Who is Liable for the Mistake in Treatment Performed by an Autonomous Robot-surgeon – An Interdisciplinary Analysis from a medical and criminal law perspective

Igor Vuletić

PhD in Criminal Law, Associate Professor of Criminal Law

Ivan Koprivčić

MD, PhD, specialist in general surgery and orthopaedic traumatology

1. Introduction

The modern era can be marked as an era of the evolution of advanced technologies. The process of automation and the powerful incursion of technologies based on artificial intelligence (AI) are gradually becoming a part of everyday life. Medicine is one of the fields in which this phenomenon has come to particular prominence. In some fields of medicine, such as surgery, robotic technology has gradually been introduced during the last thirty years, precisely with the aim of reducing the risk of mistakes and improving the quality of delivered services (Camarillo, Krummel & Salisbury, 2004). This development entered a new dimension at the beginning of 2017 when a group of American scientists presented the robotic-surgeon project known as the Smart Tissue Autonomous Robot or STAR which completely autonomously, independent of human guidance, performed surgery on a pig's small intestines and did this quicker and more precisely than the team of surgeons who performed the same operation on another pig (Strickland, 2017). This successful experiment heralded a new era of surgery based on AI, which is capable of performing surgery independently of humans and make independent decisions, and based on conclusions from the field and relevant information.

AI is poised to revolutionize the field of surgery, making procedures safer, more accurate, and more effective than ever before. Current trends include the following: computer vision - AI algorithms can analyze images and videos from surgical procedures in real-time, allowing surgeons to make more accurate diagnoses and perform more precise procedures. Robotics: robotic surgery is becoming more common, allowing surgeons to perform procedures with greater precision and less risk of complications. AI algorithms

are also being used to control surgical robots, making them more intelligent and capable. Predictive analytics: machine learning algorithms are being used to analyze patient data, including medical records and imaging results, to predict the likelihood of surgical complications and recommend the best course of treatment. Augmented reality: AR technology is being used to create 3D models of patient anatomy, allowing surgeons to visualize the surgical site in greater detail and perform procedures with greater precision. Virtual assistants: AI-powered virtual assistants can provide surgeons with real-time information and guidance during surgeries, improving efficiency and reducing the risk of errors.

The authors use interdisciplinary approach. First, they present the current state and future expectations of AI in surgery. They explain the benefits and potential challenges of AI technology from a surgeon's perspective and elaborate their standpoint on the issue of desirability of AI in the OR. Afterwards, they adopt the role of legal experts and examine the problem through a legal lens. They then explore the question of whether the current legal framework (in comparative legal systems) is sufficient to provide adequate legal protection for patients in cases of medical error, assuming that artificial intelligence will play a significant role in patient treatment in the near future. They compare the legal systems of certain European countries, which treat medical errors under criminal law, with the legal system of the United States (which approaches this through civil liability). The criminal law approach is common in continental Europe, although not all countries have the same system. Some, like Germany and Austria, treat medical errors as bodily injury or negligent homicide, while others, like Croatia and Slovenia, provide for a specialized criminal offense of negligent treatment. The purpose of this overview is to provide an overview of legal arrangements in Western and Eastern European countries (where healthcare is still predominantly a public sector concern) and compare them to the system in the United States, where healthcare is largely in the private sector.

The authors analyze the problem from the perspectives of both a lawyer and a surgeon, juxtaposing their arguments and attempting to find a common middle ground. In this way, the authors are actually engaging in a certain scientific interdisciplinary dialogue. Using a combined approach of comparative, deductive, and case study methods, the authors discuss two highly relevant questions: whether the current concept of criminal liability for medical errors is adequate, and in which direction legal regulations in this field should

be developed to provide effective legal protection while not hindering technological progress. The specific contribution of this study lies precisely in the comparison of the European and American approaches to the subject matter. It is noteworthy that, in the existing literature, no such attempt has been made in this particular context. However, the primary purpose of the paper is not comparative, but instead to present an interdisciplinary conclusion regarding the desirability of AI technology penetration in the field of surgery, by taking into account both its advantages and drawbacks. The authors advocate a position in which the contribution of AI is undoubtedly beneficial to a reasonable extent, but a scenario in which AI would assume complete control over operations is undesirable from both medical and legal perspectives. Consequently, such a direction of technological development should not be supported in the future.

2. Levels of robotic automation in surgery

Firstly it should be mentioned that the development of AI is also making strides in other fields of medicine, and not only in surgery. Thus, AI is used to a great extent in diagnostics (Stone et al., 2016). However, this paper will only focus on robotic surgery as this is an area in which scientists, as initially noted, have developed a completely autonomous system. In addition, this is an area in which the risk of immediate and direct occurrence of harmful consequences in the nature of things is quite high making it particularly interesting from a criminal law perspective. Of course, this does not mean that the considerations and conclusions in this paper should only be limited to surgery. On the contrary, they should be applied mutatis mutandis to other areas of medicine, in accordance with the dynamics of the development of AI in the relevant fields. Further in the text this will be limited to the most essential depiction of the possible levels of robotic autonomy in surgery, with the purpose of better understanding of further displays in the legal context.

The more serious role of robotics in the operating room began with the development of new surgical techniques in the late 1980s. At that time minimally invasive surgery was being developed which no longer required surgeons to use their hands to enter a patient's body, but allowed the performing of the surgery externally and with the help of instruments which represented a kind of extension of the surgeon's hand (Camarillo, Krummel & Salisbury, 2004). Such an automated approach brought many advantages, of which increased precision and consistency in performing tasks, avoidance of aggravating factors such as tremors and greater agility, and the possibility of always performing the same procedure at the same level of quality are particularly emphasized (Taylor, Menciassi, Fichtinger & Dario, 2008). In addition to expediting procedures, AI-based surgery can significantly reduce treatment costs and make quality medical service more accessible to a broader population, which is why some authors advocate for broader implementation of such technology in the near future (Rasouli et al, 2020).

