There are currently two major trends in medicine. The first is digitalization: In radiology and conventional laboratory medicine, the daily routine has already been digital for quite some time. Patient files, pathology, microbiology, and virology are all digitalized to varying degrees but complete digitalization can soon be expected.

The second major trend in medicine is “personalization”: At present, this primarily refers to personalized pharmacotherapy which takes the individual physiological constitution as well as molecular-biological constellations into consideration.

If you look outside of the medical field, there is a further megatrend: Artificial intelligence (AI). For reasons that will be explained below, this trend has not yet properly arrived in the medical field but has become indispensable in industrial manufacturing and in the processing of large amounts of data. Facebook, Amazon and Google would not function without AI.

The challenge now is to implement digitalization and AI in medicine in such a way that personalized medicine becomes routine. Personalized medicine cannot become a reality without AI: The amount of data is so immense that it will no longer be possible to rely on the personal knowledge of a few experts.

Why an essay?

You were probably expecting a scientific review describing the possibilities for using AI in radiology and medicine. I selected the essay format because black swans cannot be convincingly presented on a scientific basis. The black swan is a symbol for the unexpected (if you would like to learn more about this topic, read the book “The Black Swan” by TALEB). Let’s look at the terror attacks on September 11, 2001 as an example of a black swan: Any security expert could have predicted this terror scenario in an academic paper. It was possible to hijack a plane. It was possible for a hijacker to have a pilot’s license. It was possible to switch off the airplane’s transponder so that its altitude, call sign, and speed were no longer being recorded. Even if a security expert had predicted this scenario in a paper, it probably would not have been accepted for publication. Thinking the unthinkable causes too great a disruption of our world view.

The automobile was also a black swan. The route from Mannheim to Pforzheim would have been traveled much faster by horse while avoiding the hassle of having to buy fuel at a pharmacy. An article regarding the potential benefits of the automobile would have been rejected by reviewers as implausible. Bertha Benz’s historic journey was the more important story.

Other black swans include the demise of the Siemens telephony branch, the disappearance of typewriter manufacturers from the market, the abandoning of atomic energy driven by the CDU, the ascent of Google and the downfall of Blackberry. What can we learn from this? Black swans are actually not all that rare. They are repeatedly underestimated by science and can turn the world upside-down.

OK, but we are in the medical field. Black swans are not taboo. Let’s take a look at the gastric and duodenal ulcer. It was considered psychosomatic until the middle of the 1980’s. A hypothesis was developed and confirmation was sought until everyone believed – that is until two Australian colleagues proved their theory regarding the bacterial origin of the ulcer in a self-experiment, in a manner that was completely unmedical and essayistic in a way. Or let’s talk about anorexia, a highly monomorphic disease caused, based on current understanding, by a combination of mental and social factors. At some point we find the true cause with the help of hypothesis-free research and AI.
By the end of this section you should be mentally prepared to recognize the existence of black swans and consider it possible for black swans to change medicine. And you should be prepared to accept an essay in a scientific journal.

What is the difference between CAD and AI?

With digitalization, radiology began developing algorithms to make diagnosis easier. Computer-aided diagnosis (CAD) is the keyword here. Many AI skeptics refer to the lack of breakthrough of CAD systems as justification for their skepticism.

As a rule, CAD systems require a hypothesis and this hypothesis must then be mathematically modeled. Let’s use CAD systems for MR mammography as an example. The hypothesis is that contrast enhancement and washout are the decisive characteristics of breast cancer. Therefore, techniques were developed to clearly visualize these contrast kinetics on MR, standard populations were examined, and the differences between normal tissue and cancerous tissue were mathematically modeled. If the hypothesis is correct, good results can be achieved with such CAD systems. An AI system learns without hypotheses, i.e., not only the contrast kinetics but also many other types of image data (T2, SWI, diffusion, everything the system contains) are applied to the algorithm. Clinical data, laboratory results, and genetic information can also be entered into the system for learning purposes. The algorithm then uses all of this data to determine the parameters that indicate or rule out breast cancer with the greatest probability. Contrast kinetics will certainly ultimately play a role in the decision but the algorithm will also find that other parameters are of prognostic importance with varying degrees of weighting. Without a hypothesis. My intention here is not to discuss the various forms of AI with keywords like “machine learning”, “deep learning”, “supervised”, and “unsupervised”. The point is that AI learns without a hypothesis. And as I will discuss later, the quality of AI applications depends on the quality of the data used to train the system.

