s-secret-project-nightingale-gathers-personal-health-data-on-millions-of-americans-11573496790>. Accessed January 14, 2020.
5. Pilkington E Google's secret cache of medical data includes names and full details of millions-whistleblower. *The Guardian*. Available at: <https://www.theguardian.com/technology/2019/nov/12/google-medical-data-project-nightingale-secret-transfer-us-health-information>. Accessed January 14, 2020.

6. Galvin M Human Rights in Age of Social Media, Big Data, and AI. The National Academies of Science Engineering Medicine. Available at: <http://www8.nationalacademies.org/onpinews/newsitem.aspx?RecordID=9302019>. Accessed December 20, 2019.
7. Hannun AY, Rajpurkar P, Haghpanahi M, et al. Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network. *Nat Med.* 2019;25:65–69. [PubMed: 30617320]
8. McKinney SM, Sieniek M, Godbole V, et al. International evaluation of an AI system for breast cancer screening. *Nature.* 2020;577:89–94. [PubMed: 31894144]
9. Marr B First FDA Approval For Clinical Cloud-Based Deep Learning In Healthcare. *Forbes.* Available at: <https://www.forbes.com/sites/bernardmarr/2017/01/20/first-fda-approval-for-clinical-cloud-based-deep-learning-inhealthcare/#63e0d15c161c>. Accessed December 20, 2019.
10. Hemmerling TM. Robots will perform anesthesia in the near future. *Anesthesiology.* 2020;132:219–220. [PubMed: 31939836]
11. Joosten A, Rinehart J, Bardaji A, et al. Anesthetic management using multiple closed-loop systems and delayed neurocognitive recovery: a randomized controlled trial. *Anesthesiology.* 2020;132:253–266. [PubMed: 31939839]
12. Nair BG, Gabel E, Hofer I, Schwid HA, Cannesson M. Intraoperative clinical decision support for anesthesia: a narrative review of available systems. *Anesth Analg.* 2017;124:603–617. [PubMed: 28099325]
13. Goldman JM, Weininger S, Jaffe MB. Applying medical device informatics to enable safe and secure interoperable systems: medical device interface data sheets. *Anesth Analg.* 2019 [Epub ahead of print].
14. Weininger S, Jaffe MB, Goldman JM. The need to apply medical device informatics in developing standards for safe interoperable medical systems. *Anesth Analg.* 2017;124:127–135. [PubMed: 27584685]
15. Weininger S, Jaffe MB, Rausch T, Goldman JM. Capturing essential information to achieve safe interoperability. *Anesth Analg.* 2017;124:83–94. [PubMed: 27387840]
16. Stark H Prepare Yourself, Robots Will Soon Replace Doctors in Healthcare. *Forbes.* Available at: <https://www.forbes.com/sites/haroldstark/2017/07/10/prepare-yourself-robots-will-soon-replace-doctors-inhealthcare/#5511e95f52b5>. Accessed December 20, 2019.
17. Cannesson M, Rice MJ. Insight into our technologies: a new series of manuscripts in *Anesthesia & Analgesia*. *Anesth Analg.* 2018;126:25–26. [PubMed: 29252481]
18. Connor CW. Artificial intelligence and machine learning in anesthesiology. *Anesthesiology.* 2019;131:1346–1359. [PubMed: 30973516]
19. Shahid N, Rappon T, Berta W. Applications of artificial neural networks in health care organizational decision-making: a scoping review. *PLoS One.* 2019;14:e0212356. [PubMed: 30779785]
20. Cannesson M, Shafer SL. All boxes are black. *Anesth Analg.* 2016;122:309–317. [PubMed: 26797548]
21. Shelley KH, Barker SJ. Disclosures, what is necessary and sufficient? *Anesth Analg.* 2016;122:307–308. [PubMed: 26797547]
22. Yu KH, Beam AL, Kohane IS. Artificial intelligence in healthcare. *Nat Biomed Eng.* 2018;2:719–731. [PubMed: 31015651]
23. Augmented Intelligence in Health Care. American Medical Association. Available at: <https://www.ama-assn.org/system/files/2019-08/ai-2018-board-policy-summary.pdf>. Accessed December 20, 2019.
24. Verghese A, Shah NH, Harrington RA. What this computer needs is a physician: humanism and artificial intelligence. *JAMA.* 2018;319:19–20. [PubMed: 29261830]
25. Canales C, Strom S, Anderson CT, et al. Humanistic medicine in anaesthesiology: development and assessment of a curriculum in humanism for postgraduate anaesthesiology trainees. *Br J Anaesth.* 2019;123:887–897. [PubMed: 31591019]
26. Kelly ML. How much should big tech know about our personal health data and history? National Public Radio. Available at: <https://www.npr.org/2019/11/25/782732974/how-much-should-big-tech-know-about-our-personal-health-data-and-history>. Accessed December 20, 2019.

