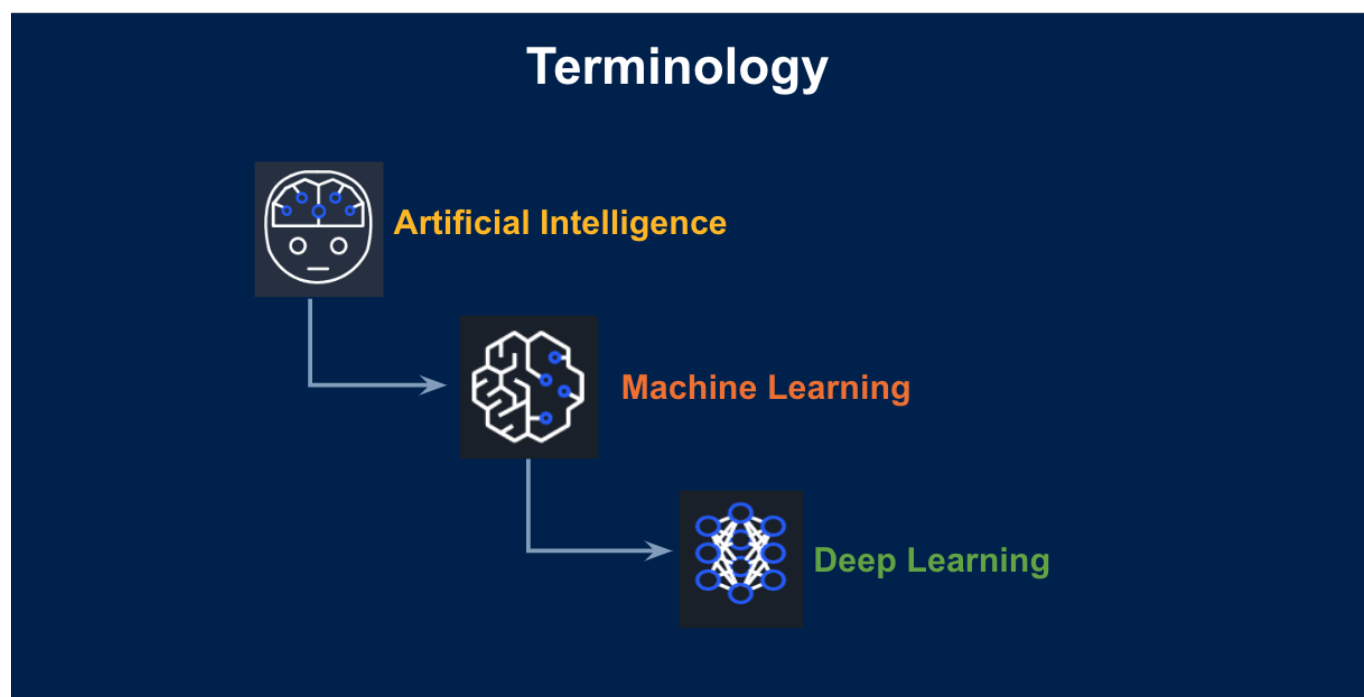# Cancerworld

# AI in Cancer Care: an overview

Adriana Albini / 14 December 2022

The first webinar on AI in Cancer Care was chaired by Eduardo Farina, Radiology Resident at UNIFESP (Federal University of São Paulo, Brazil) and AI Developer for healthcare at DASA (da América SA), SP, Brazil, who welcomed everyone and introduced the speakers.

What is AI in Cancer: Basics and Settings of Application

Felipe Kitamura, Head of Applied Innovation and AI at DASA, and Neuroradiologist at UNIFESP, started his presentation by looking at the three main types of machine learning (ML) algorithms: **supervised** learning, **unsupervised** learning, and **reinforcement** learning. Supervised learning is one of the most used group of techniques that depend on having the answer to the problem. We need to pair the input to a specific output. For instance, if we want to create a model to predict if a patient will respond to a specific treatment, we need a dataset with that information for hundreds, if not thousands, of patients in order to train a supervised model. Unsupervised techniques, on the other hand, will be able to cluster groups with specific patterns from unstructured datasets even in the absence of a desired output. Especially in cancer care, unsupervised learning is quite useful. Reinforcement learning has been used recently in imaging analysis for training agents to do specific tasks. This type is mostly used for teaching computers how to play games, but it can also be employed in healthcare, albeit less so than the first two. All of these three techniques can be applied to medical images, through a subfield of AI called Computer Vision. They can also be used and applied to medical texts using the technology of Natural Language Processing.

