Ethical Implications of Digitalization in Healthcare

Guido Lerzynski

1 Introduction

The digital transformation is profoundly changing healthcare, medicine, and nursing. Whether it is the storage of personal health information in electronic health and patient files, the creation and networking of medical databases, the use of artificial intelligence in diagnostics and therapy, or the deployment of health-related apps, the digital transformation is all-encompassing and rapid, with a significant impact on patients and the healthcare system. However, this transformation is inherently neither ethically good nor problematic. Rather, an ethical evaluation of each digital application is required, which relates to its specific utilisation (Mittelstadt et al. 2016; Wagner et al. 2017: 12). This evaluation must be based on certain parameters, as shown in Fig. 1. The evaluation of digital applications based on these parameters results in individual opportunity and risk profiles.

1.1 Responsibility of Patients

With regard to the patient, two questions arise when considering these parameters:

1. What is then ethically permitted or prohibited?
2. What rights and obligations does the healthcare system have towards the patient?

Conflicts of interest can arise at this point. For example, patients can only benefit from improved therapies if they disclose parts of their private data in return. The health of the patient may thus be at odds with their privacy and individual self-

G. Lerzynski (✉)
St. Marien-Hospital, Cologne, Germany
e-mail: guido.lerzynski@cellitinnen.de

Fig. 1 Core ethical principles. Data from Jannes et al. (2018). Source: Author

determination. In this respect, all parties involved in the healthcare system must weigh up which of the patient's rights are affected, and to what extent an impairment is justifiable. Healthcare professionals have a responsibility to respect the above-mentioned ethical principles towards the patient.

1.2 Responsibility of Institutions

Not only individuals are stakeholders in the development of the digital transformation in the healthcare system. Institutions such as data protection supervisory authorities must ensure that sensitive information is protected against unauthorised access. These institutions are responsible for creating framework conditions in which health-relevant data is processed appropriately and used in the best interests of healthcare. In many cases, they are faced with ethical challenges that cannot be overcome by a single stakeholder. However, in order to enable individual stakeholders in healthcare to deal with ethical challenges appropriately, institutional framework conditions are essential.

The increased use of digital technologies will lead to fundamental changes in the professional and activity profiles of medical professionals (Amarasingham et al. 2016). Wherever digitalisation can achieve better results than humans using traditional methods, corresponding tasks will be delegated to such systems. If an algorithm, e.g. for the analysis of images for the early detection of lung disease, achieves better results than human experts, it doesn't seem to make much sense to train and employ corresponding professionals in the current form. In the development of training occupations in the healthcare sector, they must in future aim to train professionals to use algorithm-based systems and to interpret and check the automatically generated results (Wang et al. 2016).

Institutions in the healthcare system should explicitly implement ethical principles and design structures in such a way that appropriate action by employees is encouraged. These structures can be designed to respond well to ethical challenges, because there is a high degree of mutual trust and competence, or in such a way that the individual can hardly hope for support within institutional healthcare

structures (Jannes et al. 2018). This becomes important when it comes to questions of responsibility in case of mistakes. Can the healthcare worker be held responsible for an error of an algorithm or more likely the software designer? To eliminate uncertainties, legally relevant questions must be reconciled with ethically acceptable approaches.

1.3   Responsibilities of Society

Ultimately, the digital transformation must be viewed in the context of social challenges. One of the aims of digitalisation in the healthcare system is a general improvement in healthcare and the early detection of diseases (Wilder et al. 2018). Therefore, digital applications can increasingly link and analyse data from different areas of life. This can result in both advantages and disadvantages for specific groups in society. For example, discrimination against marginalised or disadvantaged groups is possible. This is to be feared if algorithms are used to investigate the influence of lifestyle on the development of specific diseases. People who lead a lifestyle associated with an increased risk of disease could also be identified by the algorithm and excluded from certain medical services (Lippert-Rasmussen 2016). Linking the advantages and disadvantages with individualised insurance conditions can be highly problematic and ethically reprehensible. Therefore, core ethical principles must provide guidance for stakeholders and—more broadly speaking—for societies.

2   Pitfalls of Digital Applications

When selecting data to be processed for digital applications, standards and values should underlie the design of algorithms, which all have an ethical dimension (Kraemer et al. 2011; Mittelstadt et al. 2016). Algorithms are trained to process specific types of data. A set of basic data is used as a reference. This reference may already contain a bias, e.g. in the form of a prejudice, which determines the overall performance of the algorithm. An example is the malfunction of face recognition in a photo app provided by Google. The algorithm used there had been trained on the basis of image data which mainly included photos of people with fair skin. As a result of the limited data set, the programme was not trained to recognise people with dark skin colour as human beings. Instead, the automatic keywording function referred to them as gorillas (Jannes et al. 2018; Kasperkevic 2015). The discrimination against people associated with such a false classification is ethically unacceptable in any way. In the field of medical applications, it is not only hurtful, but also dangerous to health. When it comes to issues of mutual respect and security, neither individuals nor institutions alone can provide a solution to problems. Social discourse and political solutions are required here (legal regulations). Above all, there is a need for socio-political debate on the goals and purposes to be pursued. Should algorithms be used with the primary goal of reducing healthcare costs?

