THE PRICE OF AUTONOMY: LIABILITY STANDARDS FOR COMPLEMENTARY AND SUBSTITUTIVE MEDICAL ROBOTICS AND ARTIFICIAL INTELLIGENCE

Frank Anthony Pasquale III*
Brooklyn Law School

Abstract
When AI or robotics assist a professional, they are tools. In medicine, the doctrine of “competent human intervention” has shifted liability away from those who make devices and toward the professionals who use them. However, the professional in such scenarios should not bear the entire burden of responsibility. Tools can be defective, and vendors of defective, complementary AI and robotics should be held responsible for negligence. The burden of proof will still be on the plaintiff to demonstrate that not only a skilled medical professional, but also the maker of the tools used by such a professional, should be held liable for a preventable adverse outcome.

When AI and robotics replace, rather than merely assist, a skilled medical professional, the burden should shift. The vendor of such computational systems needs to take on responsibility for errors and accidents. In the medical field, there has long been a standard of competent professional supervision of the deployment of advanced technology. When substitutive automation short-circuits that review, it is both defective and unreasonably dangerous. Nevertheless, at the damages phase of litigation, the vendor of the substitutive AI should be entitled to explain how damages should be mitigated based on its AI’s performance relative to the extant human- or human-machine based standard of care. Such responsibility for explanation will serve an important information-forcing function in areas where public understanding is often limited by trade secrecy.

As law and political economy methods demonstrate, law cannot be neutral with respect to markets for new technology. It constructs these markets, making certain futures more or less likely. Distinguishing between technology that substitutes for human expertise and that which complements professionals is fundamental not just to labor policy and the political economy of automation, but also to tort law.

Keywords
artificial intelligence, liability, tort.

Resumen
Cuando la IA o la robótica ayudan a un profesional, son herramientas. En medicina, la doctrina de la “intervención humana competente” ha desplazado la responsabilidad de quienes fabrican dispositivos hacia los profesionales que los utilizan. Sin embargo, el profesional en tales escenarios no debe cargar con todo el peso de la responsabilidad. Las herramientas pueden ser defectuosas, y los proveedores de IA y robótica complementaria defectuosa deben ser considerados responsables por negligencia. La carga de la prueba aún recaerá en el demandante para demostrar que no solo el profesional médico capacitado, sino también

---

* Frank Anthony Pasquale III is an expert on the law of artificial intelligence (AI), algorithms, and machine learning. He is a Professor of Law at Brooklyn Law School, Brooklyn, U.S.A, a Visiting Scholar at the AI Now Institute, an Affiliate Fellow at Yale University’s Information Society Project, and a member of the American Law Institute. Before coming to Brooklyn Law, he was Piper & Marbury Professor of Law at the University of Maryland. He is co-editor-in-chief of the Journal of Cross-Disciplinary Research in Computational Law (CRCL), based in the Netherlands, and a member of an Australian Research Council (ARC) Centre of Excellence on Automated Decision-Making & Society (ADM+S). His book The Black Box Society: The Secret Algorithms That Control Money and Information (Harvard University Press 2015) has been recognized internationally as an important study of the law and political economy of information. His latest book, New Laws of Robotics: Defending Human Expertise in the Age of AI (Harvard University Press 2020) develops a political economy of automation focused on professionalization, in which human capacities are the irreplaceable center of an inclusive economy. Professor of Law, Brooklyn Law School. Email: pasquale.frank@gmail.com. I wish to thank Prof. Dr. Carolina Riveros Ferrada for the invitation to publish this piece. I wish to thank Brooklyn Law School’s summer research fund for supporting this research. I also wish to thank anonymous reviewers and Prof. Anita Bernstein for thoughtful comments. Any errors remain my own responsibility. I also wish to note that this essay was commissioned by the Academia Sinica Law Journal, and is planned to be published there, as well as in a collection on AI and Robotics law.
el fabricante de las herramientas utilizadas por dicho profesional, debe ser considerado responsable de un resultado adverso prevenible.

Cuando la IA y la robótica reemplacen, en lugar de simplemente ayudar, a un profesional médico calificado, la carga debería cambiar. El proveedor de dichos sistemas computacionales debe asumir la responsabilidad por errores y accidentes. En el campo de la medicina, ha existido durante mucho tiempo un estándar de supervisión profesional competente del despliegue de tecnología avanzada. Cuando la automatización sustitutiva cortocircuita esa revisión, es defectuosa e irrazonablemente peligrosa. Sin embargo, en la fase de daños del litigio, el proveedor de la IA sustitutiva debe tener derecho a explicar cómo se deben mitigar los daños en función del desempeño de su IA en relación con el estándar de atención existente basado en humanos o máquinas humanas. Tal responsabilidad de explicación cumplirá una importante función de obtención de información en áreas donde la comprensión del público a menudo se ve limitada por el secreto comercial.

Como demuestran los métodos del derecho y la economía política, el derecho no puede ser neutral con respecto a los mercados de nuevas tecnologías. Construye estos mercados, haciendo ciertos futuros más o menos probables. Distinguir entre la tecnología que sustituye a la experiencia humana y la que complementa a los profesionales es fundamental no solo para la política laboral y la economía política de la automatización, sino también para la responsabilidad civil extracontractual.

