JOURNAL OF MODERN SCIENCE

6/60/2024

www.jomswsge.com

DOI: 10.13166/jms/197008

Małgorzata Suchacka
University of Silesia in Katowice, Poland
ORCID iD: 0000-0002-3769-5892

Mariusz Wojewoda
University of Silesia in Katowice, Poland
ORCID iD: 0000-0003-0732-7500

Zbigniew Nawrat
Medical University of Silesia, Poland
ORCID iD: 0000-0003-0638-3789

Artur Domurat
University of Silesia in Katowice, Poland
ORCID iD: 0000-0001-5533-9106

Angelika Magdalena Pabian
University of Silesia in Katowice, Poland
ORCID iD: 0000-0002-8958-4262

TECHNICAL, SOCIAL, PSYCHOLOGICAL AND ETHICAL ASPECTS OF HUMAN-ROBOT COLLABORATION: A CASE STUDY OF THE 'ROBIN HEART' MEDICAL ROBOT

Abstract

Objectives: The purpose of this paper is to emphasize the need for a comprehensive examination of the technical, social, psychological, and ethical aspects of human-machine cooperation, using the Robin Heart medical robot as a case study.

Methods: This research is a literature review and discussion. The conclusions are derived from non-empirical research based on the analysis of secondary sources, specifically scientific literature, within the inductive approach employed to scientific inference. The information sources utilized are interdisciplinary, encompassing management sciences, psychology, sociology, ethics, and medical physics.

Results: The analysis posits those medical robots, including *Robin Heart,* are integral components of a *technology system.* The study identifies interactions among humans (Agent 1: medical personnel; Agent 2: patients) and technological artifacts (Agent 3: the robot). Agent 1 relates to individuals working in medical facilities, who may act as robot creators, programmers, device owners, operators, or robot trainers. Agent 2 refers to patients utilizing medical services. Agent 3 is the AI-equipped robot, which can be viewed either as a significant extension of human capabilities, or as an autonomous entity in medical actions.

Conclusions: The integration of robots into healthcare represents evolutionary transgression from mere usage to active cooperation with machines. Enhancing the human-machine relationship is likely to be a central focus for multi-disciplinary teams that integrate humanistic, biological, and technical approaches in the forthcoming years. The findings underscore the importance of fostering an awareness of the technological system through a holistic view that incorporates interdisciplinary advancements.

Streszczenie

Cel: Celem niniejszego artykułu jest podkreślenie potrzeby wszechstronnej analizy aspektów technicznych, społecznych, psychologicznych i etycznych współpracy człowieka z maszyną, wykorzystując jako studium przypadku robota medycznego Robin Heart.

Metody: Niniejsze badanie jest przeglądem literatury i dyskusją. Zaprezentowane w artykule wnioski stanowią efekt badań niereaktywnych w oparciu o analizę źródeł wtórnych – literatury naukowej. Metoda wnioskowania naukowego miała charakter indukcyjny. Wykorzystane źródła informacji są interdyscyplinarne, obejmują nauki o zarządzaniu, psychologię, socjologię, etykę i fizykę medyczną.

Wyniki: Analiza zakłada, że roboty medyczne, w tym *Robin Heart*, są integralnymi składnikami *systemu technologicznego*. Badanie identyfikuje interakcje między ludźmi (Agent 1: personel medyczny; Agent 2: pacjenci) i artefaktami technologicznymi (Agent 3: robot). Agent 1 odnosi się do osób pracujących w placówkach medycznych, które mogą działać jako twórcy robotów, programiści, właściciele urządzeń, operatorzy lub trenerzy robotów. Agent 2 odnosi się do pacjentów korzystających z usług medycznych. Agent 3 to robot wyposażony w sztuczną inteligencję, który może być postrzegany jako znaczące rozszerzenie ludzkich możliwości lub jako autonomiczny podmiot w działaniach medycznych.

Wnioski: Integracja robotów z opieką zdrowotną stanowi ewolucyjną transgresję od zwykłego użytkowania do aktywnej współpracy z maszynami. Poprawa relacji człowiek-maszyna prawdopodobnie będzie centralnym punktem zainteresowania dla zespołów multidyscyplinarnych, które integrują podejścia humanistyczne, biologiczne i techniczne w nadchodzących latach. Wyniki podkreślają znaczenie promowania świadomości systemu technologicznego poprzez holistyczną perspektywę, która obejmuje postęp interdyscyplinarny.

Keywords: *medical robots, human-machine relations, social system, new technology*

Słowa kluczowe: *roboty medyczne, relacje człowiek maszyna, system społeczny, nowe technologie*

Introduction – the purpose of analysis

Implementing innovative technological solutions provides measurable benefits for the economy, society, and the environment, and poses new development challenges. Humanity is currently witnessing the expansion of modern, digital medical technologies aimed at raising standards, increasing access to medical services, and introducing personalized optimization of patient care, treatment, and rehabilitation. The greatest practical and functional advancements of new technologies in medicine are achieved, among others, in diagnostics and surgery. Artificial intelligence (AI) is one of the main drivers of the development of diagnostic methods and surgical procedures used in the patient treatment process. The application of these solutions is not only a subject for analysis for experts in IT or technical solutions but, which is sometimes forgotten, an important issue for ethicists, sociologists, psychologists, and management specialists. Among the most intriguing topics in these dimensions, the issue of collaboration between medical machines and humans comes to the fore.