Routine uses of AI in radiology

An advantage of computers is that they do not tire and can perform routine tasks in a highly reliable manner. What are some of these tasks in radiology? There are none? Well in my opinion counting MS plaques, determining ventricular size following SAB, measuring liver metastases during treatment, assessing hand bone age, and measuring angles in scoliosis are not particularly exciting tasks. AI will quickly take over such routine tasks. All we need is algorithms that were trained to reliably detect the target organ and the lesion. And then the algorithms must be able to quantify. These applications are already available but largely still in the initial stages. OK, it’s not perfect, but we must remind ourselves that the first road trip was also slow. Ultimately, instead of a flowery description of MS plaques at various locations in the brain with or without subtle contrast enhancement, the finding will simply be “plus 19” (treatment is not going well) or “minus 4” (looking good). The radiologist can then concentrate on minimizing one of the most common errors in radiology – satisfaction of search: An aneurysm of the anterior communicating branch or a pituitary adenoma is overlooked because of all of the plaque counting. Another application that is currently being pursued at many places around the world is the evaluation of conventional chest X-rays. An incorrect endotracheal tube position, pacemaker electrodes, a central venous catheter, a pneumothorax, tuberculosis, pneumonia in the medial lobe as well as 40 other pathologies can already be reliably detected. There are algorithms that can detect and automatically interpret 70 or even 90 of these pathologies. One problem with these applications is that they are currently poorly integrated in the workflow of the radiology department or the radiologist. But this can quickly improve with our help. Finally, a small company presented a product capable of diagnosing 40 typical pathologies on conventional chest X-rays. The algorithm was perfect for simple things (central venous catheter, endotracheal tube, etc.) but catastrophic in the case of pulmonary fibrosis. Why? The training data came from a third-class hospital. Therefore, another and much more serious problem with respect to establishing AI in radiology is the training data, as I will discuss in the following.

I will conclude this section with a brief discussion of screening. Imagine that China started a mammography screening program. It could never be run by radiologists. There are far too few radiologists and far too many women requiring screening. Of course, such screening in China would have to be performed by properly trained algorithms. And now imagine a lung screening program in the USA and Germany. What radiologist would enjoy spending the day looking at screening CT scans of the lung? During a presentation of AI applications at the ECR 2018, a discussion was held as to whether it would be better to use the “AI plus 1 radiologist” principle instead of the “four eyes principle” with 2 radiologists in mammography screening. In other words, the black swan was discussed. Of course, we can resist. But will that make a difference?

Added diagnostic value of AI: Radiomics

Image data may be able to provide more information than previously thought, i.e., a black swan. However, the amount of data is so large that this additional information is missed by the radiologist. In radiomics, an algorithm is again trained with known data. Let’s use glioblastomas as an example. The algorithm receives all MR image data (from T1 native to T2, SWI, DWI, and contrast-enhanced images) as well as molecular-genetic information. The algorithm can then detect not only the molecular differentiation of the tumor but also the extent of the blood-brain barrier dysfunction in unknown MR data sets without the administration of contrast agent. Utopia? No, it’s already a reality. Not everywhere but yes in some departments.

Let me give you a few more examples: We trained an algorithm to detect with over 95% probability whether a patient with cerebral cancer has organ metastases simply by viewing the tumor on PET-MRI scans. With appropriate training data, AI can detect the molecular-biological typing of tumors without biopsy, predict whether a tumor will respond to radiotherapy, and predict the probability of response to chemotherapies. All just a Utopian
dream? At this point allow me to briefly deviate from the essay style and refer to the list of references: It provides an alphabetized list of various examples of the use of AI in radiology [1–39].

Why has AI not yet become properly established in medicine?

AI is already the standard in many areas of manufacturing and robotics. But we only recently started talking about AI in medicine. Why is the medical field lagging behind when it comes to AI? Perhaps because medicine (including radiology) is much too complicated? Certainly not. Pattern recognition – and this applies to more than just radiology (as I will explain later) – is a strength of AI. Military applications of AI are largely based on the recognition of patterns. A fundamental requirement for training intelligent algorithms is the use of valid training datasets. If we take a look at the players on the AI scene (Google, Microsoft, Amazon, etc.), it becomes clear that they do not actually have access to valid medical datasets. Google probably knows better than any public health officer where the next flu epidemic will occur because people increasingly use the search engine to search for symptoms. However, valid radiology datasets are not so easy to obtain. But what about Siemens, GE and Philips? OK, they could theoretically look at numerous CT and MRI datasets (for our purposes here we will disregard the issue of data protection) but, in simple terms, the diagnoses don’t have to necessarily be correct. “Old economy” and “new economy” cannot introduce AI in medicine (and radiology) without some effort.