Palabras clave
inteeligencia artificial, responsabilidad, derecho delictivo.

1. Introduction

Robotics and AI in medicine raise critical liability questions for the medical profession. Consider the case of robotically assistive surgical devices (RASDs) which surgeons use to control small cutting and grasping devices. If a surgeon’s hand slips with a scalpel, and a vital tendon is cut, our intuitive sense is that the surgeon bears the primary responsibility for the resultant malpractice suit. But the vendor of an RASD may eventually market a machine which has a special “tendon avoidance subroutine,” akin to the alarms that automobiles now sound when their sensors indicate a likely collision. If the tendon sensors fail, and the warning does not sound before an errant cut is made, may the harmed patient sue the vendor of the RASD? Or only the physician who relied on it?

Similar problems arise in the context of some therapy apps. For example, a counselor may tell a patient with substance use disorder (SUD) to use an app in order to track cravings, states of mind, and other information helpful to those trying to cure addictions. The app may recommend certain actions in case the counselor cannot be reached. If these actions are contraindicated and result in harm to the patient or others, is the app to blame? Or the doctor who prescribed it? Home health aide businesses may encounter similar dilemmas as they deploy so-called “care robots”\(^1\).

Of course, in neither the surgical nor the mental health scenario is the answer necessarily binary. There may be shared liability, based on an apportionment of responsibility. But before courts can trigger such an apportionment, they must have a clear theory upon which to base the responsibility of vendors of technology.

This article develops such an approach. What is offered here is less a detailed blueprint for liability determinations than a binary approach to structure policy discussions on liability for harm caused by AI and robotics in medical contexts\(^2\). The binary is the distinction between substitutive and complementary automation\(^3\). When AI and robotics substitutes for a physician,
strict liability is more appropriate than standard negligence doctrine. When the same technology merely assists a professional, a less stringent standard is appropriate. Such standards will help ensure that the deployment of advanced medical technologies is accomplished in a way that complements extant professionals’ skills, while promoting patient safety.

As law and political economy methods demonstrate, law cannot be neutral with respect to markets for new technology. It constructs these markets, making certain futures more or less likely. Distinguishing between technology that substitutes for human expertise and that which complements professionals is fundamental to both labor policy and the political economy of automation.

For example, in the case of computerized physician order entry (CPOE) for prescriptions, a “drug-drug interaction” alert (DDI) could simply warn a physician about possible side effects from simultaneous ingestion of two pills. That is complementary automation. If the DDI alert were, in fact, incorrect, a harmed patient could sue both the physician and the vendor of the CPOE system, but the burden should be on the patient to demonstrate the CPOE system’s vendor failed to follow the proper standard of care in updating data or improving algorithms in order to avoid the problem. And the physician might still bear all or most of the responsibility, under the doctrine of competent human intervention.

By contrast, some CPOE systems of the future may simply “decide everything” with respect to the prescription of the two pills, preventing the doctor from prescribing them together. In such a scenario, the physician is no longer responsible—she or he cannot override the system. Given this extraordinary deviation from ordinary professional standards in medicine—which require a skilled person to mediate between technology and the patient—it is appropriate to impose strict liability up and down the distribution chain of such a substitutive AI. Under a strict liability standard, in case of a preventable adverse event, the manufacturer, distributor, and retailer of the product may be liable, even if they were not at fault.

This may seem like an unduly harsh standard. However, the doctrine of strict liability arose in response to those who sold “any product in a defective condition unreasonably dangerous to the user or consumer or to his property”. In the medical field, there has long been a standard of competent professional supervision and monitoring of the deployment of advanced technology. When substitutive automation short-circuits that review, it is both defective and unreasonably dangerous. It also tends toward the diminution of the distributed expertise so critical to medical advance.

This article develops the complementary and substitutive categories via two case studies. Part 2 explores the complementary role of robotically-assistive surgical devices (RASDs), and some litigation that has arisen regarding them. Part 3 introduces substitutive AI and robotics, and demonstrates the ways in which strict liability standards are likely necessary to promote accountability in their development and deployment, to preserve a unitary standard of care, and to promote public awareness of their shortcomings. Part 4 concludes with reflections on the current utility of, and potential challenges to, the substitutive/complementary dichotomy.

4 MCCLUSKEY et al. (2016).
5 For a good typology of potential scenarios arising in the context of assistive AI, see generally PRICE II et al. (2019).
6 Restatement (Second) of Torts, § 402a.
7 This article addresses policymakers governing health systems with this standard of care. Those in charge of less developed health systems (coping with physician or other staff shortages) may well decide that strict liability is too harsh a standard: if there is no viable human alternative, why discourage direct access to a machine? Those in many of the more developed health systems need to acknowledge their own responsibility for this state of affairs. See PASQUALE (2010a) (describing direct and indirect ways in which medical resources are directed away from the developing and toward the developed world).
8 For an extended argument for the ideal of distributed expertise, see generally PASQUALE (2020).
2. A Negligence Standard for Complementary Robotics

Prostate surgery has seen rapid adoption of robotics with over 80% of the surgeries performed robotically. Hundreds of urological surgeons have adopted the Da Vinci Surgical Robot over the past decade. The rapid adoption of the robotically-assistive surgical devices (RASDs) in prostate surgery demonstrates just how fast a new machine can change the face of practice for hundreds of thousands of patients. Surgical robots are now spreading to head and neck, heart, and thoracic surgery departments.