The article aims to present the main principles of the use of the *Robin Heart* surgical robot – its advantages and potential use, as well as the issues related to further research on its improvement. Around these issues, the threads of interest to researchers in social and human sciences related to the social, psychological, and ethical aspects of human interaction with a medical robot will be highlighted.

Research methods and own theoretical assumptions

The relationship between a human and a medical robot can be subject to hermeneutic interpretation. The article is of a review and discussion nature. The conclusions presented in the articles contain the results of non-reactive research based on reliable secondary sources – scientific literature. The method of scientific reasoning was inductive. We recognize that Robin Heart (Agent 3) is an element of a *technology system* operating in a medical institution, in which important agents, in addition to doctors (Agent 1) and patients (Agent 2), are engineers and computer scientists responsible for creating and training the skills of the AI-based robot. We consider the analysis of human potential related to the performance of the previously indicated roles as well as the capabilities and limitations arising from the functions and skills of the robot. The authors' considerations in this article are primarily theoretical, involving the analysis of the complex aspects of the human-machine/robot/AI relations from psychological, ethical and social perspectives (Dusek, 2006). In this case, a mutual impact is assumed, meaning that the use of robots in medical practices changes humans, their attitudes, habits and calculations of what is possible or impossible to accomplish. It seems that over time this will affect the understanding of the value of caring for health. The use of robots and AI will become a benchmark for properly understood care for the health and life of patients. At the same time, it will require consideration of several aspects of their functioning, including technical, psychological, social and ethical ones.

Robots are increasingly autonomous and capable of performing complex tasks without the need for human supervision or continuous human input. In some cases, they engage in human collaboration that involves new forms of human-machine-human interaction, where robots behave as equal team partners. (Fig. 1)

Figure 1. *Model of the relationship agent 1-agent 2-agent 3*

Source: *own elaboration.*

Humans are increasingly aware of the fact that they have a valuable *partner* (agent 3) who complements their abilities in shaping the environment and influencing the quality of their lives. Health is a valuable asset; hence the acceptance of autonomous medical robots raises concerns among patients and some segments of the medical community.

The issue of work autonomy of the robots and their collaboration with humans can be considered in five aspects: (1) independence from the power source – so far it is difficult to imagine a robot independent of the power source (2) *extension* of human abilities: for example, the arm of a medical robot can *perform* actions with greater precision than the human hand. (3) the robot receives a task from a human, so Agent 1 determines the ways of accomplishing this task. (4) the robot receives a task from agent 1 but independently chooses the ways of its implementation. The human surgeon controls the execution of the task. 5) the robot independently determines the tasks to be performed and chooses the ways of its implementation. The human surgeon only controls the result of the task performed. (Glaser, Rossbach, 2011).

Terms and principles of operation of medical robots – Robin Heart project

A robot, distinct from an automaton, represents a smart fusion of sensory input (perception) and physical output (action). It includes a mechanical arm, a control system, and specialized software. Rooted in the idea of creating robots as artificial humans, as first introduced by Čapek in 1920, modern robots combine the capability to perform mechanical and intellectual work (1920). With advancements in machine learning and its practical applications, artificial intelligence (AI) now plays a significant role in robotics, including medical robotics. This is particularly crucial in the field of medical robotics, as AI also plays a vital role in facilitating communication between robotic systems and human intelligence.

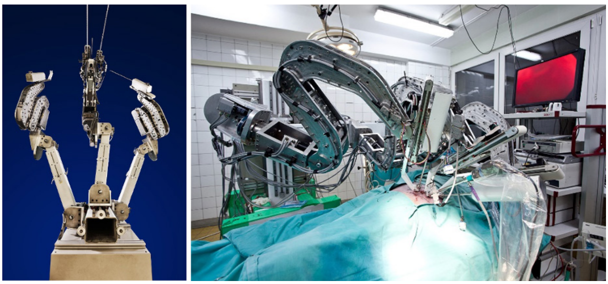
Medical robots can be categorized into several types, including diagnostic robots, surgical robots, caregiving and social robots, rehabilitation robots, rescue robots, artificial organs, biorobots (scientific robots mimicking humans or animals for research purposes), and educational medical robots.

During surgery, the surgeon employs both physical fitness and coordination skills. Robots currently used in surgery are telemanipulators. The surgeon, remote from the patient, sitting in front of the console monitors, uses appropriate user interfaces (joysticks) to move the robot tools inserted through small holes in the patient's body. The ergonomic design of the robot control console aids in overcoming physical limitations often encountered in conventional and laparoscopic surgery. An ideal telemanipulator should effectively utilize the operator's motor skills. It operates by replicating the surgeon's movements (surgeon-agent 1) to control the robot arm and tools by generating appropriate control signals for its drives. This control system ensures the required accuracy, enables scaling of set values to enhance positioning precision, and mitigates the effects of hand tremors. Additionally, the robot's control system allows for the execution of specific tasks by the robot's effector. In case of signal transmission system failures or mechanical malfunctions, it is crucial to establish a management hierarchy to ensure patient and staff safety. In such situations, medical professionals may revert to conventional methods.

