Artificial intelligence in healthcare: directions of standardization

Ad-Hoc Group on Application of AI Technologies in Health Informatics (AHG2 TC215 ISO)*
Corresponding Author: Professor DR. Georgy Lebedev
E:lebedev@d-health.institute

Abstract

Artificial intelligence (AI) can have a significant positive impact on health and healthcare. AI can be used to improve the quality, efficiency and equity of health care. However, AI has the potential to have significant negative impacts. Therefore, AI medical applications should be designed and deployed in accordance with established guidelines and legislation. There may be gaps or questions in the current regulatory framework related to the interpretation and application of the existing regulatory framework to healthcare applications that include artificial intelligence solutions. Global standardization maintains a consistent approach and can reduce the burden on stakeholders when it comes to establishing regulatory frameworks, interpreting and complying with regulatory requirements. While AI is far from new, it has only recently become mainstream. This chapter outlines the research of the authors who are members of the Hoc Group on Application of AI Technologies in Health Informatics (ISO AHG2 TC215), which was formed by ISO Technical Committee 215 to define goals and directions for standardization in the field of AI in health care.

Keywords: Artificial Intelligence, Deep Learning, Electronic Health Records, Standardization of Digital Health

1 Introduction

Artificial Intelligence (AI) has the potential to have a significant positive impact on health and healthcare. AI can be used to improve the quality, efficiency, and equity of healthcare delivery. Nevertheless, AI has potential for significant negative impacts. Consequently, AI health applications should be developed and deployed according to established principles, as well as to comply with jurisdictional regulations. Current regulatory frameworks may have gaps, or there may be questions related to interpreting and applying existing regulatory frameworks to health applications that incorporate AI solutions. Global standardization supports a harmonized approach and can reduce the burden on stakeholders when it comes to establishing regulatory frameworks and interpreting and fulfilling regulatory requirements. While AI is far from being new, it has only recently become ‘mainstream’. Progress in computing and transmission hardware and software has paved the way for embedding AI components in many products and services for the general public.

The following three stakeholder groups are impacted by AI in healthcare:

1.Healthcare, public health and research community at large including physicians of various clinical sub-specialties, nurses, administrators, researchers, pharmacists, laboratory staff, executives and other healthcare professionals

2. Health Information Technology (HIT) and AI technology solutions developers and

3.End users – consumers of healthcare and public health services (patients including children, family members, care givers, and the public at large).

Each group plays a critical role in ensuring productive, safe and ethical use of AI in health-related information sharing and use.
management and candidate ML approaches for combining the value of these complementary yet disparate data resources for patient-specific risk prediction modelling.

2 Definition of Artificial Intelligence (AI)

Numerous definitions of "artificial intelligence (AI)" are exist. Some of them focused on philosophical aspects, others on mathematical issues or computer science. For example, the definition proposed by Academy of Medical Royal Colleges: "The simulation of human intelligence processes by machines, especially computer systems. These processes include learning (the acquisition of information and rules for using the information), reasoning (using rules to reach approximate or definite conclusions) and self-correction." It seems to be too wide and non-specific. Really, there is no agreed definition, including field of healthcare. In general, the term "AI" broadly refers to systems and technologies that resemble processes associated with human intelligence: first reasoning, learning and adaptation, sensory understanding, also as interaction.

There is the specific definition at ISO/IEC 22989 Artificial Intelligence Concepts and Terminology, N695 draft, August 18, 2020: *artificial intelligence: <system> capability to acquire, process, create and apply knowledge, held in the form of a model, to conduct one or more given tasks*. This is an excellent one from the technical point of view. In addition, it is very affordable for standard development process. Otherwise, specific aspects of medicine and healthcare were neglected.

In 2018 Expert group of European Commission proposed a capacious definition, which perfectly reflects the features of AI in medicine and healthcare also: «Artificial intelligence (AI) systems are software (and possibly also hardware) systems designed by humans that, given a complex goal, act in the physical or digital dimension by perceiving their environment through data acquisition, interpreting the collected structured or unstructured data, reasoning on the knowledge, or processing the information, derived from this data and deciding the best action(s) to take to achieve the given goal. AI systems can either use symbolic rules or learn a numeric model, and they can also adapt their behavior by analyzing how the environment is affected by their previous actions[1]».

Can be propose a definition for medicine and healthcare, based on cited one (with respect to patient and health practitioner needs and rights, and medical practice features).

Artificial intelligence (AI) systems in healthcare are software (alone or embedded into hardware) systems:

- designed by humans ***teams led by health practitioner***,
- ***on standardized and prepared on evidence-based approach data***,
- given a ***specific clinical or management*** goal,
- act in the physical or digital dimension ***(with integration into hospital information systems, applications)***,
- by perceiving their environment through data acquisition, interpreting the collected structured or unstructured data ***from health records, medical devices, patient itself, follow-ups, own previous actions***,
- reasoning on the knowledge, or processing the information, derived from this data and deciding the best action(s) to take to achieve the given goal ***with considering patient safety, evidence-based practice, prevalence of human (doctor, nurse) decision***.

The components of the definition will be discuses in the following sections.

3 History

The study of mechanical or "formal" reasoning began with ancient philosophers and mathematicians in antiquity. In 1832 Semen Korsakov (1787-1853) invents five mechanical devices - so-called "intelligent machines" - for the partial “mechanization of mental activity in the tasks of search, comparison and classification of information”. His concept of artificial amplifying of natural intelligence echoes the modern concept proposed by The American Medical Association. They are recently defined the role of AI in healthcare as “augmented intelligence,” stating

that AI will be designed and used to enhance human intelligence rather than replace it[2]. In 1950 Alan Turing (1912-1954) proposed the imitation game later known as “The Turing test”. This is the test of a machine's ability to exhibit intelligent behaviour equivalent to human. In 1956 the term “artificial intelligence” is coined by John McCarthy (1927-2011) at a Dartmouth conference. AI is founded as an academic discipline. 1956–1980 are worldwide golden years of AI science, expert systems development, funding. Computer automated analysis of medical information (electrocardiogram, spirometry, other functional tests) was successfully implemented in the USSR and the USA in parallel. Several systems have been used in clinical practice. Around 1980 the first “AI winter”, with reduced funding and interest in AI research due to limited capacities of available computers. Up to 1987 the rise of interest due to creation of knowledge-based expert systems conception. The first clinical decision support systems are used in medicine. It was rule-based systems for diagnosis, especially in complex patient cases, choose appropriate treatments, and provide interpretations of clinical reasoning. However, rule-based systems are costly to build, they were critically limited by the comprehensiveness of prior medical knowledge. It was hard to implement in clinical practice a system that integrates deterministic and probabilistic reasoning. Thus, rule-based approach was unsuccessful. In 1987–1993 happens the second “AI winter”, as even knowledge-based expert systems show their serious limitations and prove expensive to update and maintain. 1993–2011 Returning of some optimism. New successes are marked with the help of increased computational power and AI becomes data-driven. Some AI-based software can beats human champions at chess and Jeopardy. 2012–today incredible progress of computational power, data transmission speeds, also as availability of data allow for breakthroughs in machine learning, neural networks and deep learning. Progress of AI development for medical imaging and records analyzing. Commercial developers’ hype mainly based on unpublished, untested and unverifiable results. The prevalence of mathematics over medicine in researches. Restrained attitude of health practitioners due to lack of evidence. Arising of evidence-based approach for AI in healthcare.

