# Artificial Intelligence-based algorithms for chest X-rays: a development with worldwide potential

*For several years now, the private radiology practice of MVZ Prof Dr. Uhlenbrock situated in the Ruhr area of Germany has been cooperating with Siemens Healthineers on the development and validation of AI-based algorithms for the interpretation of chest X-rays and the provision of clinical decision support. The results have been extremely promising and the resulting software package, the AI – Rad Companion for Chest X-ray is now CE-labelled and available commercially.*

Dr. Karsten Ridder is
Lead radiologist in the practice of
MZV Prof Dr. Uhlenbrock & Partner.
E-mail : k.ridder@mvz-uhlenbrock.de

*We spoke to Dr. Karsten Ridder, lead radiologist of the group practice on his experience of the rigors of the intensive development process and the satisfaction of seeing the end-result performing well in real-life clinical practice.*

**Q** *Before we get into the details of chest X-ray and the AI software, please give us a brief description of your practice.*

The practice was founded in 1993 by Prof. Detlev Uhlenbrock and since then has grown and established a reputation as being one of the leading medical centres in the Ruhr area in Westphalia, Germany. The practice is composed of 13 individual clinics, most of which are associated with, and linked to, a hospital. I am principally involved with the St. Lukas group of hospitals in Dortmund. Each year in the group we see more than 290 000 patients of whom three-quarters are out-patients. We are equipped with a complete range of modern equipment offering all imaging modalities ranging from DR X-Ray to high-end cardiac CT, from plain ultrasound up to PET-CT; and from MSK-MRI up to 3Tesla neuro-imaging.
Our philosophy has always been to be ready to take advantage of the latest technologies, provided always that there is a clinical benefit.

**Q** *Now let's focus on chest X-rays.*

Chest X-ray examinations (CXR) are one of the most frequent radiological procedures carried out world-wide, and this is of course also true in our practice where last year we carried out as many as 30000 chest examinations, a number that was slightly higher than usual because of the COVID pandemic. All the individual clinics in our practice are equipped to carry out chest X-rays, using equipment from various manufacturers, such as Siemens Healthineers, Carestream and Agfa. The vast majority of these are DR systems, although we still have a few old CR-based devices, which are scheduled to be updated in a couple of months.

**Q** *And now let's turn to AI-based algorithms used for interpreting the CXR images, and your collaboration with Siemens Healthineers on their AI-Rad Companion Chest X-ray project.*

In fact our collaboration with Siemens Healthineers on this topic goes back four years. At that time new German government legislation had just been introduced which required us to carry out a CXR examination for every newly arrived immigrant who had settled in our catchment area. At the time there was a real surge in the numbers of immigrants coming to our part of Germany, so the practical effect on us was that we were suddenly faced with a huge increase in our workload, with the numbers of CXR exams sometimes reaching as many as 100 procedures per day. Somehow or another we managed to struggle through this tough period, but for me the message was already clear — we urgently needed support to ease our workload. I immediately thought of the potential of systems based on artificial intelligence to help us and so reached out to Siemens Healthineers. After a fair number of email exchanges and phone calls, we initiated our joint collaboration project

**Q** *What exactly is the role of your group in this collaboration?*

In fact we have several important roles. One is the supply of a large part of the image dataset which is a vital part of the process of using AI methods to train an algorithm for the interpretation of radiology images. It is well known that the performance of a final algorithm is directly dependent on the quantity and quality of the information used to train it. In radiology that doesn't mean just high quality images — without accurate annotations and interpretative reports, images alone are of no use. So it all boils down to

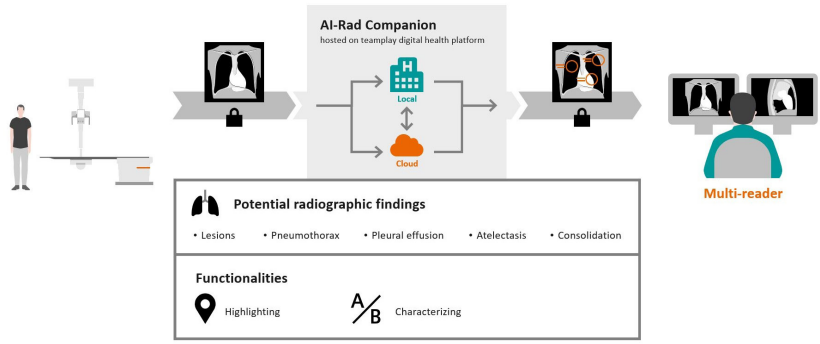


**Figure 1.** The development of an algorithm *per se* is of limited use if it cannot be efficiently integrated into the clincal routine. The system is intended to be used by a radiologist concurrently with original images before a final decision is made on a case The above diagram illustrates where the AI-RAD Companion Chest X-ray package is located in the clinical workflow.

good images and accurate reports (and in sufficient numbers).
In practice the process of introducing images into the training pool is really rigorous and quite time-consuming. Every image must be validated technically. The corresponding reports are then extracted, evaluated and compared by specialized radiologists. If there are any doubts, misinterpretations or low-quality images – then that individual dataset is excluded. Once the rigorous control and acceptability criteria of new input data have been satisfied, you progressively end up with a very large and solid database that can be used to train the algorithm. Such a database is an extremely valuable resource for companies developing the algorithm. For this part of our collaboration we are fortunate that in Dortmund we have a very strong IT team, who dealt with the practical coordination and data interfacing with Siemens Healthineers.