At the outset, it is important to be clear about the terminology and effects of machines like the da Vinci robot. The device itself does not complete the surgery. Rather, it is an extremely sophisticated tool deployed by a skilled surgeon. The surgeon operates at a console, manipulating instruments from afar. What distinguishes robotic surgery from its predecessor, laparoscopic surgery, is that rather than merely deploying a tube with a cutter and a grabber at each end, the surgeon using an RASD has more dexterity—the device can twist around and act as a second wrist or eleventh finger. These robotically assisted surgical devices took off at first in urology, because many urological and gynecological procedures involve very sensitive tissue that can only be accessed through a small opening at the base of the pelvic bowl. The device can achieve forms of movement and illumination of tissue that would be impossible using human hands alone.

That is not to say that the transition from open to robotic prostatectomy was an easy one. Surgeons who had worked their entire lives through direct manual touch had to adopt their practice to what could start off as an unintuitive imaging and manipulation system. At the beginning, for many surgeons, the lack of direct touch—the so-called haptic interface—made surgery more difficult or time-consuming. However, over time, surgeons developed the ability to detect other cues for the feel of tissue—for example, how quickly it moves once probed, or how blood vessels blanch when the metallic ends of the robotic probe contact them. For a surgeon who has already seen and more directly prodded bodily tissue hundreds or thousands of times, the association of certain sights with other feelings—of softness or hardness, thickness or thinness—provide a reservoir of intuition about what the video from the RASD is showing. One can think of adapting to a surgical RASD as the development of skills somewhat like those required in a videogame console control—so that just as the flick of a trigger on a joystick could result in a kick for a player’s avatar in a game, so too can a small movement in the Da Vinci console cause a cut or lift a vein. Moreover, as video proliferates, the “second nature” of the screen may become a “first nature” for trainees, and a source of data for machine learning programs to identify past errors and deter future ones.

According to some critics of health technology, the dissemination of RASDs is yet another tale of healthcare spending gone out of control. The devices can cost over $1 million, with high upkeep and maintenance fees. Surgeons must invest valuable time to learn the ins and outs of the new system. Some have questioned the value of the technology.

But it is important to take this early medical evidence with a grain of salt. One key problem emerges in many areas of clinical innovation—those completing robotic surgeries in the first decade of studies could only have had 5 to 10 years of experience using the robot, since it was so new, while their output was sometimes being compared to the surgeons of those who had perfected their skills in open prostatectomies for decades. Outcome measures can also be unfairly narrow. For example, according to some accounts, those who undergo a robotic surgery for prostate cancer can often return home after just four days at the hospital, while those undergoing open prostatectomies often take six or seven days. In the case of kidney cancer, the smaller incision used for robotic surgeries can lead to less pain and shorter
recovery times. Surgeons who use the robotically assisted surgical devices tend to agree that the tools make surgery much easier than pure manual manipulation. The human hand has not evolved to manipulate a scalpel to make fine distinctions between healthy and cancerous tissue; surgical robots can be specifically designed to take on this task. Videorecording via miniaturized cameras may also enable new research on body tissue. This recording already helps speed the diffusion of surgical innovations, as doctors share videos of particularly successful surgical techniques at medical conferences.

Complementary robotics are dominant now. To promote its regulatory agenda in the area of robotic surgery in 2015, the Food and Drug Administration announced a public workshop on Robotically-Assisted Surgical Devices. Speakers included cutting edge physicians and firms. Neither actual implementations of, nor planned development of, fully autonomous surgical devices were high on the agenda. Admittedly, firms planning fully autonomous systems may be in stealth mode—they could lose current surgeons as clients if they talked too openly about replacing them. And in 2016, a stitching robot did mark one notable exception to this pattern. While acknowledging that “the current paradigm of robot-assisted surgeries (RASs) depends entirely on an individual surgeon’s manual capability,” inventors demonstrated that a robot could stitch a split pig intestine together, besting the performance of human surgeons. Billed as the “first autonomous robot” to operate, it managed to bind a hole in soft tissue with speed and precision. The question, now, is whether further industrial development in this area should try to change the dominant trend in robotics, by replacing human surgeons—or if the present path of human-computer interaction is something to maintain.

In theory, it would seem obvious that a robot with minuscule, nimble, even laparoscopic probes would be a better interventionist than the average surgeon—and perhaps, eventually, even the best ones. “We rely on the dexterity of human surgeons but now we know machines are quite a bit more precise than humans. If you want to do things with extreme precision, a machine would be better,” said one Google researcher. And if the “Smart Tissue Autonomous Robot (STAR) could sew more evenly and consistently than even an experienced surgeon” on a pig intestine, there is no reason in principle it could not do the same with human flesh. However, “STAR was still dependent on a surgeon to make the initial incision, take out the bowel, and line up the pieces” before it began suturing. As leading health technology scholars observed in a review article, it will likely be decades until fully autonomous robots take on a surgery from start to finish.