Thus, artificial intelligence includes a range of methods that allow computers to perform tasks typically thought to require human reasoning and skills. Worldwide algorithms based on rules and logic specified by humans has been use in healthcare since the 1970s. During last 20 years, there have been huge technological developments, including two main components:

1. Incredible increasing of hardware computing capabilities and data exchange rate.
2. Mathematical progress of artificial neural networks and machine learning methodologies.

Progress of hardware and mathematics allows computers learn from examples rather than explicit programming now.

4 AI features and development

Rule-based expert systems contains preset answer options or backgrounds on some statistical methods (logic regression, etc.). Usually, they just able to evaluate the question according to the specified criteria and choose the most appropriate answer from the list. AI is something else. The mathematical model (neural network) is train on a prepared dataset. After that, the model becomes able to interpret new data based on internal algorithms and previous experience. The model has ability to learn, accumulate and analyze own experience – this is an "artificial intelligence".

As a scientific medical discipline, AI includes several approaches and techniques:

1. Machine learning (deep learning and reinforcement learning). For example, diagnostic imaging recognition and interpretation, electronic health records analysis.
2. Machine reasoning (planning, scheduling, knowledge representation and reasoning, search, and optimization). For example, decision-making support tools integrated into hospital information systems, data extraction systems for electronic libraries.

3. Robotics (control, perception, sensors and actuators, integration of all other techniques into cyber-physical systems). For example, automatic injections, equipment supervision, robot-assisted surgery.

Machine learning (ML) - a field of computer science that uses algorithms to identify patterns in data, it represents the dominant approach in AI[3].

Progress of ML is responsible for most of the recent advancements in the field. Usually, ML refers to a system that trains a predictive model by identifying patterns of data from input, then uses such a model to make useful predictions from new, never-before-seen data. Such algorithms can automatically learn and improve from experience without being explicitly programmed, and such "learnability" represents a key feature of AI as was mentioned above[4].

The most common ML algorithms are supervised learning, unsupervised learning, reinforcement learning, and deep learning.

A very good tutorial on this topic has been published quite recently[5].

Supervised ML - a type of machine-learning task that aims at predicting the desired output (such as the presence or absence of disease, symptom) based on the input data (such as diagnostic images, health records, laboratory tests). Supervised machine-learning methods work by identifying the input–output correlation in the 'training' phase and by using the identified correlation to predict the correct output of the new cases[6]. Supervised ML based on datasets as input and some known, labelled outcomes (tagged dataset) as output. This type of ML has been widely applied to healthcare, providing data-driven clinical decision support for mapping input variables into discrete categories (for example, structured data in radiology imaging to diagnose and stage tumor) and predictive analytics within a continuous output (for example, personal risk assessment and prognosis based on unstructured data in electronic health records).

Unsupervised ML - a type of machine-learning task that aims at inferring underlying patterns in unlabeled data (untagged datasets).

Shortly, it is use to discover the structure of data and make predictions based on input alone.

Unsupervised ML allows making an algorithm, which can find sub-clusters of the original data, identify outliers in the data, or produce low-dimensional representations of the data[7]. This type of ML do not widely used in healthcare - it is more prone to errors because it may use trivial features of the data to make predictions. One of the few applications are predicting individual disease risks using genetic biomarkers or designing personalized treatments based on genomic variations.

In some sense learning without human's labelling data it is closer to "true AI", nevertheless risks of errors is to high for healthcare. There is combination of methods that call **semi-supervised learning**. Supervised and unsupervised ML applying together by joint use of a large amount of unlabeled dataset for training with only a small proportion of tagged data. This approach is more applicable for healthcare in case of insufficient of labeled data, but it needs strict and thorough blind external validation.

Reinforcement learning is a more autonomous learning algorithm that allows a model to take actions and interact with the environment using rewards and errors as the feedback to guide training[8]. It is somewhat real self-learning approach because the model learns from its own experience without either data or tagged datasets. In healthcare reinforcement learning is applicable for situations in which AI needs to continuously interact with the environment and adjust its actions based on the feedback from the environment. Therefore, it can be tasks for optimizing or treatment (medication therapy) design or robotic-assisted surgery, manipulation and diagnostics (like intravenous injection, ultrasound examinations).

Deep learning (DL) - a subfield of the larger discipline of ML, which employs artificial neural networks with many layers to identify patterns in data. It discovers the intricate structure in large datasets by using a backpropagation algorithm operating on multiple levels of abstraction. The key ability of DL it is adding of "hidden layers" of artificial neural networks, which are increase the capacity of algorithms

for solving complex real-world problems. DL is a perfect approach in cases that result rely heavily on feature detection and real big data like genomics, unstructured health records in hospital archive, drug and biomarkers discovery, speech and language recognition.

Natural language processing (NLP) - uses computational methods to automatically analyze and represent human languages. In healthcare it is applicable for:

- Doctor's speech recognition and automated documents filing,
- Patient speech and write recognition for identification, some diseases or symptoms screening (so-called symptom-checkers), navigation and information,
- Health practitioners' and patients' personal assistance,
- Equipment control,
- Quality control,
- Health record analysis for personal risk assessment and other various tasks.

There is a very large amount of unstructured textual data in healthcare (history, doctors' notes, test results, lab and radiology reports, patients' diary, medication orders, discharge instructions, etc). NLP can extract critical information for various tasks. However, much more impressive is abilities of ***combination of ML and NLP*** for medicine and healthcare. It will enable health practitioners to make timely diagnoses and treatment decisions, which can have profound impact on health service delivery, particularly on the ways that patients are treated.

This combination is open ways for ***medical robots***. Potentially, they can help with surgical operations, diagnostic and treatment manipulations, rehabilitation, social interaction, assisted living, quality control, and more. Rights now there are AI-assisted surgical robots for in neurology, orthopedic, and various laparoscopic procedures. They can analyze data from preoperative health records to physically guide a surgeon's instrument in real time during a minimally invasive procedure. There are evidences that such robot-assisted surgery allows reduce hospital stay, complications, and errors. In nearest future AI-assisted robots will be used for rehabilitation (after serious trauma, stroke or

other neurologic disease). They would assist in the care of the elderly individuals, monitor vital signs and take proper actions when needed.

The process of ***AI development*** with mentioned above methods consists from “medical” and “technical” parts.

1. Medical part:

- goal setting,
- data selection,
- data tagging,
- dataset formation.

2. Technical part:

- creation of the mathematical model,
- training the model,
- calibration the model,
- internal validation,
- external validation.

It is clear that AI-model can continue learning after official development process ending. There are various dynamics of AI changes. As a product or service on healthcare market AI-model can be:

1. Locked – a model may learn in the field, usually through analyzing of feedback at the developer site. These models does not change during practical use.

2. Change by user – the same way of learning during practical use, but health practitioner can select an appropriate working point.

3. Discrete change through learning – a model learns in the field itself. Update with explicit/distinct update by developer or health practitioner.