The Foundation of Cardiac Surgery Development, set up by Professor Zbigniew Religa in Zabrze, Poland (FRK), stands at the forefront of medical robotics in Poland, specializing in the development of cardiac prosthetics and surgical procedures. The Robin Heart family, a series of Polish surgical robots tailored for heart and soft tissue surgeries, is a notable creation of FRK. This series includes the early models Robin Heart 0, 1, and 2 from 2000-2003, the video surgical Robin Heart Vision from 2007-2008, the modular Robin Heart mc^2 since 2009, and TeleRobin from 2014 with its innovative tool platform (Fig.2 A, B). The Robin Heart mc^2 was recognized as the world's largest surgical robot, featuring three arms – two serving as assistants and a central arm bearing an instrument platform. The efficacy of the Robin Heart system has been validated through animal trials (Nawrat, 2020, FRK1, FRK2).

Figure 2. *Robin Heart Tele and Robin mc^2 – archive of Professor Zbigniew Religa Foundation of Cardiac Surgery*

A. Robin Heart Tele and original, patented robot motion controller with force feedback



B. The largest of the Polish robot family, Robin Heart mc^2, and its tests on animals at the Center for Experimental Medicine of the Silesian Medical University in Katowice in 2009. A robot that, controlled by a single surgeon, could perform the tasks of three people at the operating table: the main surgeon and two assistants

Source: *archive of Professor Zbigniew Religa Foundation of Cardiac Surgery Development.*

When controlling the telemanipulator, the operator processes various information sources, including the patient's digitized medical history, diagnostic data, real-time anatomical information, feedback on tool-tissue interactions through visual observation or force feedback, real-time vital organ data during surgery, educational knowledge, and experience. The robot, or more precisely a telemanipulator, is the first tool for surgeons that enables the direct application of surgical simulation and planning methods (Nawrat, 2011).

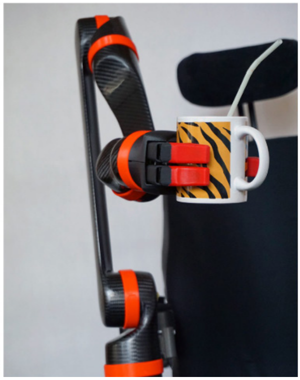
The surgeon (agent 1) – telemanipulator operator – is responsible for all decisions and work methods. The telemanipulator uses a human (brain, motor skills) in the control system loop. When controlling the telemanipulator, information is processed. To enhance perception capabilities during telemanipulation, additional information based on sensory measurements is introduced. Force feedback is a typical example of such work aimed at increasing safety and precise control of tasks performed by tools. In this case, the robot operator observes the movement of the tools they control on the monitor but also feels in his hand (haptic motion controller) the force of the tool's interaction with the tissues (agent 1 – agent 3). This creates a new type of human participation in the information loop of the telerobot system. (Nawrat, 2019).

If we want the robot to work independently, autonomously, we must equip it with systems collecting data from the environment (sensors) and a decision-making system (AI). Robot control involves perception, data processing, and action. Access to information and the ability to analyze it are essential for robots to make informed decisions. Sensors are responsible for acquiring environmental information and monitoring the robot's components and drive systems' current states. Artificial intelligence contributes to error reduction, improved surgical standards, enhanced surgical quality, and increased patient safety.

Surgical operations are teamwork. Progressing robotization, replacement of subsequent team members by robots, and the need to coordinate the operation of devices provoke the need for direct communication between devices and robots (Internet of Things IoT). Already today, a range of medical services are supported by robots, enhancing precision, expanding the pool of patients for whom surgical interventions become feasible, replacing humans in tasks that pose health risks to personnel or are economically unviable. Robots assist in rehabilitation, sterilize rooms, and serve as a crucial supportive arm for individuals with disabilities (Fig. 3). All these tasks require interaction with humans and their environment.

Figure 3. *Examples of medical robots*

A



*A. **Polish robot LUNA EMG** is an example of a robot used for limb rehabilitation that adapts its operation to the patient based on muscle activity measurement (EMG – Electromyography). This interactive approach enables the robot to provide support when needed while encouraging the patient's active participation in the rehabilitation process, fostering better outcomes through human-robot collaboration*

Source: https://www.egzotech.com/pl/luna-emg/ access 24.11.2024

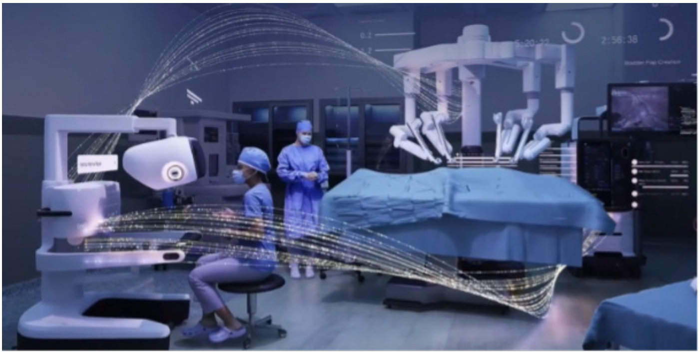
*B. **ARIA – Wheelchair** robot mounted on a wheelchair provides support in everyday activities – eating, drinking, opening the door, etc. It is ultra-light, safe, and very compact which provides a high level of convenience and independence*

Source: *https://accrea.com/aria/* access 24.11.2024

*C. **The Symani** robot is the first of its kind robotic system used for performing micro and super microsurgical procedures on blood vessels, nerves, and lymphatic vessels.*

Symani combines the world's smallest 3-millimeter wristed microinstruments with advanced technologies for tremor reduction and motion scaling (capable of slowing down the operator's hand movements by up to 20 times). This telemanipulator robot is an excellent example of how the use of robots in medical practice transforms humans, their attitudes, habits, and calculations of what is practically achievable

Source: *https://synektik.com.pl/produkty/roboty-mikrochirurgia-symani/* access 24.11.2024



*D. **Da Vinci 5 (Intuitive Surgical)** is easy to learn, with simplified setup, guided tool change, and task automation. Surgeons with varying experience levels can deliver up to 43% less force on tissue by sensing the push and pull forces at the instrument tip through Force Feedback technology. The new generation da Vinci robot offers fully digital documentation of every task performed*

Source: *https://www.intuitive.com/en-us/products-and-services/da-vinci/5* access 24.11.2024.