At the theoretical level, this refers to the multiple levels of robotic autonomy in the operating room. It should be noted however that the most independent concepts (such as the one described in the introduction) are still in the experimental phase. Yip and Das (2017) differentiate four possible levels of robotic-surgeon autonomy with regards to the level of the robotic-surgeon's (in)dependence on the human-surgeon, that is in its ability to have an influence on the course and outcome of the surgical procedure. Their division, as one that is very clear and easy to understand, will be accepted for the purposes of this work. The mentioned authors distinguish four levels of dependence, namely: direct control, shared control, supervised autonomy, and full autonomy.

In the direct control model, the human-surgeon retains complete control over the machine. The human-surgeon uses the machine (whether through manual control or teleoperating technology) as a kind of extension of one's own hands and manipulates the robot. This is still the most common approach nowadays and one which can be found in most teleoperation systems. Yip and Das (2017) cite the da Vinci Surgical System for laparoscopic surgery, developed back in 2000, as a typical example of this model.

The next model is that of shared control, where the human-surgeon and the robot work in a kind of synergy which is manifested through the combination of the surgeon's movement on the one side, and corresponding robotic reactions on the other. In this way, the robot corrects the surgeon's movements and eliminates any deficiencies such as tremors or

similar. Control of the movements of the surgical instruments is thus distributed between the human-surgeon and the robot. A typical example here is the Steady Hand robot which was developed at Johns Hopkins University in the United States (Yip & Das, 2017).

The third model – supervised autonomy – implies that the robot performs surgical tasks under the supervision of a human-surgeon, who gives orders and oversees the operation the entire time. In this model, the robot has certain autonomy in carrying out actual surgical procedures but the surgeon oversees the entire process and at any given moment has the possibility of intervening and directing the robot to work differently. Here the surgeon actually participates less in the actual performance of the surgical procedure itself than in the previous two models, but remains the primary (and only) one responsible in the decision-making. A typical example for this would be the CyberKnife system that was conceived at Stanford University (Yip & Das, 2017).

Finally, the model of full autonomy is one in which a robotic system completely replaces the human-surgeon. Such a system fully and autonomously plans and carries out the surgical operation by extracting the necessary information from data related to the specific patient (test results). In such a system, the human-surgeon actually becomes redundant. This level of full autonomy is still far from being used in modern medicine but the domination of such systems in the near future is noted in literature as the ultimate goal of the healthcare industry (Sharkey & Sharkey, 2013). The STAR system mentioned in the introduction is an example of such a model and indicates all the potential advantages in terms of speed and quality of performance. It can also be assumed that introducing such systems would significantly reduce costs and waiting lists (especially because of the possibility of the robot working an increased capacity of working hours). However, from the standpoint of criminal law, the question that arises is what if a robot-surgeon makes a mistake (chooses the wrong method of work or applies the right method but in the wrong way) and provokes the deteriorated health or even death of a patient.

3. **A surgeon's perspective: do we really want a fully autonomous robot-surgeon in the OR?**

At the outset of this discourse, we must emphasize that we are by no means opposed to the integration of technology into surgical procedures. On the contrary, we are of the opinion that technological innovations, including those based on AI, can significantly enhance the surgical profession and augment the well-being of patients as ultimate beneficiaries of the service. In this regard, the introduction of AI technology into various segments of surgical work is undoubtedly beneficial and should be supported.

Nevertheless, this should be approached with a degree of caution to avoid potential adverse counter-effects. In surgical literature that delves into the question of the benefits and drawbacks of a higher degree of automation in surgery, a valid warning is raised that not every surgical field is identical (as not every part of the human body functions in the same way). For instance, in abdominal surgery, there is a cautionary note that complete delegation of decision-making during surgery might not be advisable, as certain procedures involve multiple intricate steps that are not uniform but vary from one operation to another and from one organism to another. The case of pancreatic surgery is mentioned, wherein it is always necessary to perform many diverse surgical steps including lymph node dissection around the pancreas or the major vessels as well as intricate intestinal anastomoses. Given that the abdomen does not always conform to a standardized anatomy, individual surgical judgment and appropriate experience (and at times, sound intuition) are required to carry out the mentioned procedures in a manner that ensures the utmost safety and benefit for the patient (Kinoshita and Komatsu, 2023). Hence, such a procedure simply cannot be wholly entrusted to AI, as technology cannot make judgments based on intuition and sensation, but rather relies on objective data (which are not always feasible in this context).

In other surgical domains as well, the wholesale adoption of AI approaches is met with a significant degree of skepticism. For instance, in orthopedics, it is cautioned that current research has demonstrated the rapid expansion of AI potentials in joint and shoulder surgery, enabling better risk assessment, thereby facilitating informed decision-making for patients. However, it is equally underscored that the performance of existing models remains highly constrained and lacks sufficient external validation, necessitating further rigorous scientific investigation before entrusting this endeavor more extensively to AI (Gupta et al, 2023). The aforementioned reasons have led to the fact that, despite the fact that AI has been penetrating the field of surgery intensively since the beginning of the

millennium, the majority of surgeons today are still skeptical about the widespread use of such technology (Gumb set al, 2022; Lisacek-Kiosoglous et al, 2023).