4. Continuous change through learning - learns in the field also, but update of a model happens without explicit manufacturer or user interaction.

5. Hybrid form - a model continues to adapt, until a human decides otherwise and returns it to a prior state.

Usually, the medical device regulations impose strict limitations with regards to the significance of changes allowed by the AI manufacturer before a new conformity assessment is required.

Algorithms that change themselves during use can only do so within predefined boundaries taken into account during the conformity assessment.

Labeled (tagged) datasets are extremely important for AI development and validation in healthcare.

Reference dataset is a structured set of biological, medical, health, social, demographic and other related data that has been pre-prepared (processed, tagged, labeled) according specific clinical task with preservation of data anonymity and patients' rights. In medicine and healthcare, reference dataset can include diagnostic images and information on pathological changes on images (annotations); structured clinical cases and related documents from EHR; libraries of keywords, phrases and their critical combinations. If the dataset contains confirmed information on the final diagnosis for each case, then it is call "verified".

According to Sergey Morozov et al, 2019 the reference dataset should meet the following requirements[9]:

- the normal-to-abnormal ratio should reflect the prevalence of the target pathology in the population;
- several medical centers should source the reference dataset to introduce the data heterogeneity;
- demographic, socio-economic characteristics and basic health indicators in the reference dataset should correspond to the population's average characteristics in the target region;
- the proposed size of the reference dataset should be justified per statistical considerations, and the desired diagnostic accuracy by the main metrics indicated above;
- reference datasets used in clinical tests for registering the software as a medical device should not be publicly available (to exclude the possibility of training AI algorithms on reference datasets).

The methodology of reference datasets preparation is a specific topic discussed elsewhere.

There are three main types of companies that are providing a AI solutions for healthcare and related areas:

1. Vendors of EHR and PACS, which add AI capabilities in their products (for example, algorithms for image analysis, NLP to support clinical decision-making, etc).
2. Big tech companies, which are providing AI cloud platforms, services, and ML algorithms for health organizations to build, manage, and deploy various AI applications with massive data.
3. Specialized healthcare AI companies and start-ups. They producing various kinds of AI healthcare applications usually focused on small specific tasks.

In general, most applications of AI are narrow, current solutions are only able to carry out specific tasks or solve pre-defined problems. However, for the healthcare this approach seems to be effective and widely adopted. In conditions of huge heterogeneity of data and colossal risks, it is easier and more reliable to develop and train an AI-based system for clear very specific clinical task.

5. Problems and challenges

Results of AI health applications depend on 1) data quality, 2) quality of algorithms (and hence of software programming generally), 3) limits of validity of their applicability, and 4) their proper implementation (e.g., in workflows) and other operational considerations. Beyond the availability of sufficient data, data quality is one of the primary constraints in implementing AI health applications (particularly machine learning, deep learning, and other applications dependent on big data). It is also the main factor determining the validity of results, captured in the well-worn phrase from the start of the computer-age, *garbage in garbage out*.

AI has the serious potential to help solve important healthcare challenges, but might be limited by 2 serious issues:

- quality and availability of standardized health data,
- inability to display some human characteristics as clinical thinking and reasoning, compassion, emotional behavior, life experience sharing, also as intuition.

A key challenge is ensuring that AI is developed and used in a way that is transparent, explainable, safe and compatible with the public (including doctor, patients, society, industry) interests. AI in healthcare promises great benefits to patients and health practitioners, otherwise it equally presents risks to patient safety, health and data security. There is an only one reasonable way to ensure that the benefits are maximize and the risks are minimize. Health practitioners have to take an active role in the development of AI-based technologies. Their medical knowledge and clinical experience are vital for their involvement for reasonable task definition, standards and methodologies creation, and limitations overcoming, also as for dataset preparation, systems validation and following the evidence-based approach. For example, according to European Society of Radiology almost 100% of doctors believe that radiologists will play a role in the development and validation of AI-based software. Majority thinks that they should supervise all development stages of an AI system (>64%) or helping in task definition (>53%). Otherwise, one third only focused on just providing of labelled images (>29%) and directly developing of AI-based applications (>27%). More than 20% of radiologists are already involved in AI systems development and testing[10].

Thus, health practitioners can and have to be part of the development and use of AI. It will require rethinking and changes in behavior and attitude to education and careers. In nearest future, principles and basic methodologies of data science would be a part of doctor's competences as auscultation or injection. Anyway, AI-based software must be develop in a regulated way in partnership between clinicians and computer scientists.

There is no differences on patients' rights and safety depending on presence or absence of AI in clinical activity. Safety have to be remain paramount. It is important to understand safety as a need of a patient and health practitioner both. Patient are carry about health and life. Doctor or nurse are carry about good practice and trust to technology.

There is a critical aspect is the conformity with the national or regional data protection environment. On a global level, availability of data must be balanced by individuals right to personal data protection.

AI can combine data coming from different information sources, each one containing anonymized data but when that data is combined using AI, identification of individuals is possible. An ethical principle should be: *Do not use AI in the Healthcare sector for unwanted or unintended re-identification of individuals by combining anonymized data from different sources*.

Explainability. Regulations across the world requires medical treatments be explainable and understood by patients so that patients can give their informed consent to a planned medical intervention. However, AI systems may utilize advanced statistical and computational methods to determine a course of treatment that may not be easily understand by health practitioners, let alone patients.

Modern machine learning algorithms are usually describe as a "black box", because it is too difficult for a human to understand how the conclusion based on the huge number of connections between artificial "neurons" was reached. Meanwhile, doctors can trust to "black box" because unclear decision-making creates serious risks to a patient. Doctors and nurses need to be able to inferentially authorize their decisions, recommendations, diagnoses and predictions and take responsibility for them.

There are two critical barriers for AI in healthcare, how mentioned by Bert Heinrichs et al, 2019: "The first issue is that epistemic opacity is at odds with a common desire of understanding and potentially undermines information rights. The second (related) issue concerns the assignment of responsibility in cases of failure. Subsequently, we elaborate these issues in detail.[11] ". Conception of explainable AI potentially can help to overcome these problems.

First step here is making AI interpretable. Interpretability it is understanding a links between a cause and effect within an AI model. Observer can to *predict* what is going to happen, given a change in input or algorithmic parameters. Explainability it is a next step. This is understanding and explanation in human terms internal mechanics of a machine or deep learning system. Explainable AI means implementation of transparency and traceability of statistical "black-box" machine and deep learning methods. Explainability of AI became a mandatory

requirement in healthcare. Explainability should be described from patient and health practitioner point of view.

Patient view. Explainability is best understand as effective contestability. Explanation of this conception are exist in patient-centric approach to AI usage in medicine (particularly in diagnostics) had propose by Thomas Ploug and Søren Holm, 2020[12]. According to the approach, patients should be able to contest the diagnoses of AI diagnostic systems. It is necessary to ensure the availability of four types of information for this purpose:

1. How to the AI system's use of data.
2. The system's potential biases.
3. The system performance.
4. The division of labor between the system and healthcare professionals.