Artificial intelligence, which evolved from human intelligence, is based on the experience of its predecessors, standards and diagnostic data, introduced into the process of medical care and therapy is an important element of such a team because it allows communication between humans and robots, and communication between machines and the team with patient (updated assessment of the patient's condition, surgical field, etc.), increases the safety and quality of the service. The AI agent is a real member of the team. Literature studies on the development of medical robotics shows that the cooperation of various technical and humanities specialties in the design thinking process is very important because we create tools that will work independently (another higher level of evolution) and universally (as we expect) (Nawrat, 2019). Medical robots represent the only chance to achieve broad, democratic access to good quality healthcare services shortly.

Robotization as a trend in global management and social development

Robots, autonomous systems and artificial intelligence are part of the emerging Industry 4.0 and the evolving human work environment. From the social sciences point of view research on the technomorphization of humans and anthropomorphization of robots are very interesting (Szpunar, 2023). The progressing digitization, automation and robotization are a signal that the world is evolving, which will result in groundbreaking changes that may occur in just a few years. It proves that it will be necessary to redefine many phenomena or processes or design new business models (Zakręcki, Zawada, 2024).

AI is increasingly used in medicine, e.g., algorithms and software that analyze data and digital information to diagnose and suggest therapies. Increasing indicators proving the growing level of robotization in many areas of human activity (including medicine) significantly impact changes in the student education process. Research conducted by J. Kimmerle, J. Timm, T. Festl-Wietek, U. Cress, and A. Herrmann-Werner indicates that medical students must have knowledge and skills of practical cooperation with AI and the ethical aspects of this cooperation. It is necessary to make future doctors aware of the limitations of the algorithms used and the need to monitor the results of their work constantly. Study program changes are met with students' interest, as evidenced by the results (Kimmerle, Timm, Festl-Wietek, Cress, Herrmann-Werner, 2023). Other studies on attitudes toward new-generation technologies indicate that the implementation of solutions based on AI in medicine requires its promotion by superiors, including training and workshops (Lin, Tu, Hwang, Huang, 2021).

It is an undeniable fact that robots will be implemented on a massive scale shortly, but how quickly the adaptation process on the human-machine line will proceed should be discussed and planned today. (Libert, Cadieux, Mosconi, 2020). Any social integration of AI into the human social system would require some form of relations between humans and AI at various levels in society. This means that humans will always actively participate in certain decision-making loops that will affect (AI) operations. Many autonomous artificial intelligence programs raise concerns and lack of trust in the proper operation of the system, like, for example, autopilot in an aeroplane. Similarly, many programs (AI) working with the use of measuring sensors and exceeding the permissible values of the set parameters are endowed with limited trust. There is also limited trust in the use of artificial organs in medical procedures (Nawrat, Religa, 2009).

Therefore, it is necessary to take actions that will reduce the concerns of human employees regarding collaboration with digital employees (agent 1 – agent 3). In the context of many anticipated changes and initial technological disruptions, designing *employee experience* becomes incredibly important. This experience will make it easier to convince employees that technology has been created to play only an additive role, and thereby it will affect work efficiency, faster and more accurate decision-making, or it will create space for their creative and innovative solutions, without replacing people and without causing them to be pushed to the margins of the labour market (Kaczmarek, 2020). Enterprises are becoming increasingly aware in this respect, feeling the burden of maintaining a balance between social expectations and introduced innovations. The answer is a new trend in corporate social responsibility, i.e., Corporate Digital Responsibility (CDR), (Suchacka, 2020).

Robots equipped with AI are already a very important supplement to human activities. Over time, robots will become active agents in medical actions, remotely controlled, and capable of undertaking highly specialized tasks based on the information collected from the environment using sensors, transmitting it to network resources, and allowing other devices to utilize those resources.

Uncertainty and trust in technology from the perspective of the psychology of decision-making

It is a kind of paradox that when we seek assistance from highly autonomous machines, we must also consider how to control their actions. On the one hand, we desire that robots can *think*, recognize unusual situations best, and respond effectively, even without our awareness. On the other hand, we cannot allow them to be too innovative in problem-solving. Using autonomous systems leads to a situation in which the human decision-maker (agent 1) may feel uncertain and have limited trust in the collaborating machine. These feelings are closely related. The essence of the functioning of autonomous machines, AI systems, and robots is their action with many degrees of freedom. Therefore, by definition, they are capable of generating unpredictable outcomes.