In an effort to contribute our experience to such discussions, we align ourselves with proponents who assert that AI technology, despite its unquestionable utility, should never completely take over the surgical process. On the contrary, in our view, the responsibility and authority for making critical decisions during surgery must remain with the human surgeon. This conclusion stems from the fact that a surgeon must possess two components: the one we might call the professional and the one we might label as instinctual-emotional or, more simply, human. The operative procedure, in fact, is consistently unique, and each surgery, no matter how many times it has been performed, can present varying circumstances, at times resembling a first-time procedure. This can arise, as already mentioned, due to diverse factors: unexpected clinical presentations, atypical anatomical structures, rare and unforeseen complications, previous incorrect or incomplete diagnoses, extraordinary conditions under which the surgery is conducted, and so forth. In such situations, the surgeon transcends being merely an expert, following rigidly the so-called "rules of the profession" and proceeding "by the book." Conversely, a different, human component activates, and the surgeon makes decisions guided not only by professional protocols but also by experience, emotions, and instinct. Such conduct constitutes an integral aspect of the surgical profession and simply cannot subsist without this facet. As an illustrative example, we can reference the wartime circumstances of 1991 and 1992 in our home country, when we were confronted with the need to rapidly address a substantial number of severe war injuries under difficult conditions, previously unexperienced. In one such scenario, a patient underwent surgery on a carotid artery that had been injured by a gunshot wound to the neck. The dictates of professional standards and literature in this situation prescribed the need for arterial reconstruction—something an AI mechanism, programmed to adhere to professional guidelines, would undoubtedly execute. However, instead, the surgeons decided to ligate the artery, an action not advisable as it could lead to an immediate stroke and death. Therefore, assuming that risk was necessary, with the aim of avoiding a day-long procedure due to the need to operate on other severely injured patients. The surgeons consciously and in good faith took on that risk, relying primarily on their experience with artery surgeries, and the patient ultimately survived without complications. In other words, this was a situation frequently described in legal literature as one of necessity, in which the responsible individual (in this case, the

surgeon) must weigh the interests and make a prudent and clear decision. These capacities are simply beyond what AI can possess.

Even in peacetime conditions, various situations can arise where practice demonstrates that strictly adhering to rules can sometimes render procedures less effective or even more harmful than the surgical approach a specific surgeon selects based on experience and instinct in a given situation. For instance, femoral osteosynthesis using the older DHS screw method has proven more effective and less invasive for the patient when accessing the femur involves muscle splitting, rather than circumventing the muscle (as recommended by professional guidelines), as this approach causes less damage to the insertion site.

In conclusion, these and other examples illustrate how the surgical profession is of such a nature that it demands decision-making that cannot be uniform. On the contrary, in addition to knowledge of literature and anatomy, experience, instinct, and to some extent emotional judgment are necessary to arrive at the best decision for the patient. In this context, the notion of the best decision is relative: in a certain situation, it may involve sacrificing a minor interest to preserve a greater one. For these reasons, we advocate an approach in which AI technology should continue to develop in a manner that assists the surgeon both in mechanical aspects of surgery and in decision-making and diagnosis, yet the ultimate decision would always remain with the human surgeon.

4. A lawyer`s perspective: who (what) is to blame for malpractice?

The issue of criminalizing medical errors is highly controversial. There are different perspectives, with some advocating that medical errors should not be treated under criminal law at all, but solely within milder branches of law, such as civil and tort law. A comparative analysis reveals that the latter approach is characteristic of countries like the United States, whose healthcare system relies on the private sector and where a developed system of insurance for medical errors and settlement agreements in litigation exists. In continental Europe, however, the situation is somewhat different. There, it is considered necessary to provide criminal law protection for the health of the patient (as a constitutional category), and thus, various approaches to the criminal treatment of

medical errors are being developed in those countries. Some countries regulate this issue through specially designed, specialized criminal offenses, while others address it through general criminal offenses against life or the body.

In contemporary times, there is an escalating global initiative, particularly within the medical community, advocating for the decriminalization of medical errors. Among proponents of abolishing criminal liability for medical errors, particular emphasis is placed on the observation that various countries around the world lack clear and uniform standards (criteria) by which medical errors can be consistently judged, or when they can be deemed as criminal offenses. This absence of well-defined criteria leads to discrepancies in handling such cases and contributes to legal uncertainty. Furthermore, medical standards may differ from one country to another and vary across different cultures. It is widely acknowledged that not all healthcare systems worldwide are at the same level of development, resulting in significant variations in standards of best practice. Another argument revolves around the association between the fear of criminal accountability and the emergence of undesired phenomena, with the most prevalent and potentially harmful being the phenomenon known as defensive medicine. Defensive medicine can be defined as the occurrence where physicians, in an effort to avoid later legal issues, subject patients to a significant and objectively unnecessary number of tests, prescribe unnecessary medications, request superfluous examinations, order objectively unnecessary consultations with other physicians, and so forth (Pichler & Vuletić, 2020).

On the other hand, proponents of criminalizing medical errors frequently emphasize that only through criminal legal protection can adequate safeguards be fully provided for the health of patients, which is regarded as one of the highest social goods. Moreover, criminal repression, coupled with heightened awareness of the possibility of criminal prosecution, influences physicians to exercise greater caution in their practice. Furthermore, the determination of criminal liability often constitutes a preliminary issue in legal proceedings, and efficiently resolving this matter facilitates and expedites the litigation process. Perhaps the most crucial argument relates to the necessity for an effective level of societal protection (as well as individual protection) and for general deterrence (Dekker, 2011).

Without intending to take sides, it can be observed that in multiple countries worldwide, the concept of criminal liability for medical errors is currently accepted, and there is

almost no developed system that entirely absolves physicians from any form of criminal responsibility in this regard (Pichler & Vuletić, 2020). Although debates on whether to abandon criminalization are present in the literature of many countries, comparative legislations and practices, at present, do not seem to adopt such a significant step. Therefore, in the following discussion, we will proceed from the premise that this form of criminal liability exists and analyze it from the perspective of a situation in which AI assumes an increasing degree of autonomy within the field of medicine, specifically surgery.

The basis for further elaborations can be found in the concept of medical error in terms of criminal law and it needs to be clearly defined and classified separately from the concept of complication. A medical error in the descriptive sense is considered to be every procedure (act or omission) by a medical professional that deviates from the accepted rules of the medical profession and which leads to health consequences for the patient (Kuhn, 1994). In contrast, a complication is considered to be an undesirable consequence of a medical procedure which, however, is performed in accordance with the rules of the profession and the necessary degree of care (Šepec, 2018). Thus the constitutive premise for criminal liability is to establish breach of the rules of a specific medical profession, whether written or unwritten (Korošec, 2016). This means establishing the starting point for proving the guilt of the perpetrator (Mrčela & Vuletić, 2017). Another assumption is determining the causality between such omission and the consequences. After both objective elements have been established, it is necessary to prove the subjective element, which consists of the existence of a suitable form of guilt.