27. Konecny J, McMahan HB, Ramage D. Federated optimization: distributed machine learning for on-device intelligence. ArXiv. 2016;1610.02527.
28. Johnson AE, Pollard TJ, Shen L, et al. MIMIC-III, a freely accessible critical care database. Sci Data. 2016;3:160035. [PubMed: 27219127]
29. Sharma P, Shamout FE, Clifton DA. Preserving patient privacy while training a predictive model of in-hospital mortality. ArXiv. 2019;1912.00354.
30. Tanner A For sale: your medical records. Sci Am. 2016;314:26–27.
31. Kalra D, Gertz R, Singleton P, Inskip HM. Confidentiality of personal health information used for research. BMJ. 2006;333:196–198. [PubMed: 16858053]
32. Haendel MA, Vasilevsky NA, Wirz JA. Dealing with data: a case study on information and data management literacy. PLoS Biol. 2012;10:e1001339. [PubMed: 22666180]
33. Kolata G When patients' records are commodities for sale. N Y Times Web. 1995;A1:C14.
34. Stiegler MP, Neelankavil JP, Canales C, Dhillion A. Cognitive errors detected in anaesthesiology: a literature review and pilot study. Br J Anaesth. 2012;108:229–235. [PubMed: 22157846]
35. Landry LG, Ali N, Williams DR, Rehm HL, Bonham VL. Lack of diversity in genomic databases is a barrier to translating precision medicine research into practice. Health Aff (Millwood). 2018;37:780–785. [PubMed: 29733732]
36. Fry A, Littlejohns TJ, Sudlow C, et al. Comparison of sociodemographic and health-related characteristics of UK Biobank participants with those of the general population. Am J Epidemiol. 2017;186:1026–1034. [PubMed: 28641372]
37. Muir R Huge genetic databases are hurting marginalized people's health. Massive Science. Available at: <https://massivesci.com/articles/uk-biobank-promise-genetic-association-bias-underrepresentation-minority-groups/>. Accessed January 3, 2020.
38. Keyes KM, Westreich D. UK Biobank, big data, and the consequences of non-representativeness. Lancet. 2019; 393:1297.
39. Nelson GS. Bias in artificial intelligence. N C Med J. 2019;80:220–222. [PubMed: 31278182]
40. Zou J, Schiebinger L. AI can be sexist and racist - it's time to make it fair. Nature. 2018;559:324–326. [PubMed: 30018439]
41. Buolamwini J Gender shades: intersectional phenotypic and demographic evaluation of face datasets and gender classifiers center for civic media: MIT, 2017 Available at: <https://www.media.mit.edu/publications/full-gender-shades-thesis-17/>. Accessed January 3, 2020.
42. Adamson AS, Smith A. Machine learning and health care disparities in dermatology. JAMA Dermatol. 2018;154:1247–1248. [PubMed: 30073260]
43. Brinker TJ, Hekler A, Utikal JS, et al. Skin cancer classification using convolutional neural networks: systematic review. J Med Internet Res. 2018;20:e11936. [PubMed: 30333097]
44. Obermeyer Z, Powers B, Vogeli C, Mullainathan S. Dissecting racial bias in an algorithm used to manage the health of populations. Science. 2019;366:447–453. [PubMed: 31649194]
45. Panch T, Mattie H, Atun R. Artificial intelligence and algorithmic bias: implications for health systems. J Glob Health. 2019;9:010318.
46. Black E, Richmond R. Improving early detection of breast cancer in sub-Saharan Africa: why mammography may not be the way forward. Global Health. 2019;15:3. [PubMed: 30621753]
47. Struys MM, De Smet T, Glen JI, Vereecke HE, Absalom AR, Schnider TW. The history of target-controlled infusion. Anesth Analg. 2016;122:56–69. [PubMed: 26516804]
48. FDA. Artificial Intelligence and Machine Learning in Software as a Medical Device. Available at: <https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-software-medical-device>. Accessed January 3, 2020.
49. Puiu T AI is outpacing Moore's Law. ZME Science. Available at: <https://www.zmescience.com/science/ai-isoutpacing-moores-law/>. Accessed January 3, 2020.
50. Reason J Human error: models and management. BMJ. 2000;320:768–770. [PubMed: 10720363]
51. Schiff D, Borenstein J. How should clinicians communicate with patients about the roles of artificially intelligent team members? AMA J Ethics. 2019;21:E138–E145. [PubMed: 30794123]
52. Bathaee Y the artificial intelligence black box and the failure of intent and causation. Harv J Law Technol. 2018;31:890–934.