Decisions made by robots rely not only on the quality and quantity of input data, which depends on the user but also on the implemented algorithms and rules, which can blur responsibility. For example, who should be held responsible in an autonomous car accident may be debatable. Determining accountability may involve the manufacturer, programmer, owner, driver, or another road user. Similar dilemmas could emerge regarding the responsibility for operating a medical robot and the occurrence of unforeseen, adverse consequences. The human factor matters: if the system can autonomously decide based on prior learning, there is a risk that robots might mimic human behaviors, copy human habits, choose inappropriate actions, and develop ineffective heuristics. The human mind works on the principle of minimal effort, and humans typically strive for cognitive parsimony (Fiske, Taylor, 1991). Because analyzing information and making choices requires time and effort, humans violate the principles of rationality and are ineffective in achieving their goals (Evans, Over, 1996), making errors and being subject to cognitive biases (Baron, 2000; Kahneman, Slovic, Tversky, 1982).

Recognizing own cognitive limitations is not easy: some tasks seem difficult to people but prove to be easy when explained (Stanovich, 2009). Such tasks are usually trivial for machines that can verify and implement even very complex logical rules quickly and automatically. The problem can arise from both a lack of trust in robots and excessive routine in human activities (agent 1). On the one hand, excessive trust can lead to a loss of sense of when to take control of tasks performed by the robot. On the other hand, people may distrust a robot because of their disbelief that a machine can outperform human accuracy and efficiency. In such situations, they may inefficiently utilize a robot's capabilities. Studies indicate that smooth human-robot interaction positively correlates with the operator's work performance and satisfaction, while work overload weakens these relationships (Paliga, 2023). Such work overload might result in misinterpretation of signals generated by the robot.

Decision sciences typically categorize decisions based on their conditions: certainty, uncertainty, and risk (Luce, Raiffa, 1957; Białek, Domurat, Meyers, 2021).

Decisions under certainty refer to situations where the results are predictable from actions, and the choices always lead to determined results, as they are conditioned by the implemented method. For example, adjusting simple tools influences the nature and outcome of their use. A risky situation occurs when the results of the actions taken may be different, but the decision-maker knows the chances of their occurrence. For example, such circumstances usually accompany the production of various things manufactured with a known probability distribution, describing the dispersion of outcomes around the assumed standard. These situations usually require some form of risk management, such as maintaining machinery or adjusting the size of the items produced. Similarly, the effects of a robot's action, like any mechanical system, also have some variance. However, it is generally low: robots ensure precise manipulation, and using artificial intelligence methods enables adaptation and increases safety. For example, it is possible to position the robot's arms with high precision, estimate the frequency and magnitude of possible delays in reaction or communication between the operator and the robot, etc. Gathering and analyzing data on the environment and the robot's condition allows for minimizing errors in this area by appropriate service and calibration. Obtaining information from sensors enables better assessment and selection of actions during operations.

Uncertainty arises when the chances of obtaining different results for the considered actions or decision alternatives are unknown. This typically prompts gathering relevant empirical data to transform the uncertainty into risk. For example, economics and business aim to obtain information from empirical studies, marketing research, simulations, or historical data. However, psychologically, uncertainty can be understood more broadly, usually signaling various negative mental states (Domurat, Zieliński, 2013). In these situations, a decision-maker has to actively formulate hypotheses about the states of nature, think productively, and spend more time and effort to acquire and process information, with no guarantee of its completeness.

AI systems excel at active, ongoing collection and processing of large information volumes and generating ad hoc solutions in open situations. However, the autonomy of robots may raise doubts in the operator as to whether the autonomous system accurately assesses the situation and selects actions. This uncertainty amplifies when it is difficult to predict in advance the elements and behaviors of the robot, its structural properties, and limitations that may lead to malfunction and pose a threat to the environment and people. If the robot's actions go beyond the typical routine, it may be difficult for the operator to decide whether it is the correct reaction of the robot in an unusual situation or, for example, a system error or external interference in the form of a cyber-attack. The decision-maker may also be dismayed when the robot's sensors or systems suggest mutually exclusive actions. In such situations, where the robot's actions veer from the typical routine or when its sensors or systems provide conflicting recommendations, it becomes critical to discontinue the robot collaboration and revert to conventional methods.

Moreover, integrating generative artificial intelligence (AI) into medical robots introduces new dimensions to the psychological aspects of uncertainty and trust. Generative AI enables robots to autonomously generate solutions and responses beyond their initial programming, which can enhance their ability to handle complex and unforeseen medical scenarios. However, this increased autonomy can also exacerbate feelings of uncertainty among medical professionals (Agent 1). The unpredictability of AI-generated decisions may lead to doubts about the robot's actions, as operators might struggle to anticipate or understand the machine's reasoning processes. This *black box* nature of generative AI can undermine trust, making it challenging for practitioners to confidently rely on the robot's assistance, especially in critical situations where patient safety is paramount.

Functioning of new technologies – ethical issues and human dilemmas

For reflection on the proper use of medical robots, we refer to one of the currents of contemporary ethics, i.e., the ethics of responsibility. Theoreticians of this trend include Hans Jonas, Georg Picht. These researchers were interested in the issue of assessing the impact of technology artefacts on human life (Jonas, 1984). The solutions they propose can be successfully applied to the development of principles regarding the ethical use of medical robots and AI.