4.1. Brief comparative overview

There are diverse approaches that legal systems employ in addressing medical malpractice. Unlike some countries, medical malpractice is not exclusively confined to criminal law. The United States (USA) can be considered the birthplace of medical malpractice, as the phenomenon originated there from a legal standpoint and subsequently spread to Canada, Europe, and the rest of the world. Notably, healthcare in the USA is predominantly situated within the private sector, in contrast to most European countries where it constitutes a component of the public sector. Consequently, a

significant majority of cases involving negligent medical treatment are resolved through civil litigation or negotiated settlements, thereby avoiding trial proceedings (Di Landro, 2012). Criminal liability arises when intentional errors or errors resulting from gross negligence lead to fatal consequences. One such notable case in recent history involved the trial of an anesthesiologist, Dr. Verbrugge, who failed to monitor vital signs during an operation on an 8-year-old boy in Denver in 1993. He was accused of dozing off for 20-30 minutes during the procedure, consequently missing changes in the child's pulse and rising temperature. It was also discovered that this was not the first incident of its kind, as the hospital had previously warned him about sleeping during surgeries, though no consequences arose from those instances. Due to the failure to detect the changes in vital signs in a timely manner, which was his responsibility as the surgical team was focused on performing the procedure, the child went into shock and passed away. In the initial trial, the anesthesiologist received a conditional conviction for criminal medical negligence for a lesser offense. Subsequently, several years later, a new trial was pursued for a more serious offense (known as a felony); however, these charges were dismissed on procedural grounds (the statute of limitations had passed). Nevertheless, his medical license was revoked in the states of Colorado and California (Di Landro, 2012; West, 2007).

Another example is the case of a certain Dr. Benjamin, a gynecologist who was accused of performing an incompetent abortion, resulting in the patient bleeding excessively and dying. He was convicted of second-degree murder, and the sentence imposed ranged from 25 years in prison to life imprisonment. During the trial, it was revealed that Dr. Benjamin had previously practiced medicine in another state under a different name and had faced legal proceedings there as well. In this particular case, his negligence involved making an incision in the patient's cervix and vagina without subsequently repairing it, leaving her to bleed on the operating table (West, 2007).

It can be concluded that the American legal system only envisions criminal liability for exceptional cases involving very severe errors, leaving the majority of such situations to be addressed through civil and tort law. Consequently, it can be assumed that the issue we are discussing here - namely, the impact of introducing AI into surgery on the criminal liability for potential errors - may not carry significant practical importance in the USA. However, the situation could be considerably different in Europe, as further discussions will demonstrate.

An insight into European comparative legislation of those countries that criminalize medical errors shows that two different approaches are possible. According to the first approach – which has been adopted e.g. in Austrian and German law – medical errors are treated through criminal offenses of bodily harm (Fischer, 2018). According to the second approach – which has been adopted e.g. in Slovenian (Šepec, 2018) and Croatian law – a special criminal offense is provided for such situations, which then has precedence over criminal offenses relating to bodily harm (lex specialis derogat legi generali). What is common to these legal models is that they are designed with the basic premise to be applied to errors and omissions made by humans. It is therefore clear that both legislative approaches are necessarily based on the principle of guilt, as one of the fundamental principles of modern criminal law. In that sense, there is no question that any legislation that is based on one of the two mentioned models will have sufficient capacity for adequate treatment of a situation where a surgeon, acting with the help of one of the first three forms of AI (see section 2), causes serious consequence to a patient's health. Namely, in each of the described three levels, the surgeon retains the decisive influence over the robot and, as such, this includes the criminal liability for the consequence (of course, the greater level of autonomy of the AI, the greater the defendant's ability to successfully invoke suspending of causality or absence of guilt). That is why such situations will not be further elaborated here.

What is problematic here is the fourth situation, which is based on the model of full autonomy of the robot from the human-surgeon. The working hypothesis of this paper is that it is no longer possible to include such a situation on the basis of the current ruling postulates of modern criminal law dogma. In the following, we will endeavor to prove this hypothesis by analyzing three problems which we consider to be an obstacle to establishing criminal liability for medical errors committed by an autonomous robot-surgeon. These problems are: establishing criminal liability for a concept that requires breach of the rules of the profession and difficulties in determining the criminal liability of the person or entity standing behind the robot (be they manufacturers, programmers, health institutions) with a particular focus on the problem of causality and predictable consequences.

4.2. Breach of professional standards as an adequate criterion for criminal liability?

This section deals with the question of for what (which kind of behavior) is one criminally liable. As already pointed out, the criminal liability of a physician can only be discussed when it is determined that there was a breach of the relevant professional standards (i.e. that a procedure was not conducted lege artis). In that regard, such breach can be committed either by actions or inactions. A breach of professional standards by way of action can be made by, e.g. a physician who during surgery, with the intent of stopping excessive bleeding, uses at least 2.5 times the amount of insufficiently diluted adrenaline than permissible and this ultimately leads to the patient's death (Municipal Court in Našice, K-175/06).

On the other hand, even though criminal offenses by inaction are generally rarer than those committed by action, this does not apply to the area of medical negligence, as it is quite often the case that harm is caused as a result of the physician's failure to respond to, as a kind of guarantee of the life and/or health of a patient, in a timely or appropriate manner. Thus, a breach of duty by inaction was caused by, e.g. a physician who did not promptly request blood and urine lab tests even though the febrile state of the patient clearly required this, and by which there was no timely diagnosis of leukemia or course of treatment (Supreme Court of the Republic of Croatia, VSRH, IV Kž-120/91). A breach of the medical professional standards was also caused when a physician who was immunizing a newborn failed to warn the infant's parents of the risk factors of infection (in German Bundesgerichtshof, BGH, ruling no. 3ZR 52/93; for a commentary on the ruling see Schöch, 2010).