Direct-to-consumer medical automation and robotics is currently not a plausible step forward in many areas. First, scientific evidence is often extremely difficult for the layman to interpret. Large corporations can and often do market products in unscrupulous ways. Large pharmaceutical firms and device manufacturers have systematically skewed data to support their products. Responsibility for harms is also often defused or deflected. Information is scattered, and those untrained in medicine may not be able to interpret conflicting studies. Uncomplicated medical devices, like joint replacements and screws, have continued to be implanted in patients years after serious safety concerns were raised. Moreover, firms may impose on individuals “hold harmless” clauses, preventing future lawsuits. In other words, even in a field as technical and, in principle, automatable as surgery, it is vital to keep some person in the loop as a source of information and advice for laypeople. MIT economist David Autor offers a general reality check about automation that applies with even more force here:

---

13 SHADEMAN et al. (2016); MOLE (2016).
15 ZHANG (2016).
16 ZHANG (2016).
17 SIMSHAW et al. (2016).
18 PASQUALE (2013); GOLDACRE (2013); LENZER (2017).
19 BERNSTEIN (2019).
20 RADIN (2012).
Most automated systems lack flexibility—they are brittle. Modern automobile plants, for example, employ industrial robots to install windshields on new vehicles as they move through the assembly line. But aftermarket windshield replacement companies employ technicians, not robots, to install replacement windshields. Why not robots? Because removing a broken windshield, preparing the windshield frame to accept a replacement, and fitting a replacement into that frame demand far more real-time adaptability than any contemporary robot can approach.

Of course, futurists can probably imagine a robot in a self-driving car that can navigate itself to your car, drive it to a garage, and order other robots to replace the windshield. But even that scenario depends on a chain of contingencies and potential human interventions when things go wrong. When the stakes are higher—for instance, replacing a kidney instead of a windshield—then even more back-up systems and planning will be necessary.

To be sure, even in situations where a human health care provider is fully in charge, there is some level of responsibility, and perhaps even agency, in AI tools used. Bruno Latour puts the matter starkly, claiming that “any thing that does modify a state of affairs by making a difference is an actor.” He concedes that “it is hard to see how a hammer, a basket, a door closer, a cat, a rug, a mug, a list, or a tag could act.” However, as one of his more insightful reviewers observes:

Contrary to some caricatures, actor-network approaches do not fight sociology by proclaiming a contrary religion in which objects and scallops are mysteriously endowed with agency: they simply refuse to ignore, a priori, large realms of activity. Close empirical observation—and not only ethnography, incidentally—may reveal that objects, animals, plants, theories, chemicals, texts, and people may be jostling and pushing in ways that escape observers who already know that the truth lies under the facts, in fields, class structure, etc. What is or is not an actor is a matter for empirical investigation; there is no general theory.

A similar resistance to a totally generalized theory of tort liability is evident in U.S. doctrine, in ways that lend some credence to actor-network-driven approaches in law. For example, a leading case promoting a national (rather than local) and unitary (rather than tiered) standard of care in malpractice, also recognized a “resource-based caveat:” a doctor in a resource-constrained setting cannot be expected to meet a standard of care established in a more resource-supplied locale. Thus medical equipment in such a case may be considered something of an actant in the court’s conceptualization of medical responsibility. Similarly, the development of theories of enterprise liability in health care indicate a level of responsibility for outcomes to a non-human “person,” the corporation. Nevertheless, the distinction between a (still-putative) robotic surgeon, and a robotic tool operated by a person, is clear. Even if the tool acquires sophisticated sensing and clinical decision support capabilities, it is ultimately aiding a person in a task, rather than completing the task itself.

Even profit-minded robotics manufacturers may want to keep human beings “in the loop”—both to assure better use of their products, and to deflect liability. Legal reform will be needed to avoid opportunism built around excessively deflective doctrines. If something goes wrong with a mechanical system—be it an autopilot on a plane or a device used in surgery—doctrines of “competent human intervention,” “the learned intermediary,” or “captain of the ship” have tended to shift liability to the person operating (or merely capable of taking control from) the device, rather than the device maker. But even when robotics and AI only complement a professional, there still needs to be opportunities for plaintiffs and courts to discover whether the technology’s developers and vendors acted reasonably. Such inquiry is

21 AUTOR (2014).
22 LATOUR (2005).
23 LATOUR (2005).
24 VALVERDE (2007).
endangered by expansive interpretations of the above doctrines\textsuperscript{27}. As the example of the tendon-cutting device showed, all responsibility for an error should not rest on a doctor when complementary robotics fails to accomplish what it promised to do. To hold otherwise would again be an open invitation to technologists to rest on their laurels\textsuperscript{28}.

Even if technologists develop fully autonomous robot surgeons, the ultimate “backup system” would be a skilled human surgeon with some experience, flexibility, and creativity\textsuperscript{29}. Our aim should not be to replace such individuals, but to aid in their efficiency and effectiveness. The sequence and shape of automation in health care cannot simply be dictated from on high by engineers. Rather, domain experts need to be consulted, and they need to buy into a larger vision of progress in their field. Perhaps more of medicine should indeed be automated—but law should help ensure that physicians themselves are lasting partners in that process. They should be helped, not replaced, by machines, for the short to medium term.