53. Challen R, Denny J, Pitt M, Gompels L, Edwards T, Tsaneva-Atanasova K. Artificial intelligence, bias and clinical safety. *BMJ Qual Saf.* 2019;28:231–237.
54. Sullivan HR, Schweikart SJ. Are current tort liability doctrines adequate for addressing injury caused by AI? *AMA J Ethics.* 2019;21:E160–E166. [PubMed: 30794126]
55. Bauchner H Health care in the United States: a right or a privilege. *JAMA.* 2017;317:29. [PubMed: 28030681]
56. Chetty R, Stepner M, Abraham S, et al. The association between income and life expectancy in the United States, 2001–2014. *JAMA.* 2016;315:1750–1766. [PubMed: 27063997]
57. Douthit N, Kiv S, Dwolatzky T, Biswas S. Exposing some important barriers to health care access in the rural USA. *Public Health.* 2015;129:611–620. [PubMed: 26025176]
58. Feagin J Systemic racism and “race” categorization in US medical research and practice. *Am J Bioeth.* 2017;17: 54–56.



**Data Stewardship**

Healthcare organization collect data as standard patient care. Data is then accessed by researchers or sourced to companies .

*Do healthcare organization protect patient privacy?*

*Do patient have the right to opt in/out?*

*Does the dataset include sensitive information?*

*Who has access to the data? Sharing of data?*

*Who owns the data?*



**Bias**

Are Artificial Intelligence and Machine Learning fair?

*Are implicit or explicit bias integrated into the code?*

*Does the training data set represent the target population?*

*Is there bias in how data gets annotated?*

*Are socioeconomic or ethnic bias integrated into the system?*

*Is there insitutional, training, or treatment bias?*



**AI Implementation**

Integrating AI in to heathcare

*What evidence supports the safety and efficacy of the algorithm?*

*Have stakeholders been consulted about the introduction of AI into clinical practice?*

*Are patients informed about the use of the system?*

*Are there override safety features developed into the system?*

*How do we ensure humane treatment with AI?*



**Societal Implication**

What are the consequences of implementing the system?

*Do patients have equitable access ?*

*Who is liable for AI if there is a mistake?*

*How is wealth created from AI distributed?*

*What monitoring mechanisms are in place to assess in real life settings?*

*If AI predicts a disease, will patients have protections against discrimination?*

**Figure.**

Ethical questions in AI-dependent health care system. Ethical questions to consider in an AI-dependent health care system, in 4 categories: data stewardship, bias, artificial intelligence implementation, and societal implications. AI indicates artificial intelligence.