In practice what is most problematic is determining which professional standards have even been breached. Professional standards emerge from scientific and expert papers and from conclusions and recommendations of professional bodies at the national and supranational level. This is in fact a large set of rules and standards that are often not systematized and for whose interpretation a court would need the help of experts in the relevant profession and specialization (which is sometimes difficult or even impossible to make available to the court). Namely, in many areas of modern medicine the codes of conduct in specific situations have still not been universally and uniformly codified, and physicians are allowed some freedom in choosing the most appropriate methods. Further, these standards can differ from country to country, especially when considering the development of the health system and degree of medical advancements in a particular

country. Finally, situations where the professional standards differ from the usual practice of medicine in a certain healthcare institution, or even in a certain country, are especially problematic. This can be illustrated by two examples from Croatian case law.

In the first example, a gynecologist doctor used what is known as the Kristeller maneuver which resulted in head injuries to the infant. The gynecologist defended the use of this method, even though it has been established that this procedure is contrary to the standards of the medical profession and achievements of modern science in the field of gynecology and obstetrics, as it is common practice in Croatian hospitals in situations when there are complications during labor (Municipal Criminal Court in Zagreb, K-1291/15). In the second example, also from the field of gynecology, two gynecologists were accused of failing to promptly request the right blood tests for a patient, and did not notice that she had developed sepsis, which ultimately led to the death of the patient. In this matter, however, the court accepted their defense that the clinical picture was atypical, and even if the medical expert's findings showed that in such a situation the lege artis procedure would include requesting the said tests, a large majority of Croatian doctors would not have acted in that way. Consequently, the court absolved them from criminal liability (Municipal Court in Osijek, 1 K-615/10; for a critique of this ruling see Mrčela & Vuletić, 2017). These examples indicate how the criterion of breach of professional standards can sometimes be questionable in practice.

In the context of the issue being described here, that being the possibility of breach of professional standards by a robot acting per the fourth level of autonomy, an additional problem that we identify is the fact that such a robot will certainly be programmed according to the rules of the profession. This means that programmers will be able to feed the robot all the data related to the actual professional standards and scope of medical science in advance and that this information will be updated regularly. The robot will then draw from its database (most likely) a much wider "knowledge base" of professional standards than a human-surgeon would be able to do (after all, that is one of the benefits of robotics). It will also apply this information, according to complex algorithms, in such a way that where there are more than one set of professional standards, it will act according to those that are statistically more likely to be successful (or which minimize the risk of harm). If that is so, then it would be logical to conclude that breach of professional standards will no longer be the appropriate starting point in assessing criminal liability

for the error as the likelihood for such breach would be close to zero. In other words, the criterion for breach of professional standards and conceptualizing actus reus accordingly is based on the realization that (all) humans make mistakes. With AI, however, this reality is different. That is why with systems that are based on such a premise it will be difficult to judge the consequences resulting from treatment undertaken by autonomous robots. Therefore, another starting point will need to be found for this future concept.

In the sense described above, it is particularly interesting to analyze the German concept of criminal liability for medical errors, which has already been mentioned earlier. As noted, in Germany there is no separate criminal offense for medical negligence but rather such situations are dealt with under Article 223 of the German Criminal Code (in German: Strafgesetzbuch, STGB), which regulates the criminal offense of bodily harm (in German: Körpverletzung). Such a statutory solution is based on the legal reasoning that a medical procedure, by its legal nature, is a type of bodily injury. Consequently, such interventions, regardless of whether they are medically or non-medically indicated (e.g. interventions in esthetic surgery), always satisfy the actus reus (in German: Tatbestand) of the criminal offense from Article 223 and therefore deciding on the criminal liability of a doctor is essentially transferred from the sphere of the act (which is emphasized e.g. in Croatian law) to the sphere of unlawfulness (in German: Rechtswidrigkeit). Namely, in Germany there is no criminal liability when such an intervention (bodily injury) is performed with the informed consent of the patient and such consent represents the reason for exclusion of unlawfulness sui generis. This consequently means that the main part of the discussion on this issue in German literature and practice concentrates precisely on the problem of informed consent and implied consent, that is, on the existence of some excusable or justifiable reasons such as in an emergency situation (see e.g. Fischer, 2018; Roxin & Schroth, 2010; Hilgendorf, 2016). In Germany the emphasis on this type of criminal offense, along with the right to physical integrity, is placed on the right of the patient to self-determination. With regards to breach of professional standards, even though a small number of authors advocate the understanding under which procedures performed lege artis do not constitute an act of bodily harm, the prevailing perception in German literature is that such interventions fall under the scope of that criminal offense (Hilgendorf, 2016). This gives way to the conclusion that the German system is an example of where breach of professional standards does not have the pervasive meaning it has in, e.g. Croatian law.

Taking into consideration all of the aforementioned, we hold that in the future medical criminal law will need to move from the criterion of professional standards as the only and exclusive one and that it will need to, for the types of situations mentioned here, look to alternative criteria. Otherwise, criminal liability for escalation of an illness, deterioration of health or death of a patient, caused due to an error in treatment realized by systems of autonomous AI, will be difficult to imagine. To further illustrate what is being discussed here, we can draw a parallel with the recent case of a fatal car crash in Arizona in the U.S. caused by an Uber self-driving vehicle (set in autonomous mode where the driver does not participate in operating the vehicle but becomes a passenger, see Lari, Douma & Onyiah, 2015). The vehicle was programmed to detect and recognize other objects in traffic and, accordingly, to undertake the most appropriate maneuver. When the vehicle approached the pedestrian (a moving object), it wrongly identified her as a pole (a stationary object) and did not reduce its speed thus plowing into the pedestrian and killing her (Levin & Wong, 2018; Mrčela & Vuletić, 2018). Therefore, it can be concluded that the maneuver the vehicle undertook (continuing driving at the same speed) was lege artis in the context of the data based on which it made its decision (that maneuver would not have been if it had been based on correct data). In other words, it is possible to note that the emphasis on criminal law protection in instances of consequences caused by autonomous AI systems (whether in medicine or other areas of life) needs to be shifted to the preliminary (preparatory) stages. This can be justified on the dogmatic level with the increased risk of the onset of adverse consequences, sometimes for even a greater number of people.