First, individuals have a right to privacy, to protect themselves against harm and risks. Exercising this right to contest the use of personal health and other data are backgrounded on two types of information. So, AI-based medical services requires that individuals have access to information about:

- types of personal data used in AI diagnostics (e.g. clinical tests, images, biopsy etc.),
- sources of such data (e.g. Electronic Patient Record etc.), because sensitivity and quality of data may be critically dependent on the source).

Second, individuals have a right to protect themselves against discrimination, including due to AI’s bias. Exercising this right requires that individuals have access to information about:

- characteristics of the dataset on which the model is built and validated,
- how the data for dataset were selected and categorised by humans,
- characteristics and level of testing the AI model.

The good practice if developer make an initial general claim of potentially relevant bias. Nevertheless, an individual have a right to have individual bias investigation.

Third, the right to contest the AI model performance are directly links with the right to protect themselves against harm. Here individuals must have access to information about:

- performance of the AI model,
- trials and tests used to evaluate the performance,
- information about the key indicators of the diagnosis,
- alternatives to the suggested diagnosis,
- changes that will lead to a reconsideration of the diagnosis.

Fourth, the right to contest the division and organization of labor is also protect individuals against harm and makes responsibility clear. For exercising of this right, individuals must have access to information about:

- role of AI in the clinical work-flows,
- role of health practitioners in the clinical work-flows,
- legal responsibility for medical (diagnostic) procedures.

AI developer and health practitioner both have a duty to prepare and provide the information needed for effective contestation. In real life, most patients are unlikely to contest the AI advice. They would be satisfy by explanation of doctor about diagnosis and further tactics. Anyway, relevant information should exist and be updated regularly. One more duty exist whenever patient wish, patient have to be inform that AI system has been provide an advice and it has been used by the health practitioner.

Health practitioner view. *AI interpretability* refers to a health practitioner's ability to understand the AI-model itself or at least a summary. In clinical context, interpretability of AI provides knowledge for shared decision-making. Such AI allows humans to gain knowledge about the considered features, their integration and weighting. This information are relevant for connecting it to data from health records, laboratory tests, imaging examinations, etc. Health practitioner can process and interpret the results of the AI model relative to information

from various other sources and makes individual evaluation of the clinical case. Moreover, the final clinical decision can be clearly explain to the patient on evidence-based manner, because interpretability allows estimate the probability of a result generated by AI.

In healthcare *explainable* AI is need for many purposes including clinical practice and decision-making, professional education, research. Medical professionals must be able to understand and to retrace the machine decision process. AI explainability have to be realize on two levels[13]:

1. Model level – as a human's ability to understand the structure of the process, which provides a bridge to shared clinical decision-making.
2. Results level – as understanding why was this particular decision made in this specific clinical case.

Explainable results of AI works can be directly integrate into clinical decisions and recommendations, also as communicated to the patient.

According to modern level of technologies, developers have to make their algorithms ***interpretable*** *and* implement ***elements of explainable*** AI.

Transferability. AI can be well optimized for the specific task, but it will be incorrect imprecise and ineffective on data it has not seen before. This is a very typical situation, which occurs due to:

- AI training on limited dataset or data from only one hospital,
- Lack of independent (external) blinded evaluation on real-world data.

A number of developers uses cross-validation (leave-one-out) method. They use one dataset, which divided on 2 parts: training dataset and testing dataset. Creation and learning of the AI-model is carried out at the first one. Then, they make an internal validation of AI accuracy at testing dataset. This is not enough, since technically both “subsets” consist of data from the same source (hospital or even a number of hospitals). The problem is the medical and health data are not standardized, and always have some peculiar properties due to differs in clinical traditions and rules, population, medical devices customization,

protocols, etc. Thus, external validation on new ("previously not seen") datasets is obligatory. Moreover, it should include two stages:

- Retrospective external validation at new reference datasets,
- Prospective external validation on real-world data.

This approach allows to overcome a problem when narrow applications that cannot generalize to clinical use.

Ethics. The use of AI in clinical medicine and health researches raises many ethical issues[14]:

- potential to make erroneous decisions;
- responsibility when AI is used to support decision-making;
- difficulties in validating the outputs of AI systems;
- inherent biases in the data used to train AI systems;
- ensuring the protection of potentially sensitive data;
- securing public and professional auditorium trust in the development and use of AI technologies;
- effects on people's sense of dignity and social isolation in care situations;
- effects on the roles and skill-requirements of healthcare professionals;
- patient's preparedness and reactions to medical services (for example, image reporting) made by an AI-application alone without supervision and approval by a physician.
- potential for AI to be used for malicious purposes.

Currently, more than 84 ethics initiatives have published reports describing high-level principles, tenants, and abstract requirements for the development and deployment of AI. An analysis of 36 prominent sets of such principles revealed the following 8 themes:

- privacy,
- accountability,
- safety and security,
- transparency and explainability,
- fairness and non-discrimination,
- human control of technology,
- professional responsibility, and

- promotion of human values.

These principles should be applied to every stage of the AI health application life-cycle, from needs determination and design through decommissioning. When pertinent, to promote trustworthiness, they can be applied in basic requirements for the development, deployment, use, and evaluation of AI health applications. Moreover, relevant principles and requirements can also be applied to the development and revision of health informatics standards pertaining to or encompassing AI health applications. In particular, whenever applicable there should be full transparency about an AI product throughout its life-cycle, e.g., how the algorithm works, what data were used to train it, what tests were conducted, how the trained product performed in such tests, experience with use in practice, etc.

Unsafe AI could harm patients across the national healthcare system. Really, medical AI will help some patients but expose others to unforeseen risks. It seems like a doctor have an automatic right to over-rule an AI decision. One must always remember AI-tools could be confidently wrong, moreover a misleading algorithms hard to identify.

AI will change or at least influent the doctor-patient relationship. The health practitioner will need to behave differently to learn:

- interact with expert patients, who may have selfdiagnosed with AI tools,
- preserve of human contact to reduce patients' loneliness, safeguarding, social needs due to introduction of AI into healthcare.

Responsibility is still an unclear. A rhetorical question: who will be responsible for harm caused by AI mistakes: developer, IT-company, the regulator or the health practitioner? National regulators and international authorities (World Health organization (WHO) and International Telecommunication Union (ITU)) should solve this question, as soon as possible.

Currently, radiology is most involved in AI sphere of medicine. So the opinion of radiologists can be interpolate into other specialties. According to European Society of Radiology survey 41% of respondents believes only doctor will take responsibility for AI outcome, but other

41% choose a scenario of shared responsibilities between doctor, patient and AI developer. In exclusive responsibility of developers or insurance companies believes 10% and 3.6% of doctors accordingly. Note, more than half (55.4%) doctors believe that patients will not accept a report made by an AI alone without supervision and approval by a health practitioner. Only 11.7% claim the opposite, one third (32.9%) are still doubt about[15].

6 AI systems in healthcare

Currently, AI is used or trialled for a range of healthcare purposes, including detection, staging and monitoring of disease (screening or in clinical environments), quality control, prognosis, quantitative measurement of biomarkers, support of structured reporting, management of chronic conditions, delivery of health services, productivity increasing, support of clinical decision, drug discovery.