Basics

Supervised Learning	Unsupervised Learning	Reinforcement Learning
Computer Vision		
Natural Language Processing		
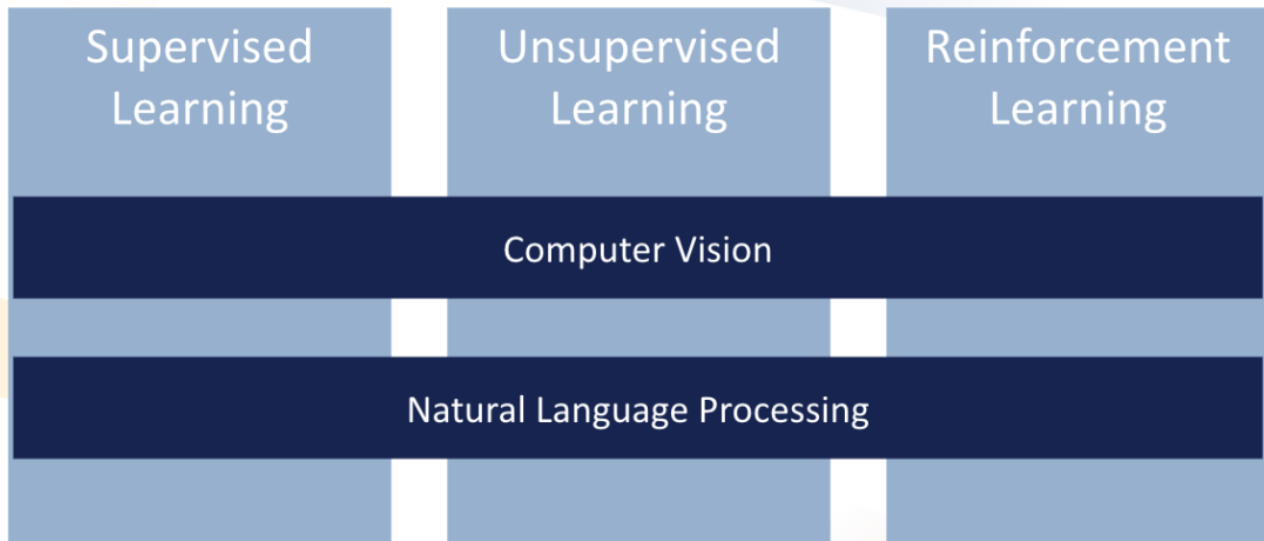
In terms of cancer **risk estimation**, the authors of a paper published in *Radiology* in 2019 developed a supervised deep learning model based on mammography, which showed to be significantly more accurate than the risk model most used today, the Tyrer-Cuzick (IBIS tool). Their model does not use AI to detect if there is cancer in the current mammogram, but to predict if a patient will develop it in the next 1 to 5 years. We know that breast density is correlated with the risk of cancer, but this study found that patients with non-dense breasts who were assessed by the model as high risk really were higher risk than those who had dense breasts but were predicted by AI as lower risk. This is an interesting finding, because it indicates that it is not all about breast density, there are other risk factors present in the mammogram.

Detection and diagnosis: The same team of researchers also published another study, in which supervised deep learning was used to detect the presence or absence of cancer in a current mammogram. It demonstrated that these models work well enough to be able to reduce the workload of a radiologist. They are not as yet widely available and used in practice, but we might gain in efficiency by using them in the future. The Prostate Gleason Score is a very important score in clinical practice, that helps identify patients at a higher risk of severe disease. Radboud and the Karolinska Institute promoted a Kaggle competition, the Prostate cANcer graDe Assessment (PANDA) challenge, where they provided a big dataset of whole-slide images, so that data scientists around the globe could try and develop the best models to predict the Gleason Score. We can also use deep learning to classify images of skin lesions, for instance to predict if they are malignant or not. This is an interesting application because it has the potential of a direct-to-consumer application. It could be used not only by dermatologists but by patients themselves as a screening technology. We are also developing deep learning models to detect polyps during colonoscopy. Another application, which is not related to images, is Natural Language Processing (NLP) to reduce time to treatment. At DASA, prof. Kitamura and his team use NLP models to read medical documents, such as radiology reports, to identify specific diseases that need follow-up, thus reducing waiting time for treatment. For instance, the time to treat breast cancer has come down to 15 days from 60 days, which was the standard of care before this tool was implemented.
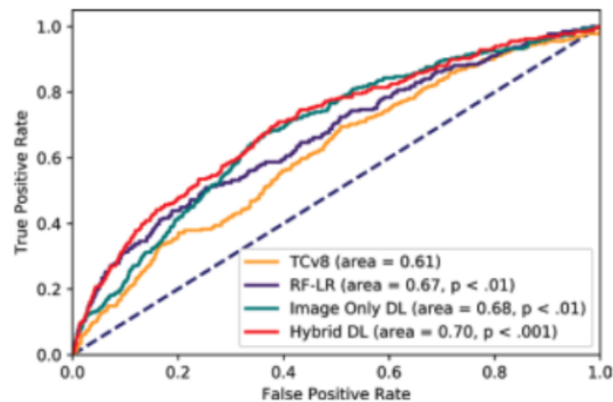
Risk estimation

A Deep Learning Mammography-based Model for Improved Breast Cancer Risk Prediction

Adam Yala, MEng • Constance Lehman, MD, PhD • Tal Schuster, MS • Tally Portnoi, BS • Regina Barzilay, PhD

- Mammography:

"When our hybrid DL model was compared with breast density, we found that patients with nondense breasts and model-assessed high risk had 3.9 times the cancer incidence of patients with dense breasts and model-assessed low risk."