Of course, in the long term, new arrangements may arise. The distinction between complementary and substitutive robotics may become more a difference of degree rather than kind in some routinized aspects of medical practice. The right tools make a job easier—and at times even more engaging. An analogy from the use of technology in driving is useful. A truck driver may find that cruise control frees his foot from the gas pedal. Automatic transmission makes it easier to shift from high to low gear. Collision avoidance software can warn him about cars in his blind spot\textsuperscript{30}. Technology can make the job much easier—until it replaces the driver altogether. So, there is delicate balance between inventions that help workers and those which replace them altogether. Economists tend to call the former “complementary” to labor, and the latter “substitutive.”

The “be careful what you wish for” story of a worker gradually replaced by his tools has a long history. Aristotle speculated about the effects of self-driving looms centuries before they transformed manufacturing\textsuperscript{31}. Hegel tells the story of a master who gradually becomes weaker and less competent in comparison with a slave whom he forces to perform ever more tasks. Labor economists have worried that “deskilling” is the natural consequence of a more mechanized workplace, paving the way to mass automation\textsuperscript{32}.

However, a smooth transition from “being helped” to “being replaced” by technology is not an inevitability. Nor should it be in medicine. While fields like driving have a relatively simple goal (getting to a destination as quickly and safely as possible), much of medicine entails difficult and subtle trade-offs. There is a much better case for aspiring to build drivers’ skills into autonomous vehicles, than trying to do the same for physicians\textsuperscript{33}. Whereas the relevant data about autonomous cars will be relatively transparent to potential buyers, performance data for autonomous medical equipment is likely to be more opaque and contested\textsuperscript{34}. For that reason alone, keeping a “human in the loop” is critical. When there are complicated value judgments at stake (for example, whether to try a riskier or experimental knee surgery in order to try to increase the patient’s ability to run afterwards), there are all manner of trade-offs that demand a skilled and experienced domain expert’s attention.

\textsuperscript{27} The learned intermediary doctrine holds that the manufacturer of a new technology “discharges their duty of care to consumers by providing adequate warnings” about its potential for harm to professionals using the technology. NELSON (2016).
\textsuperscript{28} KESSELHEIM (2010). Note that both physicians and technologists may share responsibility for preventable errors. The amount of compensation in both negligence and strict liability regimes may be limited by state legislatures to avoid over-deterring innovation. But compensation is still due.
\textsuperscript{29} CARR (2015).
\textsuperscript{30} LEVY (2015); NAT’L HIGHWAY TRAFFIC SAFETY ADMINISTRATION (2016).
\textsuperscript{31} ARISTOTLE (350 B.C.E.), if “shuttles wove and picks played kitharas [stringed instruments] by themselves, master-craftsmen would have no need of assistants and masters no need of slaves”.
\textsuperscript{32} BRAVERMAN (1974); GOLDFIN & KATZ (2008).
\textsuperscript{33} Moreover, even in the realm of driving, some firms are focused on keeping human beings in the picture. For example, Toyota has promoted cars with a spectrum of machine involvement, from chauffeur mode (which requires minimal monitoring by a driver) to guardian mode (which focuses the car’s computing systems on accident avoidance, while a person helms the vehicle). Planes have had autopilot capacities for decades, but commercial carriers still tend to have at least two persons at the helm.
\textsuperscript{34} GOLDACRE (2013).
3. Strict Liability for Substitutive Automation

Insurance contracts, licensure, and certification rules have a powerful impact on technological development. There is no autonomous robotic surgeon today—only “robotically assistive surgical devices.” Even if some genius were to invent a fully autonomous surgical machine, it would need to be vetted by years of tests and research before mass adoption occurred. Reimbursement rules may create another hurdle for rapid adoption of robotics. When robots generate marginally better outcomes, public and private insurers will think twice about paying high fees to guarantee access to them.

Liability concerns will also slow the development of autonomous systems. For example, anesthesia may seem like the ideal use case for an autonomous robot, since machine-readable reports of bodily states may, in principle, be able to indicate any untoward development meriting an intervention. However, the field still seems to be focused on assistive models. The Sedasys anesthesia machine, for instance, is licensed by the Food and Drug Administration to assist anesthesiologists in relatively straightforward operations. It can monitor patients’ breathing and heart rate, administer set doses of anesthesia, and alter those doses in response to new data. Like the “guardian” mode of cars, designed to prevent accidents, Sedasys robots could spot warning signs of adverse events in advance of their actually occurring.

The FDA observed that Sedasys’s technology might lay the foundations for higher levels of technological intervention in anesthesiology: The approval of the SEDASYS System represents a notable advancement in the field of semi-autonomous control of drug administration in medicine. The device utilizes negative feedback from specialized physiological monitors to assess and limit drug dosing and thereby control the depth of sedation. The principle of negative feedback control may be applicable to a variety of drugs and clinical scenarios different from those associated with sedation management.

However, the agency also stated that “the use of the device is restricted to settings where a practitioner trained in the administration of general anesthesia is immediately available to the user for assistance or consultation as needed. Immediate availability in this context means that an anesthesia professional will be available on site to respond to an emergency situation.” This type of safeguard both reflects and complicates the larger argument of this article that human control or monitoring should be required no matter how autonomous the robotic system. While the core case of complementarity is a physician directly operating or supervising the relevant medical AI and robotics, a secondary application of the concept may include a physician (here, an anesthesiologist) maintaining presence in case of complications.