According to Mei Chen et al, 2020[16] as a part of hospital or healthcare digital environment, AI can accomplish the following:

- Unlock the power of big data and gain insight into patients;
- Support evidence-based decision-making, improving quality, safety, and efficiency, coordinate care and foster communication;
- Improve patient experience and outcomes;
- Deliver value and reduce costs;
- Optimize health system performance.

AI should standardize assessment and treatment according to up-to-date clinical guidelines and protocols, raising minimum standards and reducing unwarranted variation. Reasonable, well-trained and well-validated AI improves access to health services, providing advice in realtime to health practitioners and patients, also as identifying critical and dangerous situations (medical emergencies).

There are number main AI applications in healthcare:

1. Clinical practice:

- Screening and detection of diseases at an early stage,

- Prognosis, risk stratification, prevention,
- Decision-making support on diagnosis,
- Decision-making support on treatment and clinical tactics,
- Management of medications (pharmacotherapy),
- Assistance during surgery/invasive manipulation, automated surgery
- Patient monitoring, pro-active screening via wearables and sensors embedded into smart environment,
- Processing of medical images and test results, health records for various clinical tasks,
- Automated filling of medical documentation (sample generation, speech recognition),
- Personal support in self-control, self-diagnosis, healthy living.

2. Healthcare management:

- Healthcare system modeling,
- Epidemiology control,
- Predictive analytics,
- Quality control in healthcare,
- Medical education.

3. Researches:

- Automated experiments,
- Automated data collection,
- Patient selection for clinical trials,
- Genome discovery,
- Biomarker discovery,
- Drug discovery and repurposing,
- Literature mining,
- “Omics” discovery.

Technically in healthcare AI performs detection, classification, segmentation, processing (including natural language processing), comparison, prediction (prognosis), and recommendation generation with three types of data:

1. Imaging (still and moving, including radiology, endoscopy, dermoscopy, patient view, etc).
2. Documents and speech (including various health records, patient or doctor speech, etc).
3. Data stream (including statistics, epidemiology data, raw diagnostic data, etc).

Thus, for proper, safe and effective using AI have to be embed into the clinical workflows to solve specific tasks at the point of care. Electronic Health Records (EHR) are the backbone of modern digital healthcare systems. Therefore, the preferable approach to integrate AI directly into EHR systems. In some countries, this approach already recommended in clinical protocols and guidelines, at least for radiology. In term on AI integration into EHR a number of abilities appears. Based on Mei Chen et al, 2020[17] they can be systematize as follows:

1. Providing clinical decision support at the point of care to improve diagnostic accuracy and treatment recommendations:

- Diagnostic analytics using medical imaging or genomic, clinical, laboratory, behavioural and other data;
- Predictive analytics and personal risk assessment (high-risk patients selection, outcomes prognosis);
- Personalized treatment recommendations based on evidence-based practice (this is an unique ability of AI to joint “narrow” individual data and “wide” clinical recommendations);
- Prediction and prevention of adverse events;
- Medication safety and reconciliation;
- Routine integration of various data for triage and critical care monitoring, diagnostic interpretation, and treatment modification.

2. Providing patient engagement technology to support self-care

- Patient empowerment via access to personal health data, prognostic information;
- Patient engagement tools (chatbots, wearables, mobile devices) for supporting patient and family members education, informed

decision-making, selfmonitoring, and self-management of chronic conditions;

- Pre-hospital support in emergency situation (including prediction of acute situation, patient support, hospital information);
- Proactive screening via smart environment;
- Channels for patients to interact with healthcare providers and on-line services;
- Crucial patient data extraction from wearable, mobile devices, sensors in smart environment, health apps, integration of these data into EHR.

3. Optimizing workflows and resource allocation, improving operational efficiency

- Predictions of the number of patients during a specific period and resources needed (for example, in situation of epidemic);
- Integrated voice technologies in EHR for clinical documentation;
- Integrated NLP for processing narrative health data and providing critical summaries of key patient information, also as for quality control;
- Simplification of operational processes through AI automation.

4. Facilitating population health monitoring and management, improving wellness via data from smart environment, social media, various information systems (with preserve of human rights, personal data and confidentiality protection):

- Population health monitoring,
- Epidemic prediction;
- Predictive analytics for health service;
- Identification of high-risk population groups;
- Prioritization of at-risk patient populations and management of proactive interventions;
- Investigation of social determinants on healthcare and management of population wellness.

5. Supporting real-world clinical research and evidence-based medicine:

- Collection and storage of real-world data for clinical research and care improvement;
- Precision medicine and clinical trial matching;
- Drug discovery;
- Patient selection for clinical trials;
- Biomarkers discovery based on various data sources.

The possibilities for using AI in healthcare are very wide as was mentioned above. There is the recognized framework for any kind of AI case in the field (fig.1).

Figure 1. Framework of AI Use Cases in Healthcare

The framework defined four broad categories, with subcategories defined for the “Individual Health” category:

1. Population Health
2. Individual Health
   a. Care Routing
   b. Care Services
   c. Prevention
   d. Diagnosis
   e. Acute Treatment
   f. Follow-up and Chronic Treatment
3. Health Systems
4. Pharma & Medtech

Almost any AI use case can be categorize within this framework. Furthermore, specific attributes associated with each category will inform whether (or not) an AI use case would be subject, for example, to regulatory approval.

The AI can enhance, extend, and expand human capabilities in medicine, delivering the types of care patients need, at the time and place they need them. In healthcare AI cannot be used alone. A human-machine partnership is a key for improving of clinical effectiveness (quality, safety, and efficiency), access, and affordability of care.

Nonetheless, complete automation is possible for some specific situations:

- tasks where AI has surpassed human performance like (screening, health records peer-review for quality control, library search for data extraction, etc),
- tasks where mistakes do not lead to serious consequences (primary prevention, flagging an at-risk population group for vaccination),
- situations where human health practitioner are unavailable but AI can help with information and support (for example, chatbot for patient support and navigation during insulin self-injection).

In clinical practice the key in “doctor-AI” partnership is to keep the delicate balance between the types of care human value and the levels of automation that technologies offer.

In educational context AI should be incorporate into simulations generating clinical scenarios across a range of specialities to enhance training and learning. Advancement of medical knowledge produce the sheer volume of new information exceeds human abilities to keep pace in real time. AI can analyze large datasets and libraries across multiple sites to condense information for the health practitioner for clinical decision-making and lifelong learning. Moreover, AI, combined with other digital technologies, can personalize education by evaluating previous experiences, responses and outcomes to model the strengths and weaknesses of individual clinicians.

7 Quality and safety of AI

Quality principles include health care that is: safe, effective, patient-centered, timely, efficient, and equitable. Healthcare system goals include:

1) enhancing patient experience,
2) improving population health,
3) reducing per capita health care costs,
4) safe-guarding/improving the work-life of health care providers,
5) improving business processes,
6) equity and inclusion.

Currently, more than 84 ethics initiatives have published reports describing high-level principles, tenants, and abstract requirements for the development and deployment of AI[18]. An analysis of 36 prominent sets of such principles revealed the following 8 themes:

1) privacy,
2) accountability,
3) safety and security,
4) transparency and explainability,
5) fairness and non-discrimination,
6) human control of technology,
7) professional responsibility, and
8) promotion of human values .