Generative AI is based on advanced deep learning models such as GPT (Generative Pre-trained Transformer), DALL E, AlphaFold, which use huge data sets to generate new content and new applications of previously known solutions. AI enables automation of decision-making processes, solving complex problems. Modern applications of generative AI are becoming important for innovation in areas such as finance, education, production, and healthcare. On the one hand, the development of AI technologies creates new opportunities, such as personalization of services, support for decisions made by specialist doctors, medical matching of potential donors and recipients of organs (Cohen, Nunes 2024). On the other hand, the use of AI is associated with threats related to the security of personal data, illegal trade in medical information about patients, unpredictable social effects. In the legal and ethical sense, this requires the creation of appropriate risk management strategies, and constant monitoring of the possible effects of using these technologies. For this reason, it is crucial to develop flexible legal standards and management systems that, on the one hand, will enable the full use of AI's potential and, on the other hand, minimize negative effects, as well as protect patient privacy and safety (Agent 2). It is postulated here to ensure the confidentiality of patient data and the transparency of algorithms used in machine learning and deep learning processes. In this field of AI applications, it is worth considering the involvement of patients in decision-making processes regarding their health, and the implementation of AI in medical techniques, etc. (Pawlak 2021).

We refer the principle of responsibility to a morally conscious entity. Admittedly, we use the term *robot ethics* (a term used by the originator of robot ethics, Isaac Asimov). The term *robot ethics* should be understood as a set of procedures developed for robots, but devoid of moral consciousness. In this context, intelligence must be distinguished from consciousness. AI is a technical creation of a human, operating based on a model observed and properly trained by humans. The effect of technical *mimicking* human behavior is an algorithm, or a set of algorithms created by the programmer. Currently, AI is also an important element of the technology system affecting the activities of individuals and the functioning of institutions and enterprises (Campa, 2011).

The principle of responsibility applies to conscious moral subjects. Therefore, the ethics of responsibility concerning medical robots is the ethics of agent 1 (functioning as a creator, software creator, robot owner, user, robot trainer) concerning agent 2 (patient). A very important aspect of this relationship is Agent 3, understood as a technical extension of Agent 1's activities.

The problem of responsibility concerning the use of medical robots should be considered in two aspects: legal liability and moral liability. The first, i.e., legal liability, concerns the assessment of the effects of the actions of Agent 1 involved in an operation using a medical robot (Wojewoda, 2015). We bear in mind events that have already taken place (post-factum). Legal responsibility should be understood concerning institutions and people employed in the hospital: doctors, IT specialists, robot skills trainers, quality controllers of its proper functioning. However, the issue of legal liability requires separate analyses. Moral responsibility as a category related to consciousness (agent 1) indicates a felt duty that precedes action (Feenberg, 2002). As part of conscious responsibility, the legitimacy of the risk taken is analyzed. Then a situation of moral dilemma is also faced. Is the risk associated with the use of the robot proportional to the risks associated with the loss of health and life of the patient? To what extent does the introduction of a medical robot for medical activities affect the effectiveness of actions taken by people, and in what sense does it cause new types of threats? Is the error of the robot an error of a human functioning in one of these roles, or an error of the machine? At the current stage of technological knowledge, it should be recognized that only human is capable of solving these dilemmas. Just because *robot ethics* is talked about does not mean that a robot is capable of solving difficult medical situations (Dazeley, Vamlew, Foale, Young, Aryal, Cruz, 2021). The manufacturer's certificate is not a confirmation of having appropriate medical skills and proper operation of the robot.

About the use of a medical robot by agent 1, the following principles should be adopted: 1) non-harm (referring to the basic principle of Hippocratic medical ethics and Asimov's ethics of robots) – robots are to serve the good of humans in the best possible way; 2) limited robot autonomy – the machine must be subject to human control when it comes to choosing the purpose of action. The autonomy of a robot may concern the choice of means to achieve a goal set by a human; 3) the principle of the patient's voluntary consent to participate in surgery involving a robot – the patient's medical relationship with the robot cannot be forced; 4) the principle of equal access – no group or person may be discriminated against in the possibility of using the services of a medical robot; 5) protection of human rights – affirmation of the dignity of the human person and protection of the patient's data.

Research limitations

Limitations of this study include its theoretical focus and the absence of empirical data. The considerations presented in this article are mainly the result of the analysis of secondary sources, but also of experiences in the field of work with the Robin Heart medical robot. Therefore, the study's main limitations should be related to its theoretical nature. The authors limited the scope of the analysis to 3 aspects, considering cooperation between humans and medical robots from a psychological, ethical, and social perspective. Additionally, the article mainly presents the perspective of the machine operator, and the experiences focus on the work of a team of medics with the Robin Heart robot. These experiences result from providing medical services to many patients, but the size of the sample (number of patients using medical services with AI support) may also be a limitation. It should be noted, however, that this is related to time and space limitations and qualification for the procedure (indication for a specific group of patients). With the next version of the implemented robot, the threads necessary to include in the analysis will multiply, and the team is ready to attempt to analyze them in the future and present a wider spectrum of variables, considering the opinions of two agents – medics and patients.

The year 2023 has been declared the year of AI, indicating the growing importance of machine learning in performing tasks previously assigned to humans. The accelerating trend related to the implementation of further innovative AI solutions may indicate new directions of research, reveal further important threads, but above all give more case studies considering variables that are not indicated in this study, mediating in the model of human-machine communication, in particular interfering variables. Future research should involve empirical studies assessing how generative AI in medical robots affects trust and decision-making among healthcare professionals.

Such studies could inform the development of training programs and guidelines to enhance human-robot interaction, ensuring that medical robots are integrated into healthcare settings effectively and ethically.