Of course the development of powerful algorithms *per se* isn't of much practical use if they can't be integrated efficiently into the routine workflow of the radiology department. We spent many hours in discussions on how best to optimize this, drawing from all of our 30 years clinical experience of radiology in the front line. To do this we had to look hard at and analyze how our existing workflows are organized. These can vary depending on whether we're dealing with out-patients or hospitalized patients. During normal working hours out-patient examinations are carried out in real time, with most of the reports being send to the referring clinician within 2 hours. For hospitalized patients, we group the cases that have occurred over a suitable preceding time period, e.g. day/night/weekend. We then present and discuss the cases every weekday in several meetings which also include non-hospital colleagues. No images will ever leave our facilities without being accompanied by an evaluation and a report by a consultant radiologist.
Just how the new package is finally integrated in the workflow is shown in Figure 1.

Yet another part of our collaboration with Siemens Heathineers is in the equally important role of the extensive rigorous evaluation, and re-evaluation, of the performance of the algorithms and comparison with the results of our own readers.
Last but not least we are now the first site actually running the system in real life, in the three hospitals of the Lukas group in Dortmund. So we are not just getting the practical benefit of its support in our day-to-day examination of patients but we are also gaining valuable experience of how the system performs in the harsh and sometimes unforgiving conditions of busy modern radiology departments.

It must be noted that there are two other groups in the collaboration with Siemens Healthineers for the development of the AI-Rad Companion, namely a unit in the Universtity Hospital Munich (LMU) and another American group in Massachussets General Hospital in Boston. We meet regularly in online-meetings which are characterized by a clear understanding of our collective aims, a cooperative team spirit and mutual respect.
The idea that this international project is one with the potential to ultimately benefit radiologists and patients throughout the world is, I admit, a powerful personal incentive for me, and makes the additional effort to my "normal" work as radiologist in a busy medical practice more palatable.

**Q** *The AI-Rad Companion Chest X-ray software accepts DICOM images from X-ray devices such as Siemens Healthineers' Ysio X.pree. Can it also accept images from other X-ray systems?*

Yes. Right from the start, the objective was to develop software capable of working with images from X-ray equipment from all vendors. At the moment the system has been shown to successfully handle images from X-ray machines from Siemens Healthineers, Agfa, Carestream, Fuji, Varian and Konica Minolta at the same quality level of acceptability and performance.

**Q** *Do you need any additional computing capacity in your center?*

No. The system is completely cloud-based and operates on Siemens Healthineers' Teamplay digital health platform. Each anonymized dataset with images and information is sent automatically from our center to the platform on a 24/7 basis. It is then processed and evaluated by the software and sent back in approx three minutes to our PACS and presented as an additional image in the case. The turn-around time could probably be shortened even further but there is a limit as to how much the data transfer to and from the cloud can be accelerated. The data security aspects of such a mode of operation are extremely important and were always a top priority right from the start of the project. However for several reasons I have no worries about data security. Firstly, we only send anonymized data to the cloud and secondly the servers that handle the data are all physically located in the EU and so subject to — and compliant with — the strict EU certification criteria needed