Figure: True Positive Rate vs False Positive Rate ROC curves:
- TCv8 (area = 0.61)
- RF-LR (area = 0.67, $p < .01$)
- Image Only DL (area = 0.68, $p < .01$)
- Hybrid DL (area = 0.70, $p < .001$)

Sharing Progress in Cancer Care

https://pubs.rsna.org/doi/pdf/10.1148/radiol.2019182716

3

Subtype classification: Cancer is a very heterogeneous group of diseases, and we can leverage a specific **unsupervised** technique developed by Laurens van der Maaten and Geoffrey Hinton, called t-SNE (t-distributed Stochastic Neighbor Embedding). As an example, prof. Kitamura took the study "Validation of Whole Genome Methylation Profiling Classifier for Central Nervous System Tumors" by Lucas Santana-Santos's team. The researchers were able to identify different groups of tumours based on their methylation profile. This was also applied for the medulloblastoma subtype classification. Again, t-SNE was applied to clustering different subtypes of group III, and group IV medulloblastomas.

A way we can use AI is to **optimise treatment.** For instance, we can try to **predict responders** to specific therapies. In the article "Molecular determinants of response to PD-L1 blockade across tumor types", Romain Banchereau's group tried to see if they could predict response to PD-L1 blockade. Unfortunately, it did not work the way they had hoped. We always think of AI as something magical that brings new advances to medicine, but it is not always the case. These techniques can fail. They are prone to bias, and they may not be able to generalise correctly. We can also **predict the prognosis**. In a study by Dong Nie, imaging data were used along with demographic and clinical data to predict the survival of patients with brain tumours. Although this is not a ready to use tool yet, it is interesting because it offers a new perspective. Patients often want to know how long they are going to live, and physicians cannot give them a precise answer. Using machine learning models, in the future we will be able to be more precise in defining the prognosis for each patient. Deep learning can be used to **segment CT scans** and help in the planning of radiotherapy. Another interesting application is to obtain **CT scans from MRI**. CT scans are needed to plan radiotherapy, yet sometimes cancer patients have undergone MRI but not CTs. Studies are carried out to use a specific deep learning technique, called Generative Adversarial Networks or GANs, to create a CT image from MRI. It has been shown that the planning of radiotherapy based on this synthetic CT works as well as if the patient had performed a real CT.

Another use of AI is to **create new molecules** or to **find the best combination** of drugs for a specific kind of cancer. As we get access to more drugs to treat cancer each year, the question we may want to ask is, what is the best combination of these drugs for each kind of cancer?

Expectation vs Reality
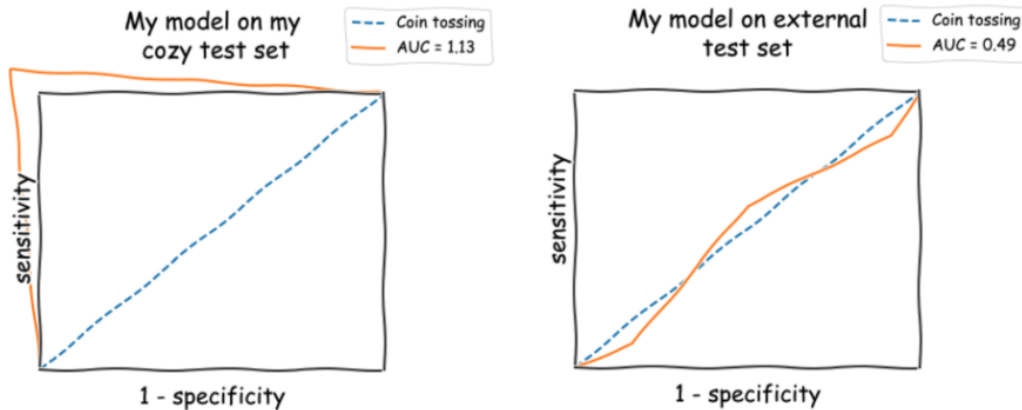
Inspired by Vinicius Ribeiro. Reframed by Felipe Kitamura

There are many different and new ideas on how machine learning can be applied to improve cancer care, but it is worth remembering that AI might not work outside of the data it was trained on.

The State-of-the Art of AI in Cancer Care: Opportunity, Bias, Barriers, and Gaps

Aziz Nazha, Executive Director of Early Clinical Development at Incyte, US, started his presentation with a brief recap of terminology. **Artificial Intelligence** is making machines think and do things like a human without explicitly programming the machine. **Machine Learning** is teaching algorithms with data. As already mentioned by prof. Kitamura, there are two main types of machine learning, supervised learning, where we know the answer, and unsupervised learning where we do not know the answer. A subset of machine learning is **Deep Learning**, which is a mathematical modelling of the neurons in the brain. The neurons in our brain are connected to each other and fire-up when we do certain tasks. The mathematical model for that is called **Deep Neural Network**. In simple terms, in machine learning we have a set of inputs, which could be a set of data, an image, a video, or a combination of these. An output is what we are trying to predict. There are two categories of outputs: **classification** and **regression.** Classification can be binary, for instance images of cats and dogs, or multiclass, when we need to classify three or more instances, such as dog, cat, giraffe. A regression problem is a continuous number. For example, if I'm building a model to predict the house prices in Philadelphia, that would be a continuous number. In between input and output we have a machine learning algorithm or a deep learning algorithm. The problem in healthcare is that everybody tends to focus on the middle, when in order to be successful we must focus first on the output: what I am trying to predict and whether predicting that outcome will be clinically meaningful. The next question should be the input, do I have the right data to answer that question? Is my data biased? And then, the last question should be, do I have the right algorithm for this?