4.2 *Liability of the person "behind the robot" and problems in suspending causality and predictability of consequences*

If we were to solve the problem being discussed here by simply shifting the liability to the person standing "behind the machine or robot" – such as the manufacturer, programmer, distributor, or the healthcare institution (hospital) using such a system – we would, in my opinion, encounter at least two difficulties. The first of these would be the concept of suspending causality, and the second would be the problem of predictability of consequences.

Earlier it was noted that regardless of which legislation regulating medical errors is adopted, the existence of criminal liability always requires the occurrence of resulting consequences for the health or life of the patient. This then refers to formal criminal offenses in the sense of acts that have consequence in its causality (Novoselec, 2016). With such acts, one of the constitutive elements of crime is the causal connection between the act (or omission) and consequence, and if no connection exists or if it cannot be proven with certainty, and it is assumed that there is none (in dubio pro reo), the defendant should be exonerated. Further, it is necessary to prove that the consequence, and with this the causal connection, was predictable to the perpetrator, taking into account their knowledge and experience, or the average knowledge and experience of persons with similar capacities as the perpetrator (Mrčela & Vuletić). These questions often arise in practice and in many cases criminal liability is absolved or criminal prosecutions are withdrawn in the earlier stages as none of the two described factors are present. To illustrate this, several examples will be presented here. In recent Croatian practice, the wife of the deceased filed a criminal charge against three doctors from the General County Hospital in Vinkovci according to which they had disregarded symptoms that pointed to the need for additional diagnostic imaging tests and failed to send the patient to a larger hospital (to the Clinical Hospital Center in Osijek) where such tests could have been performed, after which the victim died. During the inspection conducted by the State's Attorney Office, an expert evaluation was carried out which established that there had indeed been a professional breach, in the described sense, but that it could not be established with certainty (considering the victim's poor state of health) that the prompt conducting of the tests could have saved his life. Guided by the rule of in dubio pro reo, the State's Attorney Office concluded that there was the probability that the death was caused by the drastic deterioration of the patient's health (which is basically a suspension of causality) and issued a dismissal of the criminal charge (County State's Attorney of Vukovar, no. K-DO-644/2018, KPO-DO 14/2018).

A very interesting example is the matter in which a gynecologist was first convicted of medical malpractice during childbirth in 2004 after seeing signs of fetal distress in the mother's uterus (with every contraction the heart rate of the fetus slowed down and during the water breaking thick green amniotic fluid gushed out), but failing to complete the childbirth earlier by performing an emergency caesarean section, which is prescribed by the profession. During such omission, hypoxic-ischemic brain damage occurred,

resulting in a cerebral palsy diagnosis of the infant. However, during the claim for compensation of damages which was conducted several years later, medical experts established that in the meantime there were new scientific findings in the field of gynecology and obstetrics according to which it was clear that the cause of cerebral palsy occurred during pregnancy and that it had no connection to the birth itself, nor to the gynecologist's negligence. It was also found that diagnostic procedures that were available at the implicated time could not have determined the cause. Based on such findings, the court concluded that on the basis of new evidence there was high degree of probability that there was no causality and it allowed for a retrial of the earlier criminal proceedings, which at the time of writing of this paper is still pending (Municipal Court in Osijek, no. Kv-96/2015, K-422/02).

The problem of causality occurs in German law, in the context of causality of adequate informing on patient consent. An interesting decision is the one where a doctor was convicted under Article 223 when during herniated disc surgery he performed a discography and laser decompression of the nerve root which resulted in paresis and difficulty in leg movement and impotence. It was established that before the surgery the doctor had warned the patient on the risk of paresis, but not on the risk of impotence. The court concluded that with adequate information, bearing in mind the significant harm to the quality of life due to impotence, the doctor could have dissuaded the patient from surgery and led him to choose a more conservative method of treatment, whereby it sentenced the doctor (BGH, decision no. VI ZR 353/99; for a commentary see Schöch, 2010).

If we take into account the described constitutive role of causality and predictability of consequences for the existence of criminal liability of the doctor, it is apparent that this will be a very serious problem if in situations where an autonomous robot-surgeon was the one that caused the consequence and this is attributed to the persons who stand behind such a robot. In comparative law, it is not unusual that criminal liability for consequences that occurred due to defects of a certain (dangerous) product is placed on the manufacturer, programmer or distributor. If, for example, the consequences occurred as a result of negligent handling (handling contrary to instructions), then the criminal liability is attributed to the user under the relevant assumptions (in this case this would be the hospital, if it would be determined that the autonomous system was put into use

contrary to instructions or that it was not maintained properly or regularly). However, the problem here is if we consider that the systems which are being discussed here will be at the highest, fourth level of autonomy. The fact that the system is programmed in such a way as to learn from its environment and that based on the gathered data it makes its own decision also means that the creator of such a system no longer has control over its actions. If the particular system has passed all the necessary testing and received all the necessary permissions before being put to use, then it is very likely that the persons behind the machine, in case of possible criminal proceedings against them, will invoke absence of guilt (absence of negligence) and/or suspension of causality. Gless, Silverman & Weigend (2016) warn that with the development of AI autonomy, the scenario in which "no one will be criminally liable" is quite possible. In Germany this problem has an added dimension, given that local legislation still does not recognize the criminal liability of legal entities. We hold that the preceding considerations resulted in proving the initial hypothesis in which modern criminal law currently does not have sufficient mechanisms to provide adequate protection of legal interests from harm caused by autonomous AI systems. This is particularly true of the problem of protecting people's health and treatment of patients as criminal offenses from this area are based predominantly on the premise that everyone makes mistakes and that the assessment of liability is adapted to the criteria of (human) errors. Since such a situation opens up a significant legal gap and since it is not acceptable that the life and health of a patient, as the highest personal legal right and fundamental constitutional categories remain unprotected, it is evident that criminal law faces a major challenge.