Concluding remarks

New technologies, including medical robots, will be a very important aspect of our understanding of health care soon. However, they will not replace the human factor. Including the elements indicated in the study in the *system of technology* is an implication – not only for theory but also for empirical research, resulting from the conducted considerations. Such important areas of our lives as life and health require compassion, meetings, attention, dedication, support, good organization and orientation in possible solutions. Medical robots are an important *extension* of human capabilities to help sick people. In the conducted considerations, apart from presenting the principles of operation of the Robin Heart medical robot, important aspects accompanying the functioning of medical machines in the social, psychological and ethical dimensions were indicated.

The main conclusions of this study can be summarized in several basic remarks:

- in the technological dimension – the Robin Heart medical robot is a modern tool prepared to support the work of doctors – and surgeons. Robots, or rather a computer acting as an intermediary between the decision and its implementation, create an opportunity to introduce ad hoc systems of control, and supervision, presenting advice (what action will be most effective) or blocking risky action, linking the action with planning and diagnostic database. Digitization of information, both diagnostic (patient before surgery) and during surgery (patient's condition, image of the surgical field, but also the operating status of the robot, sensors) allows for personalization of actions taken and the introduction of an active improvement mechanism (machine learning, artificial intelligence). This example shows that although we started with a simple tool that is simply more precise and allows for less invasive operations, we have come to a robot (or robots) as part of a system whose subject is the patient; from the emergence of a need, through the provision of services in a hospital, to home care (e.g., rehabilitation with the help of other robots).

- in the social dimension – new challenges are emerging because in recent years there has been an increase in interest in broadly understood advanced technologies, including medical robots, which bring many advantages. Important medical centres, hospitals and clinics as well as scientific and research centres, and universities in Poland and Europe are the basis for education, the creation of innovations and developing places where advanced technologies can be implemented. Organizations will have to make an effort to train employees and develop the skills necessary for the ongoing technological transformation. In the context of new challenges, more and more attention is paid to the need to create new professions in which the doctor will be required to combine medical, social and psychological competencies with technical ones. It is also necessary to educate medical staff about the use of AI algorithms in their work and to make future doctors aware of the advantages and limitations of human-robot cooperation.
- in the psychological dimension, the issues of fear and lack of trust in new technologies are significant. People tend to have less confidence in AI and intelligent machines when they feel uncertain about their use. The potential incorporation of generative AI in medical robots intensifies these feelings, as the unpredictability of AI-generated decisions can further erode trust. Factors contributing to uncertainty and distrust include cognitive limitations, which can lead to an inadequate assessment of the capabilities and effects of both human and robot work. Addressing these concerns requires improving transparency in AI systems and enhancing the understanding of AI processes among medical professionals to foster better human-robot collaboration.
- in the ethical dimension – a certain regularity can be noticed that with the increase in the use of new technologies in the treatment process, the need to remind and emphasize the role of ethical principles in the work of doctors also increases. Due to the use of medical robots, it will be the work of doctors but also of engineers, programmers, and trainers of intelligent robot skills.

The presence of a human doctor during surgery is still necessary. Human qualities, such as the ability to act flexibly, intuition, and the ability to respond to unpredictable situations, cannot be replaced by the operation of an intelligent machine. The use of robots is another consequence of human evolution. From the stage of ordinary use, we move on to the stage of collaboration with them. For this cooperation to be possible and optimal for humans, it is necessary to know the human and the capabilities of the robot well enough. Improving the human-machine relationship will probably dominate many of the coming years of activity of multidisciplinary teams combining humanistic, biological and technological approaches.

References

Baron, J. (2000). *Thinking and Deciding (3rd ed.)*. Cambridge University Press.

Białek, M., Domurat, A., Meyers, E.A. (2021). *Decision-Making*. In: V.P. Glăveanu (ed.) *The Palgrave Encyclopedia of the Possible*. Cham: Palgrave Macmillan.

Campa, R. (2011). *Kodeksy etyczne robotów: zagadnienie kontroli sprawowanej przez człowieka*. Pomiary. Automatyka. Robotyka. 3/201, 66–70. https://yadda.icm.edu.pl/baztech/element/bwmeta1.element.baztech-article-BSW1-0078-0022?printView=true

Čapek, K. (1920). *R.U.R. (Rossum's Universal Robots) A Fantastic Melodrama in Three Acts and an Epilogue,* R.U.R. (Rossum's Universal Robots) by Karel Čapek – Free Ebook (gutenberg.org).

Cohen, J., Nunes, P. (2024). *Socioeconomic Factors Influencing the Adoption of Generative AI Technologies in Small and Medium Enterprises,* Journal of Business and Technology, nr 12(4), p. 233-245.

Dazeley, R., Vamlew, P., Foale, C., Young, C., Aryal, S., Cruz, F. (2021). *Levels of explainable artificial intelligence for human-aligned conversational explanation.* Artificial Intelligence 299, 1-29. https://doi.org/10.48550/arXiv.2107.03178

Domurat, A., Zieliński, T. (2013). *Niepewność i niejasność jako uwarunkowania wybranych decyzji ekonomicznych. [Uncertainty and ambiguity as the settings of economic decisions].* Decyzje, 10 (20), 21– 48. https://doi.org/10.7206/DEC.1733-0092.10

Dusek, V. (2006). *Philosophy of Technology. An Introduction*. Blackwell Publishing Ltd. Oxford.

Evans, J. S. B. T., Over, D. E. (1996). *Rationality and reasoning*. Psychology/Erlbaum (UK) Taylor & Francis.

Feenberg, A. (2002). *Transforming technology. Critical Theory Revisited*, Oxford University Press. Oxford–New York.