Glossary of Terms

Term	Definition
Adaptive algorithm	Any continuous learning algorithm that changes its behavior as it is run, based on predefined reward criteria and new information. With adaptation, the output of the algorithm may be different from what would have been expected before the algorithm changed with the same set of inputs.
Algorithmic bias	The concept that describes systematic and repeatable errors in a computer system that create unfair outcomes when an algorithm is applied.
Artificial intelligence	The broad concept of machines designed to understand and perform tasks on their own in a "smart" manner.
Augmented intelligence	An alternative conceptualization of artificial intelligence focusing on assisting and advancing human capabilities rather than replacing humans.
Automated complacency	In opposition to situational awareness, complacency is a self-satisfaction based on an unjustified assumption of satisfactory system state. Automated complacency refers to the propensity for humans to favor suggestions from automated decision support and to ignore contradictory information made without automation, even if correct.
Clinical decision support	A broad term for any software that provides clinically useful information that is presented in an intelligent manner with the goal of enhancing patient care. Common types include but are not limited to alerts, checklists, and diagnostic support.
Passive	Passive clinical decision support includes features that do not require a response, such as access to knowledge sources, context-sensitive documentation (ie, order sets, patient data reports, and documentation templates).
Active	Active clinical decision support includes features that require a response by the user, such as alerts, and reminders.
Closed-loop system	A system that measures, monitors, and controls a process and is capable of automatically adjusting itself through feedback to obtain a desired response. An example would be an artificial pancreas with closed-loop insulin delivery.
Computer vision	A field of study that focuses on training computers to see and understand images. Models typically focus on image classification, object detection, and facial recognition.
Data ownership	Refers to both the possession of and responsibility for information with the ability to control access, create, modify, package, derive benefit from, sell or remove data, and the right to assign these access privileges to others.
Data privacy	The relationship between the collection, storage, and dissemination of data includes consent, notice, and regulatory obligations that often revolve around how data are collected or stored, and whether or how the data are shared with third parties.
Data security	The process by which data are protected from intentional or accidental destruction, modification, or corruption of files. It also refers to the set of standards or controls to prevent unauthorized access or other threats.
Data stewardship	The practice of management and oversight of the data enterprise of an organization through established data governance practices focused on maintaining data fitness, accessibility, safety, privacy, and ethical use.
Error of commission	An error that occurs as the result of an action. Doing the wrong thing or in the wrong situation.
Error of omission	An error that occurred because of inaction. Failing to act.
Deep neural networks	Also referred to as deep learning, deep neural networks are those that have many hidden layers between the input and output, and many neurons per each hidden layer. There are many types of deep neural networks, of which the most commonly referred to are described below.
Feed-forward neural networks	The simplest form of a neural network is a fully connected feed-forward network. It is also sometimes referred to as a "vanilla" network, and is described as fully connected because every neuron in one layer connects to every neuron in the next layer, and feed-forward because all information flows from input to output with no feedback connections.
Recurrent neural networks	Recurrent neural networks utilize a "hidden" state, or a feedback connection, to represent memory of previous inputs and outputs. They are typically used for sequential problems, such as learning from a time series or sentence structure. A common variation of recurrent neural networks are long short-term memory networks.
Convolutional neural networks	Convolutional neural networks utilize convolutional layers in which the inputs are convolved into a form that is easier to use without losing important information. They are typically used for spatial and temporal problems, such as images and computer vision tasks. Within each convolutional layer, a filter moves along the input and "convolves"



<b>Term</b>	<b>Definition</b>
Federated learning	with multiplication. These convolutional layers are typically followed with pooling layers to reduce dimensionality and extract dominant features, and finally a feed-forward neural network. Also known as collaborative learning, it is a machine-learning technique that trains an algorithm without exchanging data. It enables multiple collaborators to build a common, robust machine-learning model without sharing data, thus addressing critical issues such as data privacy, security, and access rights, and access to heterogeneous data.
Interpretability	The ability to explain or to present the decisions of a machine-learning model in understandable terms, with the goal of providing insights. These are typically extracted as the importance of input features and/or their effect on the predicted output.
Machine learning	Any approach in which machines are given data and are allowed to learn from that data.
Supervised	Supervised learning refers to applications in which the output label/target is known before training. Supervised learning is referred to as classification when the desired output is discrete categories and regression when the desired output is continuous variables.
Unsupervised	Unsupervised learning refers to applications in which the output label/target is unknown. A popular form of unsupervised learning is clustering, in which the goal is to discover new groupings within the input data.
Reinforcement	Reinforcement learning refers to applications in which the goal of the algorithm is to find the action that maximizes a reward. Through trial and error, the algorithm learns how its current action affects its current reward and subsequent rewards, and chooses actions accordingly.
Right to access	A fundamental right afforded in some countries (ie, EU) toward data protection informing individuals of their right to find out whether an organization is using or storing personal data, and can exercise the right by asking for a copy of the data.
Risk homeostasis	The theory that people adjust their behavior in response to the perceived level of risk, suggesting that the safer something becomes, the more risk they are willing to take, resulting in an overall equilibrium. For example, as a procedure becomes safer, the procedure is performed on sicker or less stable patients, resulting in an overall stable risk profile.

---

Abbreviation: EU, European Union.