5. What is to be concluded: do we want AI in operating rooms?

In the context of the introductory hypothesis, which argues against the complete delegation of surgery to AI, it is evident from the preceding discussions that both medical and legal foundations support such a conclusion. From a medical (surgical) perspective, it is undesirable to entirely eliminate the human factor because numerous situations arise where the best course of action cannot be predicted or defined in advance, requiring

adequate experience and instinct. A skilled surgeon possesses not only a thorough knowledge of professional guidelines but also rich practical experience and the courage to occasionally rely on their intuition. Another important dimension of the problem is of an ethical nature. Specifically, a surgeon regularly operates according to the principles of medical ethics and deontology in their everyday work. These principles are highly abstract in theory and constitute a set of specific guidelines that should be interpreted depending on the situation, taking into account the broader context. Furthermore, the understanding of ethics can vary from country to country and from culture to culture. AI can never possess a sense of ethics in the way this sense exists in humans, and this can pose a serious problem in the context of making decisions in the best interest of patients and society as a whole.

From a legal standpoint, particularly in countries where legal (especially criminal) protection is based on the concept of accountability for errors, the absolute dominance of AI would, de facto and de iure, render patients practically incapable of obtaining effective legal protection in cases of medical malpractice. If we assume that the overarching purpose of medicine is to ensure the well-being of individuals' health, it becomes evident that the complete dominance of AI in this domain could have potentially adverse effects on the realization of this noble goal. For these reasons, we advocate against the direction of future medical development that would allow AI to entirely take over human tasks. We, therefore, maintain that AI should only be acceptable as long as it functions as an extension of the surgeon's (or general physician's) capabilities.

The human right to health is included in the majority of civilized, modern legal systems today as the highest personal right. In countries with written constitutions, health protection and the right to health are usually regulated at the constitutional level and then elaborated through various legal fields. In the criminal law sense, keeping in mind the role of criminal law as the ultima ratio that protects society from the most susceptible behavior (Jescheck & Weigend, 1996), health protection covers those situations in which a perpetrator acts with an appropriate form of guilt and causes consequences to a person's health.

It can be concluded that the technological revolution is an inevitable process in today's world. The development of AI with a high level of autonomy and independence from human influence is an integral (and very important) part of technological progress. This

is also good and socially desirable as it contributes to the quality of human life. All this, however, does not mean that this process can perform without any control and regardless of the fundamental postulates that have been formed for centuries in civilized societies. That is why legal systems need to continuously work on developing legislation and practices in order to follow the progressive development of technology. The legal field has already been faced with such challenges, first in the area of the already mentioned criminal liability of legal entities, and then in the areas of new forms of criminality which emerged with the development of society and technology. Thus in recent years the obvious global trend (material and procedural) of standardizing new emerging forms of dangerous criminal activities, such as cybercrime, cyberbullying, cyberterrorism, international criminal organizations for hacking, child pornography and child-grooming, etc.

In that sense, criminal law is the last line of defense for the social preservation and integrity of individuals. It will provide the protection of legal rights when other (milder) branches of law cannot do so effectively. The idea of the subsidiarity of criminal law is replicated in this case to the field of medicine (or robotic medicine) in such way as to advocate the introduction of new incriminations which will shift criminal law protection to a sufficiently early stage that it ensures a minimum level of protection for potentially vulnerable subjects, but also the necessary level of legal protection for those engaged in dangerous activities (perpetrators). Consequently, criminal law will not represent a hindering factor which would slow down the development of technology and the progress of humanity as a whole. Rather, it will continue to carry out its natural task of protector of fundamental social values and weigher of opposing interests. Otherwise, if such an approach would be abandoned, society would regress a thousand years back to a period where there was no legal regulation. This would de facto mean that technological progress on the one hand would advance but on the other hand it would significantly set back humankind. Such a scenario, in the nature of things, is wholly unacceptable. We do hope that this paper will simulate further discussions about this very important matter which will certainly be realized in the near future. Such discussions have recently started to take place more intensively at the level of EU legislation.

List of references:

Barthe, C. (2009). Joint Criminal Enterprise (JCE). Ein (originär) völkerstrafrechtliches Haftungsmodell mit Zukunft? (Berlin: Duncker & Humblot).

Brożek, B.; Jakubinec, M. (2017). On the legal responsibility of autonomous machines. Artificial Intelligence and Law. 25 (3). pp. 293-304. doi: 10.1007/s10506-017-9207-8.

Camarillo, M. S., Krummel, T. M. & Salisbury, K. J. (2004). Robotic technology in surgery: past, present, and future. The American Journal of Surgery. 188 (4). pp. 2-15. doi: 10.1016/j.amjsurg.2004.08.025.

Benny Chan, Applying a Common Enterprise Theory of Liability to Clinical AI Systems. American Journal of Law & Medicine, 47(4), 2021, 351-385. doi:10.1017/amj.2022.1

A. R. Di Landro, Criminal Law as a Response to Medical Malpractice: Pluses and Minuses – Comparing Italian and U. S. Experiences, medicine and Law 31 (2012.)

Fischer, T. (2018). Strafgesetzbuch mit Nebengesetzen, 65. Auflage (München: C. H. Beck).

Alberto Galasso, Hong Luo, Punishing Robots: Issues in the Economics of Tort Liability and Innovation in Artificial Intelligence in Ajay Agrawal, Joshua Gans, and Avi Goldfarb (eds.), The Economics of Artificial Intelligence: An Agenda, University of Chicago Press, 2019

Gless, S, Silverman, E. & Weigend, T. (2016). If Robots Cause Harm, Who Is to Blame? Self-Driving Cars and Criminal Liability. New Criminal Law Review. 19 (3). pp. 412-436. doi: 10.1525/NCLR-01-16-0001.

Gumbs, A.A.; Frigerio, I.; Spolverato, G.; Croner, R.; Illanes, A.; Chouillard, E.; Elyan, E. Artificial Intelligence Surgery: How Do We Get to Autonomous Actions in Surgery?. Sensors 2021, 21, 5526. https:// doi.org/10.3390/s21165526.