Terminology

Input → Algorithm (Vs) → Output: Classification, Regression
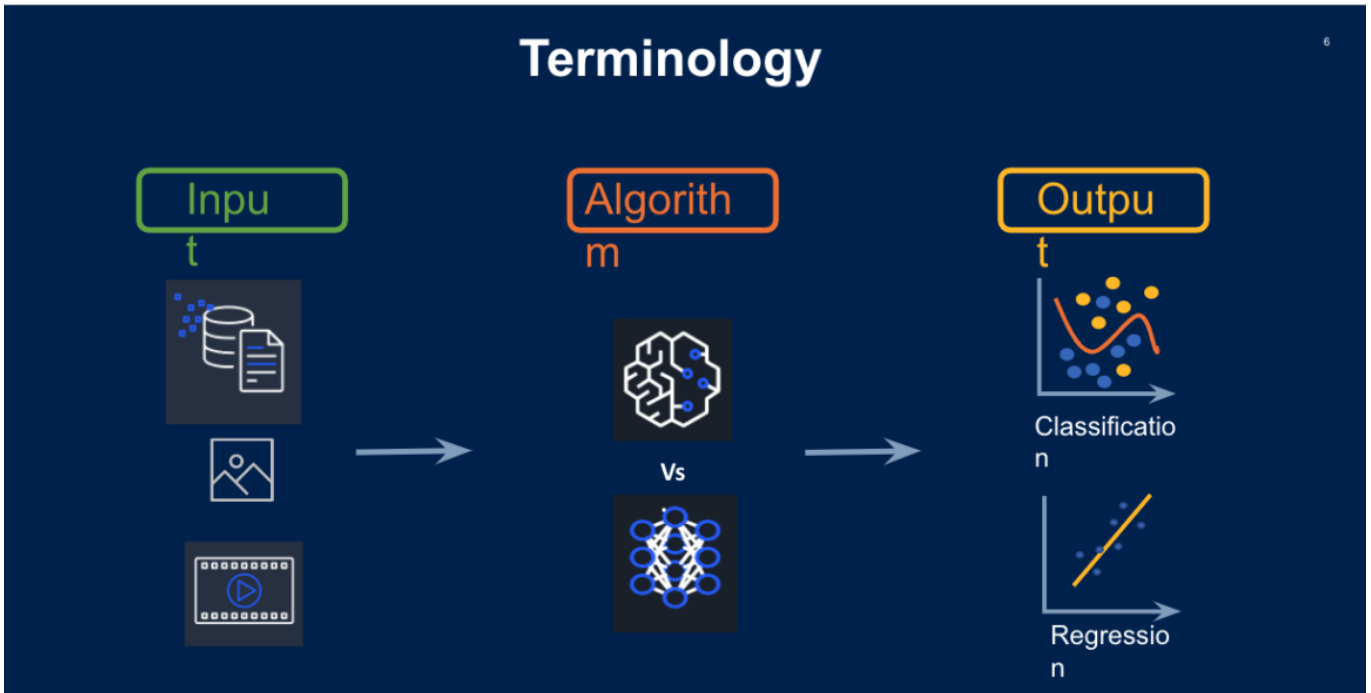
To see how AI can be applied to cancer research, dr. Nazha focussed on a method devised by his team for a blood cancer called myelodysplastic syndromes (MDS). The way it is normally diagnosed is with a bone-marrow biopsy to identify dysplasia. Then the disease gets staged. We take the patient's blood counts, the chromosome and then we have a staging tool that will enable us to define patients with higher or lower risk of progression. We treat patients with aggressive cancers more aggressively compared to lower risk cancers. But this paradigm has limitations. First of all, the diagnosis of MDs is very subjective, and pathologists may be in disagreement. Also, predictions often do not match the actual patient outcomes. And this situation is not limited to MDS, it is true for other cancers as well. Today, we give chemotherapy without knowing which patient is going to respond. We have uncertainty in predicting response or resistance to treatment. So, can we solve some of these problems using AI? Starting from diagnosis, Keiko Sasada's article published in 2018 showed that there is significant discordance in what different pathologists consider dysplastic. In the study conducted by Christian Matek and others, the investigators dissected the cells inside each of the bone-marrow biopsies they had taken, and fed those images to a deep neural network. Then, they asked the neural network to identify dysplastic versus non dysplastic cells. The investigators were able to achieve a robust model that enabled them to improve the accuracy of detection, albeit the model was still struggling to differentiate the abnormal cells, as feeding the model with the abnormal cells is still dependent on a human deciding whether or not they are abnormal.