These principles apply to the development and deployment of AI health applications. They should be applied to every stage of the AI health application life-cycle, from needs determination and design through decommissioning . When pertinent, to promote trustworthiness, they can be applied in basic requirements for the development, deployment, use, and evaluation of AI health applications. Moreover, relevant principles and requirements can also be applied to the development and revision of health informatics standards pertaining to or encompassing AI health applications. In particular, whenever applicable there should be full transparency about an AI product throughout its life-cycle, e.g., how the algorithm works, what data were used to train it, what tests were conducted, how the trained product performed in such tests, experience with use in practice, etc.

Nuffield Council on Bioethics declared: “The tech mantra of “move fast and break things” does not fit well when applied to patient care.”[19] Despite hype, artificial intelligence in healthcare is still in its infancy, and it really has hardly started. There are some positive prognosis about AI: it will deliver major improvements in healthcare quality and safety, reduce costs, and even make an imminent revolution in clinical practice. Nevertheless, there are no strong evidence for that. Medical society are very early in the evidence cycle and it is unclear how the predictions coincide with reality.

It may be difficult to apply current regulatory frameworks for health and medical technologies to applications utilizing artificial intelligence. For instance, medical device regulations across the world require validation that devices and processes produce reproducible and expected outputs or results. Many jurisdictions are modifying existing regulations and/or developing new regulatory frameworks to govern the use of AI in Healthcare. These for instance are related to the jurisdictions’ applicable regulatory frameworks for Software as a Medical Device (SaMD). Relevant initiatives are underway in the USA, the EU, Russian Federation Australia, Canada and elsewhere. Most of these jurisdictional approaches leverage a common risk categorization framework developed under the auspices of the International Medical Device Regulators Forum’s (IMDRF) SaMD working group.

AI applications can be classify by applying a risk-based approach. For example, IMDRF categorize SaMD along two dimensions and classify SaMD into one of four categories, ranging from low-risk to high-risk, taking account of:

- State of healthcare situation or condition,
- Significance of information provided by SaMD to healthcare decision.

Similarly, the EU Commission’s White Paper on Artificial Intelligence considers that in a healthcare a range of AI applications can exist, and proposes an assessment of the level of risk of a given use based on the impact on the affected parties. For AI applications in the healthcare sector, additional factors might be taken into account for establishing a risk-classification framework, for example: the degree of

adaptivity (values could be “locked”, “discrete adaptive”, “continuously adaptive”), the degree of autonomy. AI applications that have a high degree of adaptivity and/or high degree of autonomy can be regarded as potentially of higher risk. It is not immediately clear how to regulate such applications under existing legislations, how to place such systems on the market, and how to operate such systems, at least in a safe and effective manner and with certainty about potential liability.

Currently, legislators and regulatory agencies across the world are involved in active work to creation and harmonization of rules for AI in healthcare. Interpretation and guidance may be helpful for applying existing regulations to AI technology, and modifications, new regulations and new standards may be necessary to fill in possible gaps and to resolve ambiguities in existing regulations. Wherever possible, regulations and standards should be harmonized internationally to ensure that everyone has access to current state of the art technologies and to prevent the development of technical barriers to trade that raise costs and limit access to healthcare.

Based on experience with Management System Standards (MSS) such as ISO 9001:2015, certification of manufacturers of AI health applications can be expected:

1) to improve product quality,

2) to assist customers to select vendors.

Such certification may also be of value to regulators of medical products when using a risk-based approach. The term manufacturer encompasses both developers of commercial products (vendors) and also healthcare and similar organizations that develop products for their own use.

According to globally recognized practice a manufacturer can only place medical AI-based devices on the market for use on patients or their data when these are safe and effective. Once the device is on the market, the manufacturer must perform clinical evaluations throughout the entire lifetime of the device, including post-market clinical follow-up, to prove the assumptions remain valid and no risks emerge that are unacceptable. For that purpose a clinical trial (test) should be perform. The objective of clinical trial (test) is to confirm the effectiveness, safety of use, and

compliance of medical device characteristics with the intended use specified by the manufacturer. Usually, the clinical trial consist from two stages[20]:

1. Analytical validation.
2. Clinical acceptance.

Analytical validation refers to the evaluation of the correctness of input data processing by the software to create reliable output data, which is performed using reference datasets. Clinical acceptance (evaluation of the performance by using the software within a standard operating process) consists of two components:

- clinical correlation (evaluation of whether there is a reliable clinical relationship between the results and the target clinical condition),
- clinical validation (confirmation of achievement of the intended goal for the target population in the clinical workflow through the use of accurate and reliable output data).

Clinical tests are organize per national legislation and local or accepted international methodology for assessing the quality, effectiveness, and safety of medical devices.

Various metrics can be used to assess AI. The standard set of diagnostic metrics includes:

1. Sensitivity, specificity, accuracy, the likelihood ratio of a positive or negative result, positive and negative predictive value.
2. Area under receiver operating characteristic curve (ROC) as area bounded by ROC-curve and horizontal coordinate.
3. The agreement (concordance) of classification.
4. Similarity degree.
5. Timing study.
6. Retrospective per-review (audit).

Detailed information, definitions and formulas can be find elsewhere.

Standard metrics are used to compare the diagnostic performance of index-test (AI-based software) relative to the reference-test (another option for diagnostic, screening, decision-making, etc).

Thus, results of AI healthcare applications depend on:

- data quality;
- algorithms quality (and hence of software programming generally);
- limits of validity of their applicability;
- their proper implementation (e.g., in workflows) and other operational considerations.

Beyond the availability of sufficient data, data quality is one of the primary constraints in implementing AI health applications. It is also the main factor determining the validity of results, captured in the well-worn phrase from the start of the computer-age, *garbage in garbage out*.

Assessing product/algorithm performance is different from assessing (and assuring) data quality. Certain aspects of algorithm quality can be assessed by using fit-for-purpose (FFP) standard data sets 1) to ensure that an algorithm performs reliably and 2) to compare the performance of different algorithms with the same purpose. Using multiple standard FFP data sets may provide insight into the validity of outputs with different data inputs. Further, using standard degraded data sets may help to gauge use-risk, i.e., to assess potential results produced by an AI health application that was trained using FFP data when it is used with the type of real-world data (RWD) expected to be encountered in practice. The validity of results in practice depends on the quality of RWD inputs, as well as the quality of algorithms or other machine-performed processes.

8 Standardization of AI in healthcare

AI has the potential to have a significant positive impact on healthcare and improve the quality, efficiency, and equity of healthcare delivery. However, like any emerging field, there is a lack of regulatory guidance and standards regarding the use of AI in healthcare. Current regulatory frameworks may have gaps, or there may be questions related to interpreting and applying existing regulatory frameworks to health applications that incorporate AI solutions. Global standardization supports a harmonized approach and can reduce the burden on

stakeholders when it comes to establishing regulatory frameworks and interpreting and fulfilling regulatory requirements.

Standardization work should be focuse in the following areas:

- Methods to measure and to reduce bias
- Methods to measure reliability
- Notions of reproducibility in non-deterministic systems
- Methods for explainability for various kinds of AI techniques (for example, for deep-learning neural networks).