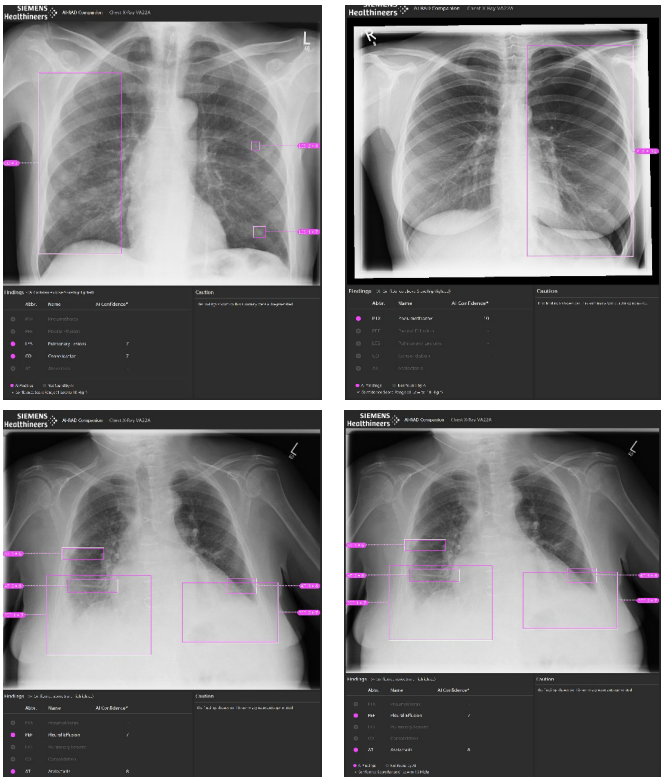

**Figure 2.** Typical examples of outputs of the AI-Rad Companion Chest X-ray software. The software identifies and highlights any "Big Five" pathologies that it detects in Chest X-ray, namely Pulmonary Lesions, Atelectasis, Consolidation, Pleural Effusions and Pneumothorax. Corresponding AI-confidence scores are also presented, indicating the level of confidence in the findings.

to handle sensitive data. But I can't say whether I would be so confident about data security if we were dealing with data processing systems in other continents.

**Q** *What about validation of the system before you implemented in routine clinical practice ?*

As I said, right from the beginning of the project, close scrutiny of the quality of the results was an on-going process and accompanied the development and improvement of the system step-by-step. It was really encouraging that even at the very initial stages of development the quality of the results looked highly promising. After several development and improvement cycles, we were confident enough to carry out a formal double-blinded multi-case, multi-reader validation study. The protocol was designed to not just compare the results generated by the software with the ground truth scenarios established by a panel of highly experienced radiologists but also to evaluate inter-reader variability. A total of 1019 patients were selected from our practice and from our two partner centers in Boston and in Munich, with the patient cases being representative of the typical patient population we see. We analyzed the classical performance metrics, namely sensitivity, specificity and calculated the area under the Receiver Operating Curve (AUC). The results showed that the AUC for the software was consistently and statistically significantly greater than that of the radiologist readers, for each of the "Big Five" pathologies in Chest X-ray, namely Pulmonary Lesions; Consolidation; Atelectasis; Pleural Effusions and of course Pneumothorax. After all the hard development work, the results of the study were extremely gratifying and encouraged us to implement the system in routine practice.

**Q** *In your hospitals, is the AI-Rad Companion software used for X-ray images acquired in the emergency room or in the radiology department or both?*

This is in fact the beauty and the power of the system — thanks to the high performance IT-network linking all of the collaborating hospitals in our group, the AI-Rad-Companion Chest X-ray can operate anywhere there is access to the IT network. Thus the algorithm can be used 24/7 just as easily in the emergency room as in the ICU and of course in the radiology department. We also use it in our regular daily meetings with our clinicians, in which typically we discuss cases "right in time" and where we usually don't have yime to prepare detailed case presentations beforehand

**Q** *The AI-Rad companion is designed as a "concurrent reader" to highlight any abnormalities on Chest X-ray. How in practice do you as a radiologist use it?*

The algorithm has been trained to detect, evaluate and highlight the "Big Five" in daily chest X-ray routine, namely consolidation, atelectasis, pleural effusion, masses and pneumothorax at a performance level which is comparable to or — as the validation study showed — even better than that of an average radiologist, For me as a senior radiologist it is not just helpful to avoid any accidental carelessness but also as an efficient means to prioritize workflow by clinical importance. In practice I look at the original images and the software - annotated images simultaneously and in parallel.

As ultimately responsible for the report and the interpretation of the images, I however still follow the good old, robust methods and guidelines for systematic scanning of Chest X-ray images. All radiologists have their own favorite routine image-reading method designed to minimise the chance of errors but it is reassuring to know that the software is there as a further guarantee against any errors caused by tiredness or disturbances.

But perhaps the biggest benefit of the system is for less experienced radiologists or even non-radiologists such as surgeons or anesthesiologists. Especially in the early formative stages of their clinical experience, such personnel can often be faced with the need to take urgent, important decisions often in the tough conditions of a crowded emergency room in the middle of the night. The support of the results from the AI-Rad Companion is extremely helpful in such situations.

**Q** *How useful are the confidence scores that the software generates?*

The software flags detected abnormalities but also shows confidence scores which are an indication of how certain the system is in the characterisation of the abnormalities [Figure 2]. These are very helpful. Currently the confidence score is expressed as figure on a scale of 1 (low confidence) to 10 (high confidence). We have found that if the score is above 7, the system is reliably correct in over 95% of cases, which of course is very impressive. We interpret a score of 6 as meaning that it is worth having a quick further look at the flagged abnormality since it isn't clear what exactly is behind it.