But do we actually need bone marrow biopsies? In an international study, dr. Nazha and colleagues used clinical and next-generation sequencing data to build a machine learning model to diagnose MDS and differentiate it from other similar myeloid malignancies without relying on bone marrow biopsy data. To build the model, genomic and clinical data were collected from three institutions in the US and Europe. Cohorts and data were externally validated. The researchers then ran the data through a machine learning algorithm. They asked the algorithm what the important features were that impacted its decision, and from that they extracted those features and plotted them on a graph to make sure that the variables were clinically important and that new information could be learnt from the algorithm. Lastly, they built the final model. There are other rare blood cancers that look exactly like MDS and sometimes it is really hard to differentiate. But this machine learning model showed 95% accuracy against the validation cohort. Moreover, this model also explains why a given patient has MDS or another blood cancer. And validating the model against an external cohort is extremely important.
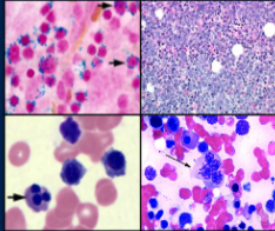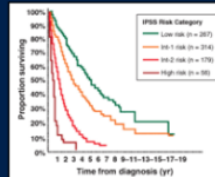
What about predicting response or resistance to therapy? For MDS, there are two FDA approved drugs, Azacitidine and Decitabine. About a third of patients respond to them, and it takes about six months for the response. In other words, we are giving this chemotherapy without knowing which patient is going to respond. Models built on traditional statistics, clinical data, genomic data, were not robust enough to be used clinically. So, dr. Nazha's team used a consumer-type **recommender system** algorithm to try and predict response resistance to chemo. For instance, a company can make a prediction of what movie a consumer is going to watch tonight on the basis of the ones they have watched last week. Swapping consumers with patients and movies with genes, the question becomes, can we predict response or resistance to chemotherapy? By using this recommender system algorithm, the team was able to identify genomic biomarkers that predicted resistance to chemotherapy with 93% accuracy against the external validation cohort. The researchers identified leukaemia cells that have certain genomic biomarkers and treated them with Decitabine and showed that these cells have primary resistance to chemo. They built a platform that uses AI and CRISPR/Cas9 to identify in silico signatures that predict resistance, to then test the results in the lab, and finally apply these models in the clinic.

Among the current major challenges of AI in healthcare is **explainability**. Besides being told an output a physician needs to understand the reason for that output. If a model predicts that a certain patient is going to be readmitted to hospital, the physician needs to know why. Another important aspect of explainability is that sometimes we need to know what the algorithm is using to make a decision, because it may be using wrong data or looking at the wrong detail in an image to make that decision, thus coming up with incorrect outputs. Another problem is **reproducibility**. if I build a model in a healthcare system using my data, can I apply the same model in a healthcare system in the United States or even around the world and reproduce that same result? Very often not, I cannot reproduce the same accuracy, the same model, using a different patient cohort. This is problematic, and it explains why we still do not see many algorithms applied in hospitals, even if they are FDA-cleared. **Bias** is also a serious topic, that is not talked about enough. There are different kinds of bias. Bias could be in the data. If I build a model on only white patients, can I apply it on African American patients? Perhaps not. We can have bias in the algorithms, meaning the output of the model is biased toward minority, males versus females or other prospects. If you start with bias data, you get a biased model. Sometimes, there is also bias in using certain algorithms instead of others that could produce different results. **Fairness** goes along with bias to a degree. When an algorithm