Many different organizations are developing or have developed papers that are potentially relevant to the update of AI standards (for example, ISO/IEC JTC1/SC42 is developing ISO/IEC 22989 which is an AI glossary. JTC 1/SC SC42's made ISO/IEC 23894 is titled "Risk Management"). The World Health Organization is working on developing a standardized assessment framework for the evaluation of AI-based methods for health, diagnosis, triage or treatment decisions. Some other organizations (like number of technical committees of ISO) are developing horizontal AI standards, many governments have published papers regarding the development and use of AI in multiple industries, regulatory agencies have published draft (or final) guidance documents specific to AI in healthcare, etc. There are many opportunities to leverage their existing work, as well as help in the development of future work products of those organizations.

There are three not exclusive categories of AI standards in healthcare:

1. AI Technologies and Applications.
2. AI in a Clinical Encounter.
3. AI in Clinical, Public Health and Research Sub-specialties.

Current standard landscape includes topics, already realized by number of organization:

- adaptive regulatory frameworks;
- definitions, vocabulary and general characteristics;
- recommended practice and basic principles of quality management;
- trustworthiness principles;

- ethical concerns;
- data privacy and safety;
- set of standards for human augmentation;
- set of standards for biotechnology;
- set of standards for AI in imaging.

Critically important for further progress standards for:

- Clinical trials, evaluating the performance and validity of AI health applications, such as both static (fixed until updated by the manufacturer) and dynamic (self-learning) algorithms used in AI/ML products;
- Dataset preparation, data labeling (tagging), including issues of describing, assessing, and communicating, data quality (to assist manufacturers and users of AI health applications to decide if available data are fit-for-purpose and/or how they differ from data that are).
- Quality management system for organizations that manufacture AI health applications, such intermediate products as standard data sets, and/or supply data for these purposes;
- Methods to measure and to reduce bias, to measure reliability and performance;
- Notions of reproducibility in non-deterministic systems;
- Methods for explainability for various categories of "AI solutions" in healthcare.

Guidance and regulation (via national and international standards) are need for manufacturers and users of AI in different sectors to increase the usability and confidence in such systems. Manufacturers would benefit from the guidance produced by establishing standards for good manufacturing processes and practices; which may vary by type of AI health application. Required standards for manufacturing such as AI health applications for machine learning, deep learning, and other data dependent AI health applications should encompass inputs (e.g., data), processes (e.g., algorithms, interpretation, display/distribution), and results (including their implementation, limits on use, evaluation, etc.); as well as applicable management modules (which settings may

determine AI application processes and/or performance) and environmental probes (which may set operational parameters). Resultant requirements should be expressed in the form of standardized checklists. While general principles may apply (so that some checklist items may be common to all such products), some checklist items may be specific to a type of AI health application. Checklist standards should specify personnel qualifications, experience, and/or training necessary to be able to use such checklists effectively (and, if applicable, associated certification requirements). Checklists could be the basis for developing a computerized decision support tool (CDST) to facilitate their use in practice and to document which requirements were considered when and with what results.

Customers of AI health application manufacturers, and such individual end-users as clinicians, need information pertaining to the correct use of an AI product and the interpretation of resultant information. The scope of such required correct-use information to be provided by AI health application manufacturers should be standardized, in terms of contents and expression. **:** Customers of AI health application manufacturers, such as health care organizations, can use such information to inform purchasing decisions, product installation, training, individual end-user guidance, etc. A corollary is that an organizational user has assessed the FFP of its available data. Further, organizational users should periodically repeat assessments of RWD quality so that they can gauge use-risks and can track the effectiveness of data quality improvement efforts. Regulators could require organizational health AI application users to submit results of their QMS assessments and/or product-produced process and outcome data to enable regulators to monitor product safety and performance across organizations and settings. Such requirements might extend to individual end-users so that organizational users of the AI health application can aggregate their experiences for reporting purposes. The information resulting from such reported data may support regulators in meeting their obligations to ensure that products on the market are safe and effective.

Various factors contribute to people having trust in AI systems, these factors are grouped into aspects of efficacy, adoptability, and understandability, respectively.

- Knowing that the AI system has been developed according to the state of the art, and is operated by skilled/well-trained persons
- The system offers insights into its decision-making, by providing some form of transparency and by offering explanations that are understandable to the target audience
- The system operates reliably according to some measure of reliability
- The system is proven to make unbiased decisions, according to some measure of bias
- The system is verified and validated according to standardized, recognized software development methods that are well-suited to the system at hand, taking into account of, for example, its degree of adaptability.

Many of these factors lend themselves well to standardization and are therefore seen as opportunities for standardization. This list is not exhaustive and just gives some examples:

- Standardized definitions (terms, concepts), once established, contribute to common understanding of various stakeholders (among them: legislators, operators, manufacturers) Regulations can refer to and use these definitions, either directly in legal text or by making use or harmonized/recognized standards
- Verification and validation of software can be covered in SW Lifecycle Standards
- Methods for Explainability and Transparency can be described in standards, based on the state of science, making such methods the state of the art.

To start addressing the need of regulatory guidance and standards, ISO/TC 215/AHG 2 was created in 2019 at the ISO/TC 215 meeting in Daegu, S. Korea. A cross-functional team formed and divided the work

into categories such as a landscape analysis, establishment of key principles, regulatory assessments, etc. Note that ISO/TC 215/AHG 2 does not create specific recommendations for specific updates to a specific standard, rather, it provides a series of resources (e.g. AI standards landscape, use case inventory, etc.) that can be used by teams performing an assessment of how AI might impact existing standards or require new standards.

There is the roadmap to future directions in developing standards for AI health applications, a fast evolving field. The ISO/TC 215 leadership should decide the direction of travel and roads to be taken. Key recommendations include the following:

- Establish a mechanism to keep the landscape map up-to-date, fit for purpose, and accessible - to avoid duplication, overlaps, and conflicts in standards.
- Establish a mechanism to develop/maintain a dictionary of key terms, synonyms, abbreviations, etc. to be used in standards development - to standardize and to avoid confusion in terminology used in standards.
- Issue guidance to TC conveners to review existing standards to establish priorities for revision to include needed but missing provisions pertaining to AI health application or missing additional standards - to ensure that ISO/TC 215 and its standards remain relevant.
- Develop/maintain a checklist of AI health application considerations for use when revising/developing standards - to ensure that all relevant considerations are addressed.
- Develop/maintain standards for manufacturing, evaluating, and using AI health applications, including a management system standard for certifying organizations involved in the AI health application life-cycle/supply chain - to foster good practices, safe and effective products and to support regulators.

5. Conclusion

AI is playing an increasingly important role in the provision of medical care, in supporting medical decision-making, and in managing patient flows. In many countries of the world, sore attention is paid to the development and application of AI in medicine. In these countries, government programs are being developed and innovative solutions are being introduced. Therefore, the formation of unified approaches, definitions, requirements for AI in medicine will significantly increase the efficiency of its development and application. The tasks solved by the ISO AHG2 TC215 are essential for the development of this direction of AI and will be extremely useful to the global community.