May algorithms developed in the healthcare sector also be used for commercial purposes? These and other questions are of a socio-political nature and require corresponding discourses and solutions. Further questions arise with regard to the possible effects of technology on future socio-cultural developments. Will there be health-related obligations in light of new technological possibilities? Will there be an obligation to record one's individual vital signs data, in order to make potential risks of illness recognisable at an early stage, and thus more cost-efficient to treat? These questions also require a broad public discourse, in which an awareness of possible developments is created. Responses to current and future challenges must be found, which meet the ethical requirements for observing the above-mentioned core principles. They should help to promote the ability to make decisions, to protect against potential harm and discrimination, and to distribute scarce resources fairly.

3 Opportunities and Challenges

3.1 Opportunities

There are many hopes and expectations associated with the use of algorithms in healthcare. Numerous current reports on projects for the development and use of algorithms in healthcare convey the impression that the realisation of fully digitalised healthcare is imminent.

In reality, many ideas and projects still have a long way to go before they can be realised and used on a practical level in healthcare, on a quality-assured basis. As in other areas, it must be expected that not all expectations will be met. The following description of the opportunities and challenges of algorithms in medicine and healthcare is to be understood as a description of expectations, wishes, and hopes. It also highlights the challenges that can be associated with the various applications. The aim here is neither a prognosis of the future nor an evaluation of the convictions and assumptions associated with the opportunities and challenges formulated. The use of algorithms in healthcare is associated with many expectations, some of them very high: a considerable increase in the speed with which health-relevant knowledge is gained in research and introduced into healthcare; a considerable broadening of the knowledge base and the range of medical services based on it; an increase in the precision of diagnoses and treatment recommendations, and associated with this, the medical safety of healthcare services (Dörn 2018: 352; Wired 2017; De Witte 2017). The automatic processing of a large set of health-related personal data is also associated with the hope of developing individualised medicine and reducing costs in the healthcare system (IBC 2017: 7; De Witte 2017).

The above-mentioned expectations of digitalised health research and care are primarily linked to the possibility of processing large amounts of data from different sources. However, the mere availability of a considerable amount of data by no means guarantees its meaningful evaluation. With regard to Big Data, experts criticise that in current applications the usual principles of science are often not

observed, and the principles of evidence-based medicine are violated (Antes 2016). The main criticism is that too little attention is paid to theory formation in data evaluation (Mayer-Schönberger et al. 2013: 70).

To be able to meaningfully analyse the data that will become available in the various fields of medicine with the ongoing digitalisation, it is necessary to edit and curate the data. This task can only be performed by human experts. However, they can receive valuable support from algorithms. Algorithms can be used to facilitate data analysis by training and using them to process precisely and exclusively the data that is necessary to achieve a specific goal, such as the prognosis of a complex disease. The use of algorithms thus promises to make it easier to handle an ever larger and more diverse set of different data generated in medical contexts.

Improvements are expected in particular from the ability of algorithm-based systems to automatically match a large amount of data in the shortest possible time. Here, the mechanical capabilities clearly exceed the corresponding capabilities of human stakeholders. Based on such data matching, algorithms can achieve the same or even higher accuracy than human experts. Especially in the case of rare diseases, they are even superior to humans in terms of diagnostics (Esteva et al. 2017; Rajpurkar et al. 2017). Algorithm-based image analysis methods allow, for example, an automatic quick check for potential skin diseases.

Moreover, algorithms are already being used to automatically detect drug interactions and side effects based on the evaluation of information from digital patient files and medical articles (Dörn 2018: 651). The number of inadequate or unnecessary treatments could also be reduced by improving findings. The use of algorithms can counteract possible errors caused by overworked employees. Algorithms thus contribute to increased security in healthcare. In addition, they can generally reduce the workload in medicine and care. They also open up new possibilities for automation processes in other areas. Many routine tasks, for example, in laboratory medicine, cardiology and radiology, could be taken over by algorithms in future (Rasche 2017).