recommends something for a patient, is it a fair recommendation? Are we denying some patients access to life saving treatments? **Regulatory approval** is still problematic. Although the FDA has cleared many algorithms, it still does not mean they have been approved. Another crucial challenge is **implementation in the workflow**. Sometimes we do not have a robust algorithm, but we try to implement it in healthcare anyway, other times the hospital does not have the infrastructure or personnel trained to implement it and monitor it. Lastly there is the so called **model drift**, when models perform worse in time. We need to monitor the model, and if its accuracy starts to drop, we need to retrain it. So, these are the challenges from a technology perspective. **What about patient concerns?** Studies have shown a number of apprehensions patients have about the application of AI in healthcare. They are concerned about the safety of the algorithm; the potential threat to patient choice; a possible increase in healthcare costs caused by the introduction of new technology. And, again, data-source bias. But the main challenge is lack of talent. We need people who understand and speak the language of healthcare, but are also that of AI, we need physicians and researchers who are also computer and data scientists with a deep understanding of both fields. To this end, dr. Nazha and his team have just launched a [course](#) to teach students, residents and fellows how to do machine learning in AI using no-code or low-code. They set up the ambitious goal to train 10.000 healthcare professionals by the end of the year.
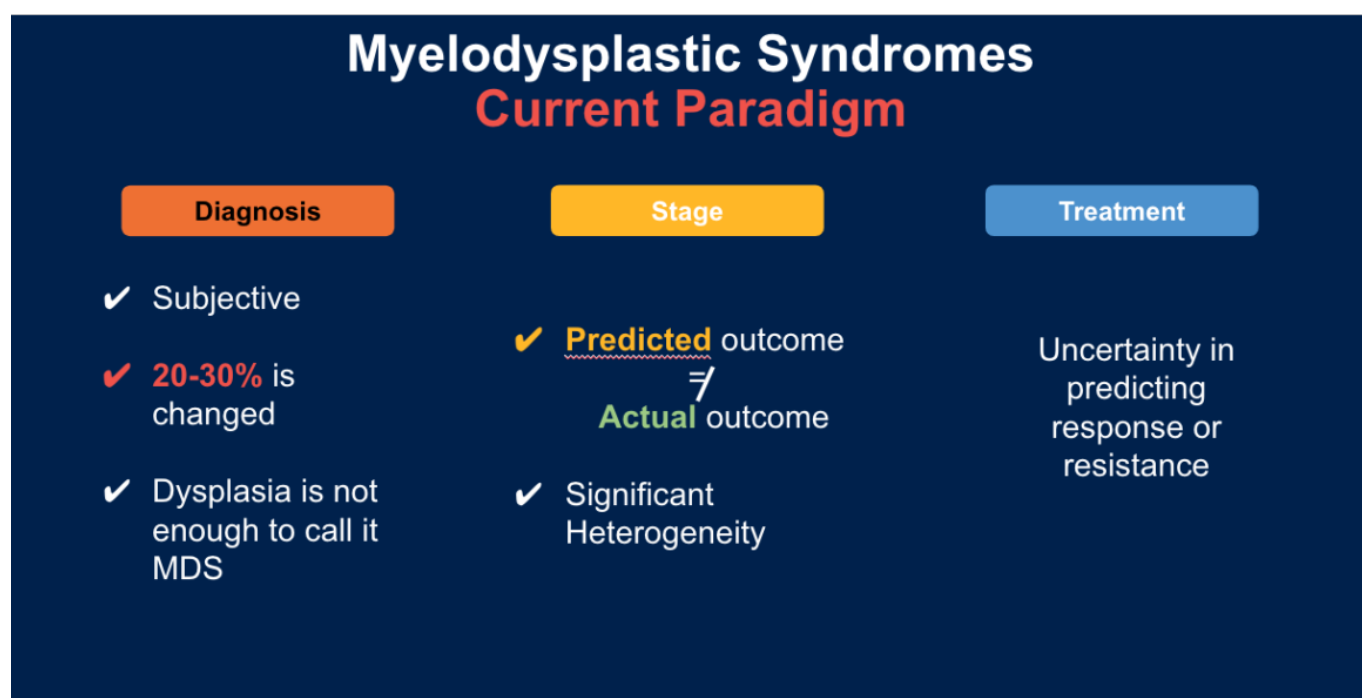
If we want to improve the outcome for our cancer patients, we have to embrace technology, and change how we think about healthcare today. AI is transforming our lives. We all use it every day. There is a huge opportunity for AI in healthcare, and more specifically in cancer. Of course, there are still limitations, but hopefully they will be overcome in the nearest future. And then, one important point to keep in mind about AI and ML in healthcare is that it is a collaboration between human and machine, with the same goal in mind, better outcomes for patients.

Value of AI in Cancer Care: a Systematic Patient-Oriented Innovation

The third speaker was **Fabio Ynoe de Moraes**, Associate Professor at the Department of Oncology Queens University, Ontario, Canada. Currently, we have three major problems in oncology care. One is that although medical knowledge is growing fast, we still need to make oncology more precise and accurate. We constantly hear about precision oncology, personalised medicine, but we are not quite there yet. The second point is that the patient journey is still very challenging. The patient should play a central role in decision-making. We need to expand access to medical knowledge and to

support value-based medicine. Last but not least, healthcare professionals need help. They are mentally and physically overwhelmed, and computers can help them improve their daily activities.

Cancer incidence is on the rise. In 2007, around 13 million people worldwide were living with the disease. By 2030 we are expecting to almost double that number, 22 million people. This is largely due to a growing and ageing population, as well as lifestyle changes. But the numbers can vary dramatically across countries and ages, and access to care is unequal. Healthcare professionals are stressed, overworked, and spend a lot of time dealing with bureaucracy and tasks not related to patient care. We need solutions. We need to **promote patient-centred care**, **good use of data** and **innovation**. Patient-centred care is where the care is organised around the comprehensive needs of a patient rather than just their disease. Innovation does not necessary mean doing something new, just doing it better, more effectively, being able to solve problems in a simple way that is fast and efficient. Computer processing capacity has increased at least 80-fold in the past 20 years. We know that hospitals are an important source of big data, if we can access it. However, the human brain is not capable of dealing with the amount of information that we are generating. It can only consider five facts per decision, while now we are dealing with thousands and thousands. In 2022 computers are an integral part of our healthcare system. AI is on the rise, yet it is still far from being generalised clinical practice. We are now generating larger amounts of data in oncology, but they need to be better standardised. We need regulatory frameworks and solutions to tackle diversity, equity and so on. We need to conquer data and build a culture of innovation so we can implement AI in practice. We will be facing a data tsunami, we cannot stop it, but at least we can prepare for it. To apply AI to the real world, we need to create and cultivate data sources, harvest information, and feed valuable knowledge. We also need to do this fast, in an agile cycle.