Puneet Gupta, Heather S. Haeberle, Zachary R. Zimmer, William N. Levine, Riley J. Williams, Prem N. Ramkumar, Artificial intelligence-based applications in shoulder surgery leaves much to be desired: a systematic review, JSES Reviews, Reports, and Techniques, Volume 3, Issue 2, 2023, pp. 189-200, ISSN 2666-6391, https://doi.org/10.1016/j.xrrt.2022.12.006.

Hage, J. (2017). Theoretical foundations for the responsibility of autonomous agents. Artificial Intelligence and Law. 25 (3). pp. 255-271. doi: 10.1007/s10506-017-9208-7.

Hilgendorf, E. (2016). Einführung in das Medizinstrafrecht (München: C. H. Beck).

Jescheck, H. H. & Weigend, T. (1996). Lehrbuch des Strafrechts, Allgemeiner Teil (Berlin: Duncker & Humblot).

Korošec, D. (2016). Criminal Law Dilemmas in Withholding and Withdrawal of Intensive Care. Medicine, Law & Society, 9(1). pp. 21-39. doi: 10.18690/24637955.9.1.21- 39(2016).

Kuhn, M. (1994). Kunst-bzw. Behandlungsfehler, In: Heinrich, H. (ed.). Handbuch des Arztrechts (Zürich: Schulthess Polygraphischer Verlag).

Lari, A, Douma F. & Onyiah, I. (2015). Self-Driving Vehicles and Policy Implications: Current Status of Autonomous Vehicle Development and Minnesota Policy Implications. Minnesota Journal of Law, Science & Technology. 16 (2). pp. 735-769.

Laufer, W. S. & Strudler, A. (2000). Corporate Intentionality, Desert, and Variants of Vicarious Liability. American Criminal Law Review. 37 (4).

Levin, S. & Wong, J. C. (2018). Self-driving Uber kills Arizona woman in first fatal crash involving pedestrian, available at: https://www.theguardian.com/technology/2018/mar/19/ uber-self-driving-car-kills-woman-arizona-tempe (26. 1. 2019.).

Kaspar Rosager Ludvigsen, Shishir Nagaraja, Dissecting liabilities in adversarial surgical robot failures: A national (Danish) and EU law perspective, Computer Law & Security Review, Volume 44, 2022, 105656, ISSN 0267-3649, https://doi.org/10.1016/j.clsr.2022.105656.

Kinoshita T, Komatsu M. Artificial Intelligence in Surgery and Its Potential for Gastric Cancer. J Gastric Cancer. 2023 Jul;23(3):400-409. doi: 10.5230/jgc.2023.23.e27. PMID: 37553128; PMCID: PMC10412972.

Lisacek-Kiosoglous, A. B., Powling, A. S., Fontalis, A., Gabr, A., Mazomenos, E., & Haddad, F. S. (2023). Artificial intelligence in orthopaedic surgery. Bone & Joint Research, 12(7), 447-454.

Markwalder, N & Simmler, M. (2017). Roboterstrafrecht. Zur strafrechtlichen Verantwortlichkeit von Roboten und künstlicher Intelligenz. Aktuelle juristiche Praxis. (2). pp. 171-182.

Mrčela, M. & Vuletić, I. (2017). Granice nehajne odgovornosti za kazneno djelo nesavjesnog liječenja. Zbornik radova Pravnog fakulteta u Splitu. 54 (3). pp. 685-704. doi: 10.31141/zrpfs.2017.54.125.685.

Mrčela, M. & Vuletić, I. (2018). Kazneno pravo pred izazovima robotike: tko je odgovoran za prometnu nesreću koju je prouzročilo neovisno vozilo? Zbornik Pravnog fakulteta u Zagrebu. 68 (3-4). pp. 465-491.

Novoselec, P. (2016). Opći dio kaznenog prava. Peto, izmijenjeno izdanje (Osijek: Pravni fakultet).

Pichler, D. and Vuletić, I. Građanskopravni i kaznenopravni aspekti pogrešaka u medicini, Pravni fakultet Osijek, 2020.

Rasouli JJ, Shao J, Neifert S, et al. Artificial Intelligence and Robotics in Spine Surgery. Global Spine Journal. 2021;11(4):556-564. doi:10.1177/2192568220915718.

Schöch, H. (2010). Die Aufklärungspflicht des Arztes und ihre Grenzen, In: Roxin, C. & Schroth, U. Handbuch des Medizinstrafrechts. 4. Auflage (Stuttgart-München-Hannover-Berlin-Weimar-Dresden: Richard Boorberg Verlag).

Sharkey, N. & Sharkey, A. (2013). Robotic Surgery: On the Cutting Edge of Ethics. Computer. 46 (1). pp. 56-64. doi: 10.1109/MC.2012.424.

Stone, P. et al. (2016). 'Artificial Intelligence and Life in 2030' One Hundred Year Study on Artificial Intelligence: Report of the 2015–2016 Study Panel, (Research Report, Stanford University 2016) (26. 1. 2019.).

Strickland, E. (2017). Autonomous Robot Surgeon Bests Humans in World First, available at: https://spectrum.ieee.org/the-human-os/robotics/medical-robots/autonomous-robot-surgeon-bests-human-surgeons-in-world-first (25. 1. 2019.).

Šepec, M. (2018). Medical Error – Should it be a Criminal Offence? Medicine, Law & Society. 11 (1). pp. 47-65. doi: 10.18690/2463-7955.11.1.47-66(2018).

Taylor, R. H., Menciassi, A., Fichtinger, G. & Dario, P. (2008). Medical Robotics and Computer-Integrated Surgery. Springer Handbook of Robotics. pp. 1199-1222. doi: 10.1007/978-3-540-30301-5_53.

Yip, M. & Das, N. (2017). Robot Autonomy for Surgery. World Scientific Review Volume. 975x65. arXiv:1707.03080 [cs.RO].

J. C. West, Criminalization of medical errors: When is a mistake more than just a mistake?, Journal of Healthcare risk managment 27 (2007.)