There are many possible future developments for such anesthesia technology. In Europe, national health authorities will be pulled in opposing directions. Health cost cutters may favor full robotization as a cost-cutting measure. On the other hand, European workers have, in general, been able to play a larger role in the deployment of technology than their American peers. That trend would encourage a slow roll-out of the devices, as they gradually...

---

35 HAMILTON-PIERCY (2007).
36 GOLDBERG (2012).
37 See, exempli gratia, UNITED HEALTHCARE (2016). “UnitedHealthcare Community Plan considers S2900 (Surgical techniques requiring use of robotic surgical system (list separately in addition to code for primary procedure)) to be a technique integral to the primary surgical procedure and not a separately reimbursed service. When a surgical procedure is performed using code S2900, reimbursement will be considered included as part of the primary surgical procedure. Use of Modifier 22 (increased procedural services) appended to the primary surgical procedure is not appropriate if used exclusively for the purpose of reporting the use of robotic assistance. Modifier 22 may only be used when substantial additional work is performed, (i.e., increased intensity, time, technical difficulty of procedure, severity of patient’s condition, and physical and mental effort required) that is unrelated to robotic assistance. Documentation must demonstrate the reason for the substantial additional work performed during the surgical procedure”. HEALTH NET (2016); ANTHEM BLUE CROSS (2015).
39 AM. ASS’N OF NURSE ANESTHETISTS (2014); AM. SOC’Y OF GASTROINTESTINAL ENDOSCOPY (2011); U.S. FOOD & DRUG ADMIN. (2015); SINGH et al. (2016).
proved their worth, and health care workers (including anesthesiologists and nurse-anesthetists) transitioned toward positions monitoring and improving the machines—or migrated toward jobs still requiring the “human touch”.

In the US, there are also conflicting political and economic currents. In the country with the highest health care expenditures on the planet, costs are always a concern. But risk-averse hospitals may only permit patients to opt for cheaper, robotic anesthesiology if they sign disclaimers of liability, promising not to sue the hospital if something goes wrong. The dubious legal status of such waivers in the past have sandbagged trends toward “consumer directed health care” in the United States. While some doctors wanted to give patients the option of “last year’s medicine at last year’s prices,” they did not want to be sued for malpractice if the cheaper option proved ineffective. Similar concerns will arise as device makers market robotic systems for hospitals and doctors’ offices.

Assume, for now, that such disclaimers and exculpatory clauses prove ineffective. What would be the proper liability regime for a fully autonomous anesthesia machine? Assume for purposes of this article that the machine’s operations were not explicable to the surgeons and other medical personnel among whom it was deployed—so that there is no sense in which it could be considered merely “helping” them. What is the proper way to assess its responsibility (or, more accurately, the responsibility of its manufacturer, distributor, and retailer) for preventable adverse outcomes?

In a jurisdiction where human anesthesiologists are widely available, strict liability is a compelling approach. In a negligence regime, there are simply too many ways to manipulate notions of the standard of care to exonerate technology developers and vendors. Consider, for instance, a machine that fails to include data about certain vulnerable groups. If someone in such a group is harmed because of such a lack of data, they may not be able to recover damages under negligence standards, since the device’s vendor may claim (if there is adequate statistical evidence) that the device met the standard of care for the physicians it replaced. By contrast, strict liability requires vendors of substitutive AI and robotics to anticipate and meet such predictable defects, and to invest in monitoring for and anticipating less predictable errors. Note, too, that strict liability does not require a vendor of substitutive AI to be “perfect.” Expected damages will be balanced against expected profits. To the extent a strict liability standard appears to be undermining valuable innovation, limits on damages can temper the potential unfairness of liability without fault. Just as they do with respect to malpractice exposure, insurers may offer liability insurance to assist innovators in risk-shifting. And if the substitutive AI has a performance record clearly as good as, or better than, the extant standard of care (be that unaided human care, or, more likely, human-machine cooperation), policymakers should step in to pay damages out of the public fisc, or to provide cheap insurance against such damages to vendors of substitutive AI and others potentially held liable under a strict liability standard.