Farming is a useful analogy for the process of creating outcome databases in healthcare. First, we need to cultivate data sources, by standardising inputs and processes. When we begin to get a good amount of data with high-volume and variability, we need to harvest that information, so as to make sure we can work with and understand these data. Thirdly, we are going to feed knowledge that improves patient care. Making sure that we have the right process to implement AI in practice takes time. But if we do it well, we can harvest quite a good number of projects and data. It is all about first things first, and the first thing that we must start from is a standardised nomenclature in documenting, communicating, contouring, planning and reporting. We need to improve how we document things on the electronic health system, on imaging, et cetera. And standard data entry must be quick and easy. We now have multiple sources of documents that can help us with minimal datasets, nomenclature standards, etc. To bring AI into the real world, we need to have good practices. And in order to have those we must remember a couple of things. First of all, we must ensure that the data we are working with is relevant to the practice, that it is acquired in a consistent, relevant, and generalisable way; that it aligns with the intended research question; that there is appropriate separation between training, tuning, and testing datasets; and that there is an appropriate level of transparency in the output algorithm. We do not want a black box that does not let us know where the data comes from, how to assess it, and if it makes sense. To avoid that, one approach is to define a clinical problem and develop an algorithm that comes from real world clinical data, test it in local studies, go to clinical trial, and then proceed to clinical registration and finally clinical deployment. The journey of developing something that will be applied to clinical practice is a long and difficult one. We need to focus on important problems and work together to extract, select and refine data, and then, we will be able to test it, compare it and use it together with humans. At that point, potentially we will need to go to prospective clinical validation. Another systematic approach that has been used is to have platform processes embedded in the electronic health system. So, we can collect standardised data, pre-process them and then start using them to help predict outcomes. DISCOVERY AI, for instance, houses various modular AI tools to process data in order to make predictions. Modules can be diagnosis, readmission, complications, and so on.

The ER at Queens University now runs an algorithm that can predict with a high degree of accuracy the number of patients who will visit the Emergency Department within the next 48 hours. That way, they can plan how many physicians, nurses and other healthcare professionals are needed to be on call. It is important to have this kind of set-up in place so we can leverage all the data that the hospital, the ER, or the cancer centre is generating in a way that will in future have an impact on the running of our clinic. Another way to systematically approach the problem is to build a big database, a so-called digital knowledge database. These databases in cancer care are very important because they will allow us to collect patient demographic, personal and clinical data, and past clinical decisions and outcomes, and make best decisions on management. After we have these data, we will need to harmonise and preprocess the data so that they will be analysed and worked with a machine learning process. Then, we will potentially be able to synthesise the results, and create algorithms that will help physicians in the clinic. All of these systematic approaches take time and commitment, and they all start with standardising the way that we collect and work with data.

Looking at some **applications of AI in oncology**, an international collaboration that prof. de Moraes worked on is a good example of a model that does not require particularly complicated algorithms. Regression models can do very important work when based on good data. The team worked with about 20,000 patients, 1 million unique data points, 55 centres, two external validation datasets, and they were able to create nine unique groups for staging prostate cancer, taking into account age, stage T category, N category, Gleason Score and so on. This system is AJCC-compliant for clinical prognostic staging. With this kind of work, we are making cancer care more precise and more personalised. Another example, this time on how to bring AI to planning in radiation oncology, is a prospective study led by Chris Mcintosh in Toronto, where they deployed and evaluated a machine learning algorithm for therapeutic curative-intent radiation therapy (RT) treatment planning for prostate cancer. ML and human-generated RT treatment plans were compared in a retrospective simulation and treating physicians assessed those plans in a blinded manner. 89% of ML-generated RT plans were considered clinically acceptable. RT planning with ML also reduced the time required for the entire process by 60%. This would mean freeing up space for humans to do other kinds of work or to engage better with their patients. Some of the areas in which we can apply AI to radiation oncology in the real world, if we have established a framework are: help in treatment decision, imaging, treatment planning, RT delivery, prediction of response to treatment. A recent study by colleagues in China shows that AI for segmentation of nasal cancer can decrease the time of contouring from about 38 minutes to 8 minutes and can also reduce more than two-folds inter-

observer variability. Follow up is something else that we are working on to make sure that we catch recurrence or toxicity before it even arises.

Limitations and challenges in moving from paper studies to the real world are, first of all, data acquisition and standardisation. Often, we do not have enough data because it is siloed within institutional centres. Other issues are privacy and security, competition between institutions and lack of data-sharing infrastructure. We have now guidelines proposed to support FAIR (findable, accessible, interoperable, reusable) data use. However, as mentioned by dr. Aziz, we have lack of human resources. We need to train people and free up their time so they can apply this kind of analysis in their centres. Healthcare is under pressure. Patient-centred care, data and innovation are the driving forces for improving healthcare. Advancement in oncology and AI presents opportunities for a major clinical impact, but we need to focus on generating evidence, supporting places in need and teaching the next generation.