3.2 Challenges

Given their high speed and their ability to process even the largest amounts of data, the performance of algorithm-based systems could easily be overestimated. Machine systems are indeed systematically superior to humans in the storage and management of data, and this superiority is likely to increase in the future. But when it comes to evaluating information, they are systematically inferior to humans. Human judgement is required in many, if not most, areas of medical/nursing care and research. If there are several diagnostic or therapeutic options, an algorithm can at best have a supporting function (Rasche 2017). This cannot replace human judgement. With algorithm-generated recommendations, it is therefore important to clearly distinguish between recommendations and decisions: digital assistance systems could make recommendations, but they cannot yet make a decision (Rasche 2017). Decision-making always falls to a human being. This also applies to the use

of algorithms in systems that, for example, automatically administer medication, trigger electrical impulses or send notifications to medical or nursing staff. One example is sensors implanted under the skin that record the blood values of diabetics in order to automatically release insulin when required.

Ethically and legally problematic implications may arise in connection with the programming, use and settings of the system, particularly with regard to the attribution of responsibility. Obviously, an algorithm can cause damage through poor-quality or even faulty programming or application. However, it would be nonsensical to claim that the algorithm is literally responsible for damage. Even highly developed algorithms are not able to assume responsibility. They do not make morally responsible decisions. Only humans can do that. So if damage occurs as a result of an algorithm application, those who were involved in the programming and application decisions are responsible. However, in view of the often-large number of people involved in such decisions, the question arises as to who is ultimately responsible for which factors and possible errors (Mittelstadt et al. 2016). Is it the programmer, the institution offering the system, the attending physician or the patient? The problem of attributing responsibility is exacerbated by technical aspects. Different types of algorithms sometimes raise different questions. In order to make decisions, people must have sufficient relevant information and practical decision-making knowledge. However, the different modes of operation of algorithms are sometimes hardly comprehensible, and sometimes not at all, even for computer scientists (European Group on Ethics in Science and New Technologies 2018).

Partially supervised or unsupervised machine learning poses the most problems. The individual steps of the respective processes are often no longer comprehensible, even for computer scientists and programmers. If the algorithm works incorrectly, people cannot recognise which step is the cause. Even with supervised learning algorithms, questions arise about transparency and the allocation of responsibility, e.g. between individual programmers and users. They are used to filter and process information. Thus, they influence human decisions. As a result, a mistake in data processing can lead to wrong human decisions, for example, if information relevant to the decision is classified as irrelevant. If experts rely on the performance of such an algorithm, decision-relevant factors can easily be overlooked. In the worst case, the awareness that decision-relevant information can be overlooked by an algorithm is lost (Mittelstadt et al. 2016).

Further challenges are associated with the so-called bias phenomenon. Bias means that the processing rules of an algorithm lead to a systematic distortion or bias. Algorithms are used to automatically analyse and group cell samples with regard to certain disease markers (Kraemer et al. 2011). In many cases, such grouping will be unambiguous. In other cases, however, the classification may be unclear. In such cases, a threshold value must be defined that determines whether a cell sample is marked as disease-relevant or not. Such a limit is a norm, and when determining it, it must be weighed up which consequence is more likely to result: a possibly more frequent false positive alarm or a possibly higher proportion of samples marked as false negative (Kraemer et al. 2011). A bias can

also result from an algorithm operating on an insufficient data basis. This may be because the algorithm, as in the above-mentioned example of Google's image recognition algorithm, was trained with insufficient and above all one-sided data sets. But it can also be due to the incompleteness or inconsistency of data sets in the process of applying a learning algorithm. Health-relevant data has often been recorded incompletely so far. Data in patient records is often coded insufficiently or inconsistently and information is incomplete. Such shortcomings have an impact on the performance of algorithms, which often cannot evaluate such data or can only do so inadequately. Further imbalances in the database can also be caused by the fact that there is a particularly large amount of data available from certain groups of people, but only little from others. Patients in hospitals that already work digitally produce more data than those in less digitalised hospitals. Such an imbalance can also lead to a bias (De Laat 2017). Bias-related failures can significantly affect the reliability of systems in practical use. The analyses generated are inevitably either incomplete or even incorrect. The above-mentioned chance that the use of algorithms can significantly improve the safety and reliability of health services is therefore currently only limited.

It remains to be seen whether the problems caused by the various types of bias can be remedied in the future. Automatic or semi-automatic processing of digital content can ultimately only work if a sufficient degree of interoperability of different systems is ensured. It is therefore important to develop and establish common standards for data exchange (cf. Chap. 2). At present, however, there are sometimes considerable deficiencies in this area.

4   Conclusion

It is indisputable that digital algorithms can contribute to improving care, but their use also raises ethical questions: about distributive justice and protection against discrimination, about liability for algorithm-based decisions, about the upcoming changes in the relationship between doctor and patient, and about trust in the healthcare system as such. A broad understanding is therefore needed about which developments we as a society should support and demand on the one hand, and where boundaries must be drawn on the other. One of the tasks of digital ethics is to establish the effects of digitalisation on society and the individual, and to develop consistent justifications for moral action and normative standards. Furthermore, it can serve as a navigational tool for questions of values and norms associated with new technologies and the resulting social-communicative practices. Its aim is to promote value-based digital literacy, in order to develop a better understanding of how algorithms work and behave.