Fiske, S. T., Taylor, S. E. (1991). *Social Cognition* (2nd ed.). McGraw-Hill Book Company.

FRK1. (2009-2012). Professor Zbigniew Religa Foundation of Cardiac Surgery Development, *Projekt* pt. *Projekt, konstrukcja, badania i optymalizacja interfejsu człowiek – robot chirurgiczny. Uniwersalna konsola sterowania telemanipulatorem Robin Heart oraz stanowiska treningowymi chirurgii małoinwazyjnej w środowisku fizycznym i wirtualnym. Narodowe Centrum Badań i Rozwoju.*

FRK2. (2012-2015). Professor Zbigniew Religa Foundation of Cardiac Surgery Development, Projekt pt: *TeleRobinSurgery – opracowanie i badania nowych rozwiązań technicznych dla zdalnie sterowanych operacji chirurgicznych za pomocą robotów Robin Heart*. Program Badań Stosowanych NCBiR.

Glaser, H.A., Rossbach, S. (2011). *The Artificial Human*. Peter Lang, Frankfurt/M., Bern, New York, p. 82– 84.

Jonas, H. (1984). *Imperative of Responsibility. In Search of an Ethics for the Technological Age.* Chicago: The University of Chicago Press.

Kaczmarek, P. (2020). *Szybki, efektywny i nie robi błędów. Korzyści robotyzacji procesów biurowych,* Personel & Zarządzanie, 3(360), pp. 34-36.

Kahneman, D., Slovic, P., Tversky, A. (Ed.). (1982). *Judgment under Uncertainty: Heuristics and Biases,* Cambridge University Press.

Kimmerle, J., Timm, J., Festl-Wietek, T., Cress, U., Herrmann-Werner, A. (2023). *Medical Students' Attitudes Toward AI in Medicine and their Expectations for Medical*

Education. Journal of Medical Education and Curricular Development, 10. https://doi.org/10.1177/23821205231219346

Libert, K., Cadieux, N., Mosconi, E. (2020). *Human-Machine Interaction and Human Resource Management Perspective for Collaborative Robotics Implementation and Adoption,* Proceedings of the 53rd Hawaii International Conference on System Sciences (HICSS 2020; pp. 533-542). Maui, Hawaii, USA. https://doi:10.24251/HICSS.2020.066

Lin, H.-C., Tu, Y.-F., Hwang, G.-J., Huang, H. (2021). *From Precision Education to Precision Medicine: Factors Affecting Medical Staff's Intention to Learn to Use AI Applications in Hospitals*. Educational Technology & Society, 24 (1), 123-137. https://www.jstor.org/stable/26977862

Luce, R. D., Raiffa, H. (1957). *Games and Decisions: Introduction and Critical Survey*. New York: John Wiley and Sons.

Nawrat, Z. (2019). *Roboty medyczne w systemach teleinformatycznych; Informatyka w medycynie.* In: M. Kurzyński, L. Bobrowski, A. Nowakowski, R. Rumiński (eds.) Seria: Inżynieria Biomedyczna, Podstawy i Zastosowania. T.7, Akademicka Oficyna Wydawnicza EXIT, p.727-760.

Nawrat, Z. (2011). *Robot chirurgiczny Robin Heart – projekty, prototypy, badania, perspektywy.* Rozprawy habilitacyjne 24/2011. Śląski Uniwersytet Medyczny w Katowicach, Katowice.

Nawrat, Z. (2020). *MIS AI – artificial intelligence application in minimally invasive surgery, Mini-invasive Surg 4:28.* https://doi:10.20517/2574-1225.2020.08

Nawrat,Z., Religa, Z. (2009). *An unfinished story*, The International Journal of Artificial Organs, 32 (6), https://doi.org/10.1177/039139880903200601

Paliga, M. (2023). *The relationships of Human-Cobot interaction fluency with job performance and job satisfaction among Cobot Operators—the moderating role of workload.* International Journal of Environmental Research and Public Health, 20, 5111, 1-19. https://doi.org/10.3390/ijerph20065111

Pawlak, A. (2021). *Etyczne aspekty zastosowań sztucznej inteligencji w medycynie*, Medycyna i Etyka, nr 2021(4), p. 22-35.

Stanovich, K. E. (2009). *What Intelligence Tests Miss: The Psychology of Rational Thought*. Yale University Press. http://www.keithstanovich.com/Site/Research_on_Reasoning_files/Stanovich_ThePsychologist.pdf

Suchacka, M. (2020). *Corporate digital responsibility – a new dimension of the human-technology relations.* System Safety: Human – Technical Facility – Environment – Czoto 2/2020, iss.1, https://doi.org/10.2478/czoto-2020-0001

Szpunar, M. (2023). *Pomiędzy antropomorfizacją maszyn a technomorfizacją człowieka*, Journal of Modern Science, 53/3/2023, s. 24-38, http://dx.doi.org/10.13166/jms/173275

Wojewoda, M. (2015). *Etyka odpowiedzialności a technika medyczna.* Medical Robotics Reports. (4), pp. 76 – 83. http://yadda.icm.edu.pl/baztech/element/bwmeta1.element.baztech-df57d795-50d4-4398-b86e-27d6f14703c5

Zakręcki, A, Zawada, P., (2024). *Analiza interesariuszy projektu drukowanej 3D ortezy przedramienia firmy Mediprintic sp. z o.o.* Journal of Modern Science, Tom 1/55/2024. https://doi.org/10.13166/jms/183726