What is New in AI in Cancer Care: Progress and Regulatory Perspectives

The final speaker was **Nishith Khandwala**, Co-Founder of Bunkerhill Health, Palo Alto, US. The journey from research to clinical practice for an AI algorithm in healthcare starts with building a proof-of-concept algorithm, one that we have perhaps trained and tested on our local dataset. After that, in order to assess how the algorithm performs on data from other scanner manufacturers, other patient population, etc., we need to validate it on external data from multiple hospitals. Once verified that the algorithm performs consistently well, we must prove its clinical or financial utility. Typically, this is done via a clinical trial, which should demonstrate how the algorithm improves the current standard of care. With that data, we can file for those regulatory approvals that are applicable and necessary. Once we have regulatory clearance, we can distribute the algorithm for clinical use and for prospective use by physicians. This presentation focussed on the regulatory stage, notably in the US. For the sake of this talk, dr. Khandwala made a couple of assumptions, that the type of algorithm falls under the FDA's Class II (medium) risk level (if the algorithm requires a physician supervision or if a physician has the ability to override its result, it likely falls under Class II.) Another assumption is that the FDA understands how to evaluate the algorithm that we are submitting, as they have previously validated a similar one, and understand the positives and negatives of the validation study that we are submitting. If these assumptions hold true, the FDA has an accelerated pathway called the **510k process**, whereby the algorithm developer demonstrates that their algorithm is similar to one the FDA has evaluated and approved in the past.

The FDA has developed tighter timelines to ensure that innovation does not get delayed and reaches the market as soon as possible. So, when we submit our device to the FDA under the 510k process it should take roughly 90 days for them to contact us and tell us whether they want to reject the application, put it on hold or need additional information. At this point we can interact with the FDA to further the application. And the outcomes can be clearance, or the FDA determines that this is not a device that they are familiar with in terms of how to validate it, and in that case, we need to go down a much lengthier process in order to get clearance. So, what does a 510k application docket look like? What pieces of information need to go in it? The first is **indications for use**. The FDA does not just clear or approve an algorithm, but an entire product, which involves how we expect it to be used. An algorithm that is used under physician supervision and an algorithm used autonomously, are both ways in which a singular algorithm can be used. And the pathway or regulatory evidence required to support the two is vastly different. The next part is the **substantial equivalence** discussion. Since we are claiming that the FDA has already approved or cleared a similar device to the one we are submitting, we need to show equivalence between the two. Then the FDA also requires a section on software. What is the software composed of? How was it created? Under what sort of software engineering paradigms? If this is an AI algorithm, as would be the case

in this discussion, what training dataset was used, what validation dataset, what type of algorithm? Was it a convolutional neural network, a transformer algorithm? Finally, perhaps the longest section and the most statistically inclined one is the **testing** itself. Based on the sophistication of our device, and on what the device that we are claiming substantial equivalence to has done, we might be asked to show evidence on both bench testing and clinical front.

As a case study, dr. Khandwala took a company called Optellum. They have a device that looks at incidental pulmonary nodules on chest CTs, especially the ones that have indeterminate risk categories, and suggests how to manage patients with those types of nodules. The tool takes in a CT scan, then the physician, usually a radiologist, tells the algorithm which nodules to focus on. The algorithm then predicts a risk score, and this leads to potentially optimal care for those patients. Typing "Optellum" on the online FDA 510k clearances database brings up a record containing information about the company and details of the algorithm. Clicking on "Summary" we can find in the indications for use that the algorithm characterises incidentally detected pulmonary nodules. It also tells us who the users are: pulmonologists and radiologists only. It also specifies that the input to the algorithm is a CT image. What is the output? It is a single value proprietary score. And then, finally, it also talks about the inclusion-exclusion criteria. We are only looking at patients who are 35-years or above. In short, the indications for use talk about how an algorithm becomes a product. Looking at the substantial equivalent section, Optellum's algorithm was compared against another algorithm from a company called QuantX that looks at mammograms, detects and assigns a risk score to lesions on mammograms. From a user point of view, the two devices are quite different, but from the FDA's perspective of risk-benefit analysis, they are actually fairly similar. Moving on to the **software** section, first there is a diagram showing how the Optellum's algorithm fits into the overall workflow, from the browser clients to PACS, to a server inside the hospital datacentre and finally to the Optellum cloud where the algorithm is hosted and orchestrated. The results come back to the server inside the hospital network, and then to the physician who ordered the scan. The software part also talks about the algorithm itself. There are sections on the convolutional neural network, the training dataset, and the output. What kind of testing did Optellum perform? They did two types of analyses for their algorithm. The first is a standalone **testing**, to show how the algorithm performs on a standalone basis on a static dataset against their goal standard. The second is to show how the algorithm improves the performance of the clinician. They looked at readers without the use of the algorithm and then at the same readers when using the algorithm. The AUC went from almost 82% to almost 89%, so a substantial improvement.

The FDA process for most algorithms is fairly clear. There are steps to expedite the process and researchers should keep them in mind when developing their algorithms and designing their validation studies, so that those algorithms can smoothly progress from research into clinical practice.