---

42 See Tunkl v. Regents of the Univ. of Cal., 383 P.2d 441 (Cal. 1963), disallowing waivers of liability; but see Colton v. N.Y. Hosp., 414 N.Y.S.2d 866 (1979), upholding exculpatory clause in a case where an experimental treatment was the only option for the patient.
44 On the critical distinction between explainable and non-explainable AI in medical automation and robotics, see EVANS & PASQUALE (2021).
45 LINDENFELD (2016); REUTIMAN (2012).
46 There are also many troubling ways to shift blame in the complex relationships between medical professionals, hospitals, vendors of AI, and others in the distribution chain—particularly given “hold harmless” clauses or other exculpatory contract terms that may be foisted on providers. PASQUALE (2019) (discussing myriad, mutually reinforcing strategies of liability-deflection).
47 By contrast, if the extant performance standards are much better than those of substitutive automation, the significant damages available pursuant to strict liability for substitutive automation will be a critical prod toward the restoration of a unitary standard of care.
48 Strict liability could also transition into a no-fault compensation scheme like the US’s National Vaccine Injury Compensation Program, which compensates those who have suffered side effects from vaccinations. See, exempli gratia, ZOLLERS et al. (2005), “we favor an approach that accepts strict product liability for software and lets the case law find the point of maximum social benefit through the application of a rigorous cost/benefit analysis. In the rare case where a beneficial software product is being kept out of the market because of liability concerns and that claim is corroborated, there are a number of public policy strategies.
However, the definition of “as good, or better” performance records needs to be granular, so as to be sensitive to historical victims of health disparities. There has been growing concern that the data used for diagnostic AI may not adequately represent all groups in society. For example, minority groups may be poorly represented in databases. Women are seriously disadvantaged as well. Diagnostic AI that ignores all these problems, and which still generally delivers better results than unaided human observation, may not be actionable under a negligence standard for those it fails to help—particularly if the standard of care is unaided human observation. However, under a strict liability standard, failure to include available, more representative databases, that leads to preventable accidents, would leave vendors liable for adverse events even if they managed to do better on average than the standard of care. Such liability may be critical to incentivizing them to address health disparities.

Efthimios Parasidis has convincingly argued that courts need to recognize and counteract automation bias—that is, the tendency of persons to assume without proper evidence that a machine has better judgment than persons. The problem of automation bias is recurrent and is a persistent temptation when often-overworked professionals seek tools to ease their workload. More stringent liability standards are a way of gradually ensuring a lower risk level in the industry.

A vehicle manufacturer may be held responsible for an accident if the manufacturer failed to design or manufacture the vehicle properly. Similarly, AI and robotics may be designed or developed in a way that fails to conform to basic standards of safety and reliability. The product analogy increases accountability for safety, reliability, and security. Unpredictability of advanced AI systems means that forms of accountability reflecting classic legal standards are critical. For example, those keeping particularly vicious or wild animals (like lions and tigers) are strictly liable if these beasts escape and cause harm. Enterprise liability “asserts that actors should bear the costs of those accidents that are characteristic of their activities and then distribute those costs among all those who benefit from the imposition of the risks at issue.” Professor Danielle Keats Citron inventively applied these ideas to the digital age by analogizing massive data holdings to early water reservoirs which, if breached, could cause death and destruction to communities immediately adjacent to them. Much the same could be said of autonomous robotics or AI in critical medical situations when an irresponsible user decides to simply let them run autonomously. When they are marketed or developed to be used in autonomous mode, their developers and vendors must take responsibility. AI and robotics systems are ultimately attributable to humans.

that can be employed in mitigation. For example, there is federal legislation limiting vaccine manufacturers’ liability for certain vaccines. The NVICP is funded by an excise tax on each vaccine dose.

49 ADAMSON & SMITH (2018). This lack of diversity also afflicts genomics research. Non-European groups tend not to be as well-represented in DNA databases than European groups. POPEJOY et al. (2018).

50 CRIADO (2019).

51 PARASIDIS (2018), “as to the coding phase of software development, a strong argument can be made that coding should be encompassed under the category of manufacturing defects [and this is critical because U.S. courts typically employ strict or product liability analysis for manufacturing defects]... [A]llowing products liability claims for CDS systems also may be a way to counter disclaimers of liability that typically are found in CDS contracts”. See also PINKNEY (2002), focusing on security-related failures (“Manufacturing software in general and software providing network functions in particular, is arguably abnormally hazardous... [and even] “if courts do not find strict liability, software manufacturers should be liable for the harm that results when they fail to take due care preventing design flaws that allow security-related software failure.”); SCOTT (2017).

52 CARR (2015).

53 ZOLLERS et al. (2005).

54 However, as Jamil Ammar warns, “From the perspective of product liability, courts in the U.S. consider computer software to be a service rather than a product. To date, courts have been reluctant to extend theories of product liability to software”. AMMAR (2019). One purpose of this article is to urge courts (both in the U.S. and elsewhere) to reconsider this approach in the context of substitutive AI.


56 KEATING (2007).

57 CITRON (2007).

58 And they should be kept that way. PASQUALE (2020), proposing a fourth law of robotics requiring all AI and robotics to be attributable to responsible persons.
A strict liability standard will be controversial. The legal scholar Ryan Abbott has argued that, if an autonomous vehicle is, in general, safer than the typical human driver, only a negligence cause of action should be available. If accepted, such a comparison would make negligence, rather than strict or product liability, the proper judicial response to errors. For Abbott, the standard of a “reasonable computer” would then supplant that of the “reasonable person,” in judicial considerations of the type and level of responsibility to assign. Such an approach would help ensure that there are not undue impediments to the development of autonomous vehicles. However, it is likely less appropriate in the medical field, since health care providers are likely to be far more valuable in guiding the deployment of technology, long-term, than “guardian drivers” who have been deployed to restrain self-driving cars when they errantly threaten to crash into a person or another vehicle. That ongoing need for guidance and development by present, expert, professionals counsels in favor of a strict liability standard in the case of substitutive automation in medicine.

It is not unreasonable for persons to expect that AIs will have some basic rules of engagement, such as being programmed and designed to limit human harm. This was articulated as the first of Isaac Asimov’s Laws of Robotics and shows up in various forms in many other works on human-machine interaction. Strict liability for machines and AI that substitute for professionals, as mitigated by economically calibrated damages limitations, is one more way to ensure that prevention of harm is prioritized while innovation is not unduly hampered.