OK then: Who has the necessary knowledge and materials? Textbook publishers would be one answer. For many years now, publishers have required authors to illustrate a disease with more than one image. Authors are expected to illustrate an epidural hematoma with a 3D dataset. Therefore, publishers have textbook knowledge, also known as “ground truth” in IT speak. We don’t need “big data” to train intelligent algorithms but rather only valid data and well annotated images (N > 1). Publishers probably simply do not have enough data to train an algorithm, for example, to determine bone age in children. And that brings us to the solution: Large hospitals and especially university hospitals have such data. Unfortunately that data is disorganized and poorly sorted and sometimes even incorrect. Therefore, we need something like “data validation” departments in large hospitals to check data (radiology, lab, pathology, electronic patient file, etc.) for validity and add it to a data pool with a structure that is accessible for AI. Not possible? We’ll see. When hospitals were still reimbursed by health insurance based on the number of occupied beds, we didn’t need any coding departments.

And now? DRG made it possible. OK, you can’t forget about the economic pressure. True. But AI creates economic opportunity! Imagine that you have significant data regarding prostate cancer and you train an algorithm with image data, lab results, and molecular-biological, genetic, and pathological data. And then you also have the follow-up data from hundreds of patients. If there is a pattern for the benign and malignant prostate cancer variants, AI will find it in the sea of data. And then you have an algorithm that no longer contains patient data but only the results of the training and can be sold as a product together with a strategic partner. Great idea, but what if everyone does it? We can’t all do everything! We could do the “liver” but not “rheumatism”. Everyone has their area of expertise. Moreover, there is more than one company manufacturing cars and MRI units. In the end, one product is perhaps better integrated into the workflow, another is faster, the third is trained with a different number of parameters and a varying degree of data validation. Therefore, there are sufficient features that differentiate individual applications from one another. And independent of economic opportunity: AI will be a major research area in medicine and radiology. A noteworthy secondary effect for university hospitals.

One last thing to conclude this section: Above I said that Google and Amazon have difficulty obtaining valid medical datasets. So far this is true. However, in the future, both of these companies could operate their own hospitals, perhaps initially only for employees and only on an outpatient basis. Then they’d have data structured according to their needs. And the health care market is so enormous that the Googles of this world are going to enter the market. You don’t think anyone will go to an Amazon hospital because data protection could be a problem? How many people buy from Amazon every day?

Super diagnostics

Can you now imagine the major effect AI will have on radiology? Independently handle boring tasks and usher in a new era from detection to prediction with radiomics. Let’s take a quick look at other diagnostic areas: AI will learn to look at the microbiologist’s petri dish and identify that the dish contains streptococcus. And AI will learn to predict resistance to HIV medication. Digital pathology is ultimately like radiology with color, also pattern recognition. AI will learn routines quickly. Pathomics will be the term for deeper analyses. AI will bring genetic analysis to another level. And in the future it will be able to diagnose diseases from certain constellations based on laboratory results.

In a nutshell: We will soon have AI in all diagnostic disciplines and the correct personalized (see above) treatment decision will increasingly come from these super diagnostic centers. Is your head spinning? Data validation department and super diagnostic center.

Too much? But if you already have a data validation department, the step to becoming an SD center is very small. It needs to be structured but in a way it is like a tumor board, albeit a scaled-down version, because some diagnoses and treatment recommendations will be made automatically, e. g., the designated standard therapy will always be recommended for a glioblastoma with a specific mutation. Of course, this applies not only to tumors but also to a broad range of diseases. Who can create such super diagnostic centers? Only university hospitals because as a rule only they provide all diagnostic disciplines and are thus able to merge the data in a structured manner. Perhaps super diagnostics will become a main feature of university medicine.
Ethical and legal dimension of AI

When I occasionally give a lecture on AI there are not infrequently people in the audience who have studied both medicine and law and there is always someone in the audience who raises a warning finger and mentions the word “ethics”.