References

Amarasingham, R., Audet, A.-M. J., Bates, D. W., Cohen, G., Entwistle, M., Escobar, G. J., Liu, V., Etheredge, L., Lo, B., Ohno-Machado, L., Ram, S., Saria, S., Schilling, L. M., Shah, A., Stewart, W. F., Steyerberg, E. W., & Xie, B. (2016). Consensus statement on electronic health predictive analytics: A guiding framework to address challenges. *eGEMs (Generating Evidence & Methods to improve patient outcomes), 4*(1), 3.

Antes, G. (2016). Ist das Zeitalter der Kausalität vorbei? *Zeitschrift für Evidenz, Fortbildung und Qualität im Gesundheitswesen, 112*(1), 16–22.

De Laat, P. B. (2017). Algorithmic Decision-Making based on machine learning from big data: Can transparency restore accountability? *Philosophy & Technology, 31*, 525–541.

De Witte, B. (2017). Algorithms will outperform doctors in just 10 years time. *Dataconomy*. http://dataconomy.com/2017/10/18532/

Dörn, S. (2018). Digitale Technologien. In S. Dörn (Pub.), *Programmieren für Ingenieure und Naturwissenschaftler* (pp. 333–370). eXamen.press. Berlin, Heidelberg.

Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., & Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. *Nature, 542*(7639), 115–118.

European Group on Ethics in Science and New Technologies. (2018). *Statement on Artificial Intelligence, Robotics and 'Autonomous' Systems*. Brussels.

International Bioethics Committee (IBC) United Nations Educational, Scientific and Cultural Organization (UNESCO). (2017, 15 September). *Report of the IBC on Big Data and Health*, Paris. http://unesdoc.unesco.org/images/0024/002487/248724e.pdf

Jannes, M., Friele, M., Jannes, C., & Woopen, C. (2018). *Algorithmen in der digitalen Gesundheitsversorgung. Eine interdisziplinäre Analyse*. Gütersloh: Bertelsmann Stiftung.

Kasperkevic, J. (2015, July 1). Google says sorry for racist auto-tag in photo app. *The Guardian*. Accessed September 7, 2020, from www.theguardian.com/technology/2015/jul/01/google-sorry-racist-autotag-photo-app

Kraemer, F., van Overveld, K., & Peterson, M. (2011). Is there an ethics of algorithms? *Ethics and Information Technology, 13*(3), 251–260.

Lippert-Rasmussen, K. J. (2016). Is health profiling morally permissible? *Medical Ethics, 42*, 330.

Mayer-Schönberger, V., Cukier, K., & Mallett, D. (2013). *Big Data. Die Revolution, die unser Leben verändern wird* (2. Auflage). Munich.

Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The ethics of algorithms: Mapping the debate. *Big Data & Society, 3*(2), 205395171667967.

Rajpurkar, P., Irvin, J., Zhu, K., Yang, B., Mehta, H., Duan, T., Ding, D., Bagul, A., Ball, R. L., Langlotz, C., Shpanskaya, K., Lungren, M. P., & Ng, A. Y. (2017). *CheXNet: RadiologistLevel pneumonia detection on chest X-rays with deep learning*. arXiv, 1711.05225v3.

Rasche, C. (2017). Digitaler Gesundheitswettbewerb: Strategien, Geschäftsmodelle, Kompetenzanforderungen. In M. A. Pfannstiel, P. Da-Cruz, & H. Mehlich (Pub.), *Digitale Transformation von Dienstleistungen im Gesundheitswesen I. Impulse für die Versorgung* (pp. 1–30). Wiesbaden.

Wagner, B., & Vieth, K., on behalf of the Bertelsmann Stiftung. (2017). *Teilhabe, ausgerechnet*. Gütersloh.

Wang, D., Khosla, A., Gargeya, R., Irshad, H., & Beck, A. H. (2016). *Deep Learning for Identifying Metastatic Breast Cancer*. arXiv preprint arXiv:1606.05718.

Wilder, B., Suen, S.-C., & Tambe, M. (2018). *Preventing infectious disease in dynamic populations under uncertainty*. Association for the Advancement of Artificial Intelligences. www.cais.usc.edu/wp-content/uploads/2017/11/aaai_tb_final.pdf

Wired. (2017). *Behandlungen und Technologien der Medizin der Zukunft*. Accessed August 22, 2020, from www.wired.de/collection/science/medizin-der-zukunft