While promoting AI as a substitute for competent medical personnel, some firms will fail to engage in the quality control and other steps necessary to avoid disastrous outcomes. Tort lawsuits will follow, with plaintiffs demanding damages for the firms’ failures to meet the relevant standard of care. Legislators and courts will need to develop approaches to liability adequate to the new technological environment. As they do so, they will effectively set nuanced and contextualized standards for the deployment of AI. Distinguishing between complementary and substitutive AI is one conceptual tool that will help them do so.

When AI or robotics simply assist a professional, they are tools. In medicine, the doctrine of “competent human intervention” has shifted liability away from those who make devices and toward the professionals who use them. However, the professional in such scenarios should not bear the entire burden of responsibility. His or her tools can be produced well or badly, and vendors of defective, complementary AI and robotics should be held responsible for negligence. Both legislators and courts will need to develop standards of care designed to incentivize proper safety, security, and risk avoidance. But the burden of proof will be on the plaintiff to demonstrate that not only a skilled medical professional, but also the maker of the tools used by such a professional, should be held liable for a preventable adverse outcome.

When AI and robotics replace a skilled medical professional, the burden shifts. The vendor of such computational systems needs to take on the responsibility for errors and accidents. At the damages phase of litigation, the vendor may explain how its damages should be mitigated based on its AI’s performance relative to the extant human or human-machine based standard of care. Such responsibility for explanation will serve an important information-forcing function in areas where public understanding is often limited by trade secrecy.

Accountability is a contested and complex concept in tort law. It is all too easy to reduce the problem of preventable adverse events in medicine as a simple matter of providers’...
responsibility to patients. However, a broader political economy perspective goes beyond the dyad of provider-patient, incorporating larger concerns about the nature of the labor force, the explainability of AI, and the power of dominant technology firms. As AI and robotics take on more roles, there will be cost pressures for technology to prematurely replace providers. Strict or enterprise liability for such adverse events arising out of particular replacements will generally help deter its happening too quickly. By ensuring that vendors of medical AI and robotics are more accountable to those whom they harm, administrative agencies and courts may renew an ongoing quality movement within the profession of medicine. They may even spark the professionalization of AI research itself, since professions are institutions of accountability that help assure ongoing self-review and improvement. And if there are concerns about liability over-detering innovation, damages caps may be imposed to calibrate incentives accordingly.

From an individualistic, utilitarian perspective (dominant in mainstream economics), substitutive automation of machines to replace humans in many fields seems to be a foregone conclusion, thanks to a set of interlinked value judgments about the value of cheapening tasks. But within a profession like medicine, matters are more complicated. A renewed political economy of automation demands a role for professionals to mediate between patients and complex technologies. Professionals enjoy forms of autonomy, and are burdened by constraints, rare in non-professional fields. Professionals are charged with protecting distinct, non-economic values that society has deemed desirable. Their labor, in turn, reflects, reproduces, and is enriched by those values. Knowledge, skill, and ethics are inextricably intertwined. In the face of well-hyped automation, professionals ought to reaffirm their own norms. The threat of tort liability for the firms they work for (including AI vendors) gives them some leverage to push back against management demands for premature automation. Indeed, when Marc Law and Sukkoo Kim examined the history of professionalization and occupational licensure, they found patterns of worker self-organization in the United States in the early 20th century that substantially increased consumer protection. By deterring premature substitutive automation, a liability regime that reduces potential exposure of AI vendors when they complement (rather than substitute for) medical professionals will help ensure a democratization of expertise, including ongoing critical evaluation of medical AI and robotics by physicians and other providers.

Of course, as courts develop such evolving standards of care, they will also face predictable efforts by owners of AI to deflect liability. For example, firms may require their customers or users to sign exculpatory clauses and other contractual limitations on liability by waiving their right to sue. Asymmetries of power are important here. For example, in medicine, courts have resisted exculpatory clauses purporting to relieve physicians of responsibility for malpractice, thanks in part to the asymmetrical power of hospitals and their patients. They should be similarly wary in AI-intensive scenarios, where even greater imbalances of power and knowledge are common.

Policymakers are currently struggling to keep pace with the speed of technological development. Legislators have been hesitant to pass broad statutes, as they are fearful of inhibiting growth and innovation in the space. However, increasingly there is public demand for policy interventions and protections regarding critical technology. These demands do not necessarily impede economic or technological advancement. Some fields may never get traction if customers cannot be assured that someone will be held accountable if an AI fails. Developing appropriate standards of responsibility along the lines prescribed in this article

---

62 On the critical distinction between explainable and non-explainable AI in medical automation and robotics, see EVANS & PASQUALE (2021).
63 PASQUALE (2015).
64 LAW & KIM (2005).
65 SAWICKI (2018).
66 SMUHA (2019).
should advance the quality of both artificial intelligence (AI) and intelligence augmentation (IA).

Cited bibliography


LENZER, JEANNE (2017): The Danger Within Us: America's Untested, Unregulated Medical Device Industry and One Man's Battle to Survive It (New York, Little Brown).