Let’s discuss ethics. We agree that medicine always has an ethical dimension. Therefore, it is important to determine the specific ethical dimension of AI. If the topic of the presentation was: “Making medicine better with more intelligence”, there wouldn’t be any ethical concerns. The issue arises because it is “artificial intelligence”. Let’s replace “artificial intelligence” with “computer intelligence”. Did anyone mention “ethics” when computed tomography replaced palpation of the liver? I can’t remember. In fact, I find it strange that an ethicist has not yet questioned the ethics of not using an AI technique when it can ensure faster and more reliable diagnoses. And now we need to talk about the law. Following one of my lectures, a judge sitting in the audience stood up and told me – well not me personally but certainly everyone with such heretical thoughts – there would be major legal problems because judges would not be able understand how AI determines a diagnosis. And what would happen if AI is wrong? Major consequences.

Time for some de-escalation. Google has been planning for some time to make online mammography interpretation possible. For one dollar. This is not yet a reality. Maybe it will cost 5 dollars. But it’s coming. Now imagine the following scenario: two German radiologists interpret a mammogram as normal and then half a year later the woman is diagnosed with liver metastases of undiagnosed breast cancer. The husband of the patient then uploads the scans to Google and the algorithm says that breast cancer was diagnosed. The husband of the patient then uploads the scans to Google and the algorithm says that breast cancer was diagnosed. The husband of the patient then uploads the scans to Google and the algorithm says that breast cancer was diagnosed. The husband of the patient then uploads the scans to Google and the algorithm says that breast cancer was diagnosed. But the back pain is still there. We are all familiar with such cases. And then after 2 years (on average), the patient undergoes more than one MRI examination of the lumbar spine over the course of 2 years (the lumbar spine shows age-typical changes that we – unfortunately – always also describe in detail) and she receives countless mud pack and fascia treatments and perhaps even undergoes lumbar spine surgery with fusion of 4 segments. But the back pain is still there. We are all familiar with such cases. And then after 2 years (on average), the patient is diagnosed with “depression”. If you analyze the photos this woman uploads on Instagram or better yet look at her “daily activity” on Facebook, you can immediately determine the diagnosis. Or you use the algorithm that was trained to detect the speech melody and pitch of voice of depressed patients. You have your answer in 2 minutes. Even for a radiologist. And as soon as digital data from narrative-based medicine is available – Google is feverishly working on it – narrative-based medicine will completely change.

But radiology will remain otherwise unchanged?

No, of course not. You have accepted that AI is pushing radiology into the radiomics dimension and super diagnostic centers are the future. But the radiology business model will also change. When I started in radiology in 1987 everything was under one roof. The physician sat at the unit and viewed the images on the monitor. And there was a film that had to be developed and was used for reporting. The images were interpreted in the radiology rooms. Today, some of my colleagues work from home. And today, MR units can even be operated remotely. Highly qualified technologists sit in a room far away from all MRI units and simultaneously perform two or three MRI examinations. In the unit and the waiting area, non-technologists perform patient management including positioning. Positioning is less complicated with new surface coils designed as blankets and thanks to AI the device detects the region to be examined increasingly independently.

And then there is also Amazon’s business model – the platform. You could also look at Uber or Airbnb. Images are uploaded to a broker platform and any radiologist with time and desire can
interpret them in a quality controlled manner. This represents a huge opportunity for hospitals that can’t find radiologists and for large collaborative practices and hospital departments to grow without making an investment. And perhaps the indication for MRI examination of the lumbar spine in back pain is no longer exclusively determined by orthopedists but by AI which accesses the patient history and uses the results of the physical examination entered by the physician. Or perhaps this is done by a physician assistant. And then the patient goes to the MRI unit that is located at Aldi around the corner and is remotely operated. And the image is then interpreted on the platform.

OK, you’re right. This sounds like a Utopian dream and perhaps won’t be a reality in the near future. Or maybe things will happen a bit differently. But think about the black swan at the beginning. Don’t think about why it’s not possible (laws, ethics, data protection, unions, medical associations, power outage). I know that searching for and finding problems is viewed as a highly intellectual pursuit in Germany. The introduction of the RIS-PACS in 2001 was greeted almost entirely by skeptics and those predicting problems both within and outside radiology. And the transition phase – previous analog images in the archive and current images in the PACS – was not fun. It will be the same with AI. But the technology will come and you have the opportunity to actively help shape it. So start thinking of solutions.

Michael Forsting

References

[28] Oorooji M et al. Combination of computer extracted shape and texture features enables discrimination of granulomas from adenocarcinoma on chest computed tomography. J Med Imaging (Bellingham) 2018; 5: 024501