One key question regarding the performance of the software is how

many abnormalities are identified that were completely missed by the radiologist. This depends totally on the level of experience of the radiologist. The current performance of the software is comparable with that of a mid-level radiologist. This is in fact a key take-home message coming out of all of our studies – the system can help relatively inexperienced radiologists, or non-radiologists, to significantly improve their performance and, crucially avoid potentially dangerous mistakes caused by misinterpretation or lack of experience.
The other side of the coin regarding the number of abnormalities detected by the software but not by the radiologist is the inverse, i.e. the number of lesions only detected by the radiologist but missed by the software. The answer to this sounds too good to be true, but the data are clear — with the current threshold setting, the results we have shown that the system effectively misses none of the Big Five CXR pathologies on which it has been trained.
Of course since the software has not (yet) been trained to detect non-lung incidental findings such as a broken rib or a projectile in the chest, the software will (for the moment) not pick them up

**Q** *What about the legal responsibility for the diagnosis when the AI Rad Companion is used?*

That's an important question but one with a clear answer. Even if the software is a very useful support to us, ultimately the radiologist remains legally responsible for the outgoing report.

**Q** *What about possible further technical developments?*

One development that could be extremely useful in the future would be the possibility for the software to automatically re-prioritize cases waiting to be signed off by the radiologist. This is currently being worked on but is quite a challenge since robust and powerful interfaces between the AI/PACS and RIS systems have to be established.
Another useful development which is also being worked on is the automatic production of a report by the software or at least the incorporation of the algorithm-derived data into a structured report and the transfer of this report to the RIS.
A relatively minor but still useful development could be the validation of the system to handle supine cases. At the moment the software only handles thorax images taken upright and posterior-anterior. These represent 80% of all the chest X-rays we take, but it would be nice to also cover upright anterior-posterior and supine images as well. I have no doubt that the next release of the software will take care of this.

**Q** *So what's your overall impression of the final resulting software?*

I hope that you have by now detected that I am extremely positive not just about what has been created and our contribution, but particularly how the software performs in daily routine. Right from the start, our strategy was that once we had clearly defined our objectives, we had to focus absolutely on achieving these aims and to "keep it simple" in getting there. We didn't want to get distracted along the way or be side-tracked into exploring too many diversions — tempting though they may be — but which on cold, hard analysis were not strictly necessary for the goal of attaining the objectives and were in fact only "nice to have" features.
Thus our focus on ensuring that the software can reliably detect and at high sensitivity the "Big Five" pathologies in chest X-rays means that we cover 95% of all pathologies that occur in daily routine.
In practice when I get a message from the software rating a case as "inconspicuous", this effectively means that I know the diagnosis can be established in a few seconds. This is a powerful feature allowing the clinician and even non-radiologists to concentrate on other clinical aspects of the patient. You can imagine how useful that is in practice. For example take a case, admitted to the emergency room at 3.00 o'clock in the morning – the patient presenting with fever, shortness of breath and serious weight- loss over the previous few weeks. In such a case, if the AI-Rad companion rates the Chest X-Ray "without finding", you can be sure that there is no consolidation, atelectasis, tumor, pneumothorax or pleural effusions. Thus a simple and rapid procedure can have a big impact on the patient's immediate medical work-up. That's the big goal. The software provides help where it is most needed – right at the side of the clinician and especially during night and weekend-shifts.

**Q** *Future prospects? Will you apply the AI-Rad companion to all your chest Xray imaging systems ?*

We will definitely apply this system to all imaging systems and locations/hospitals in our group, where there has already been a positive reaction to this idea .
There are also AI-Rad Companion modules for other organs such as the prostate. As you have no doubt noticed I am "hooked" on the AI-Rad technology approach so naturally I am eager to try to evaluate other modules everywhere we have high case numbers. This is the great advantage of the Teamplay digital health platform. Once the backbone is installed you can expand the system and scale it exactly the way you need it..

In the longer term, a key issue to be faced in radiology will inevitably be whether ultimately the software can be left on its own to make diagnoses without any human intervention. That is not the case at the moment and right now I am happy that every case is still seen by a radiologist and I think my patients are too.
However the hard fact is that our daily workload is increasing inexorably and there is a constant pressure on costs, so the day will come when we will be very happy to have tools like the AI Rad Companion to take over monotonous tasks and give us radiologists time to communicate more with our colleagues and patients. Such a hand-over is increasingly becoming less difficult to contemplate as we accumulate more data on the use of AI-Rad and our confidence and trust in it continues to grow.
Some radiologists perceive AI as a threat but my sincere opinion is that we in the radiology profession will not be able to master the challenges of the future without it. We cannot keep pace with the increasing technical and medical demands just by increasing the numbers of radiologists – in fact we'll be lucky if we can even just replace the loss of experienced colleagues as they retire. To sum up my thoughts on AI and radiology, I don't think I can express myself better than to simply repeat what I said at the last ECR meeting in the perhaps colorful but I believe accurate expression
***"AI will save radiology's ass …. "***