Acknowledgements

The authors would like to thank the leadership of Technical Committee ISO/TC 215 " Health Informatics" and Subcommittee ISO/IEC JTC 1/SC 42 "Artificial Intellegence" of ISO Technical Committee ISO/IEC JTC 1 "Information Technology" for providing the opportunity to bring together such a wonderful international team and do useful and meaningful work.

*Ad-Hoc Group on Application of AI Technologies in Health Informatics (AHG2 TC215 ISO)

Paolo Alcini[1], Pat Baird[2], Peter Williams[3], SB Bhattacharyya[6], Todd Cooper[7], Rich de la Cruz[8], Chandan Kumar[9], Gora Datta[10], Dorotea Alessandra de Marco[11], Andreas Franken[12], Regina Geierhofer[13], Peter G. Goldschmidt[14], Ilkka Juuso[15,16], Herman Klimenko[4], Antonio Kung[17], Frederic Laroche[18], Joe Lewelling[19], Martin Meyer[20], Sergey Morozov[21], Anna Orlova[22], Telonis Panagiotis[23], Thomas Penzel[24], Derek Ritz[25], Gaur Sunder[26], Soo-Yong Shin[27], Alpo Värri[28], Anton Vladzimerskiy[4,21], Georgy Lebedev[4,5]

Affiliations

[1]European Medicines Agency, [2]Royal Philips, Pleasant Prairie, USA, [3]Oracle Corporation, [4]I.M. Sechenov First Moscow State Medical University (Sechenov University), [5]Federal Research Institute for Health

Organization and Informatics, [6]Bhattacharyyas Clinical Records Research & Informatics, [7]Trusted Solutions Foundry, Inc., San Diego, USA, [8]Silver Lake Group, Inc., Minnetonka, USA, [9]Indian Standards (BIS), New Delhi, India, [10]CAL2CAL Corporation, Irvine, USA, [11]Italian Data Protection Authority, [12]Bundesverband der Arzneimittel-Hersteller e.V. (BAH), Bonn, Germany, [13]Medizinische Universität Graz, Austria, [14]World Development Group, Inc., [15]Cerenion Ltd., [16]Center for Machine Vision and Signal Analysis at the University of Oulu, [17]Trialog, Paris, France, [18]LCI Consulting Inc, Montreal, Canada, [19]Association for the Advancement of Medical Instrumentation, Arlington, USA, [20]Siemens Healthineers, [21]Research and Practical Center of Medical Radiology, Department of Health Care of Moscow, [22]School of Medicine, Tufts University, Boston, USA, [23]European Medicines Agency, Amsterdam, The Netherlands, [24]Charite University Hospital, Interdisciplinary Center of Sleep Medicine, Berlin, Germany, [25]ecGroup Inc, [26]Centre for Development of Advanced Computing (C-DAC), India, [27]SungKyunKwan University, Seoul, Republic of Korea, [28]Tampere University, Faculty of Medicine and Health Technology, Tampere, Finland

References

[1] A Definition of AI: Main Capabilities And Disciplines. European Commission High-Level Expert Group on Artificial Intelligence.08.4.2018.- https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=56341.

[2] Chen M, Decary M. Artificial intelligence in healthcare: An essential guide for health leaders. Healthc Manage Forum. 2020 Jan;33(1):10-18. doi: 10.1177/0840470419873123.

[3] Yu KH, Beam AL, Kohane IS. Artificial intelligence in healthcare. Nat Biomed Eng. 2018 Oct;2(10):719-731. doi: 10.1038/s41551-018-0305-z.

[4] Chen M, Decary M. Artificial intelligence in healthcare: An essential guide for health leaders. Healthc Manage Forum. 2020 Jan;33(1):10-18. doi: 10.1177/0840470419873123.

[5] Jussi Tohka, Mark van Gils. Evaluation of machine learning algorithms for health and wellness applications: A tutorial. // Computers in Biology and Medicine, Volume 132, May 2021, 104324, DOI: 10.1016/j.compbiomed.2021.104324

[6] Yu KH, Beam AL, Kohane IS. Artificial intelligence in healthcare. Nat Biomed Eng. 2018 Oct;2(10):719-731. doi: 10.1038/s41551-018-0305-z.

[7] Yu KH, Beam AL, Kohane IS. Artificial intelligence in healthcare. Nat Biomed Eng. 2018 Oct;2(10):719-731. doi: 10.1038/s41551-018-0305-z.

[8] Chen M, Decary M. Artificial intelligence in healthcare: An essential guide for health leaders. Healthc Manage Forum. 2020 Jan;33(1):10-18. doi: 10.1177/0840470419873123.

[9] Morozov SP, Vladzymyrskyy AV, Klyashtornyy VG, Andreychenko AE et al. Clinical acceptance of software based on artificial intelligence technologies (radiology). In: Best practices in medical imaging. 2019; 57: 45.

[10] European Society of Radiology (ESR). Impact of artificial intelligence on radiology: a EuroAIM survey among members of the European Society of Radiology. *Insights Imaging*. 2019;10(1):105. doi:10.1186/s13244-019-0798-3.

[11] Heinrichs B, Eickhoff SB. Your evidence? Machine learning algorithms for medical diagnosis and prediction. Hum Brain Mapp. 2020 Apr 15;41(6):1435-1444. doi: 10.1002/hbm.24886.

[12] Ploug T, Holm S. The four dimensions of contestable AI diagnostics - A patient-centric approach to explainable AI. Artif Intell Med. 2020 Jul;107:101901. doi: 10.1016/j.artmed.2020.101901.

[13] Heinrichs B, Eickhoff SB. Your evidence? Machine learning algorithms for medical diagnosis and prediction. Hum Brain Mapp. 2020 Apr 15;41(6):1435-1444. doi: 10.1002/hbm.24886.

[14] Bioethics briefing note: Artificial intelligence (AI) in healthcare and research. Nuffield Council on Bioethics.2018. URL: https://www.nuffieldbioethics.org/assets/pdfs/Artificial-Intelligence-AI-in-healthcare-and-research.pdf.

[15] European Society of Radiology (ESR). Impact of artificial intelligence on radiology: a EuroAIM survey among members of the European

Society of Radiology. *Insights Imaging*. 2019;10(1):105. doi:10.1186/s13244-019-0798-3.

[16] Chen M, Decary M. Artificial intelligence in healthcare: An essential guide for health leaders. Healthc Manage Forum. 2020 Jan;33(1):10-18. doi: 10.1177/0840470419873123.

[17] Chen M, Decary M. Artificial intelligence in healthcare: An essential guide for health leaders. Healthc Manage Forum. 2020 Jan;33(1):10-18. doi: 10.1177/0840470419873123.

[18] Jobin A, Ienca M, Vayena E. The global landscape of AI ethics guidelines. Nat Mach Intell 2019; 1: 389-399.

[19] Bioethics briefing note: Artificial intelligence (AI) in healthcare and research. Nuffield Council on Bioethics.2018. URL: https://www.nuffieldbioethics.org/assets/pdfs/Artificial-Intelligence-AI-in-healthcare-and-research.pdf.

[20] Morozov SP, Vladzymyrskyy AV, Klyashtornyy VG, Andreychenko AE et al. Clinical acceptance of software based on artificial intelligence technologies (radiology). In: Best practices in medical imaging. 2019; 57: 45.