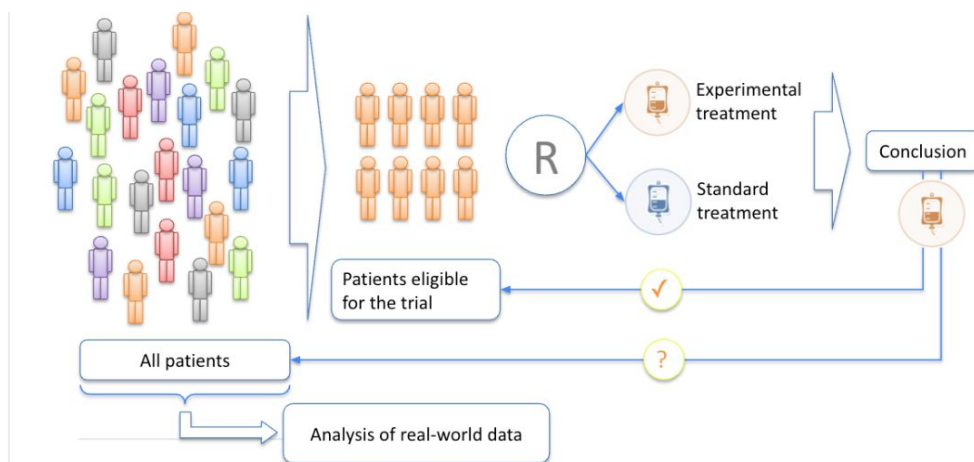


Cancerworld

AI in Genomics and Reporting for Clinical Practice

Adriana Albini / 26 January 2023

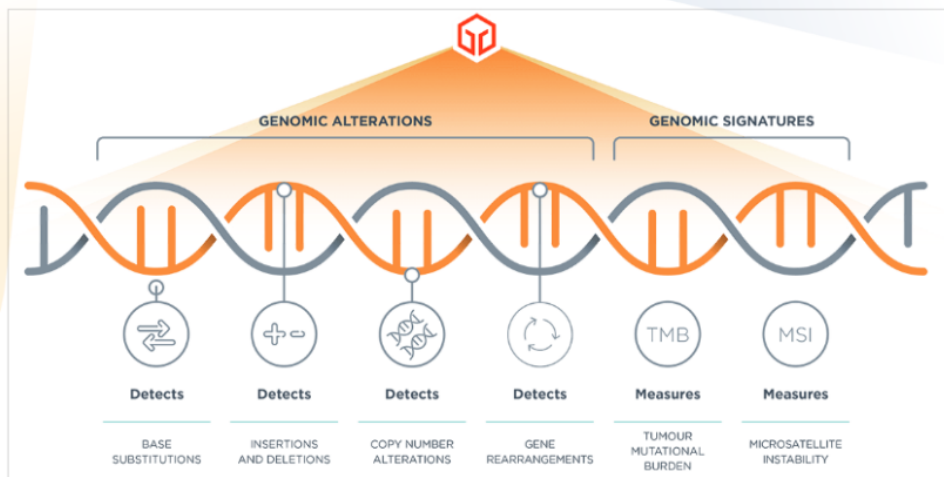


Artificial intelligence is revolutionising many aspects of our lives. Although it entered clinical practice only in recent years, its contribution to clinical research and patient data handling has a growing impact in the development of precision medicine. A panel of digital science and oncology experts discussed the advances and challenges of AI within this context in the third webinar of the SPCC Artificial Intelligence in Cancer Care 2022 series, which took place on 15 December 2022.

Using AI to improve biomarker calling in NGS data

The first speaker was **Tim Chen**, Scientist at Foundation Medicine, Boston, US. **Comprehensive Genomic Profiling (CGP)** allows for the simultaneous detection of all classes of genomic alterations in hundreds of genes with a single sample. This new approach to leveraging **next-generation sequencing (NGS)** data was pioneered by Foundation Medicine (FMI). NGS is a technology for massively parallel sequencing of DNA and RNA. There are different ways to utilise NGS. **Single Marker Tests** detect one or a few specific types of alterations within a single gene or DNA region of interest; **Multi-Gene "Hotspot" NGS** tests identify prespecified mutations occurring in limited areas of genes of interest; while the **Hybrid Capture Based CGP Platform** can analyse more than 300 genes simultaneously, saving time and preserving tissue. It detects all four main types of DNA alterations, including single-nucleotide substitutions, smaller insertions and deletions, rearrangements, and copy number variations. Since CGP profiles a broad set of genomic regions, the assay is also able to detect complex genomic signatures that predict immune-oncology (IO) response, such as tumour mutational burden (TMB) or microsatellite instability (MSI).

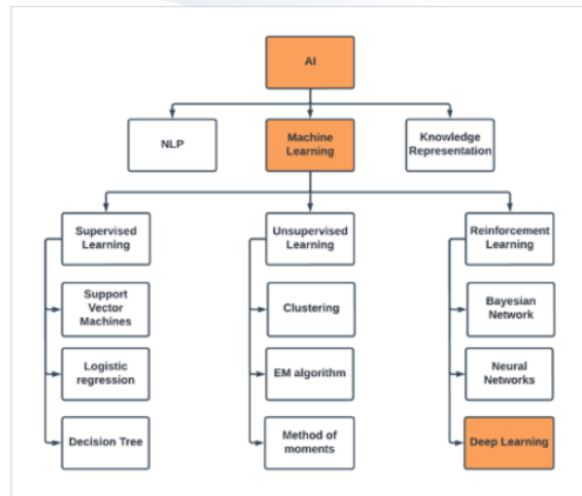
CGP can detect complex biomarker (eg MSI, TMB)



AI is any technique that enables a machine to imitate human intelligence. There are different branches of AI, such as NLP (Natural Language Processing), Machine Learning (ML), and Knowledge Representation. Machine learning is a subset of AI that enables machines to learn from experience or data. There are three types of ML: supervised, unsupervised, and reinforced. Deep learning is a subset of ML based on neural networks that let a machine train itself to perform a task. AI, of course, is not a new concept. The terms “artificial intelligence” and “machine learning” were coined as far back as the 1950s, but what made them more popular in recent times and allowed their extraordinary growth is the higher-performing computational power we have now, available at low cost. We also have large high-quality annotated datasets. However, AI-based application in healthcare is just getting started. The first medical device that used AI was approved by the FDA only in 2018 (to detect diabetic retinopathy in adults).

What is AI?

- **AI:** Any technique that enables machines to imitate human intelligence
- **ML:** A subset of AI that enables machines for learning from experience
- **Deep learning:** A subset of ML based on neural networks that let a machine to train itself to perform a task

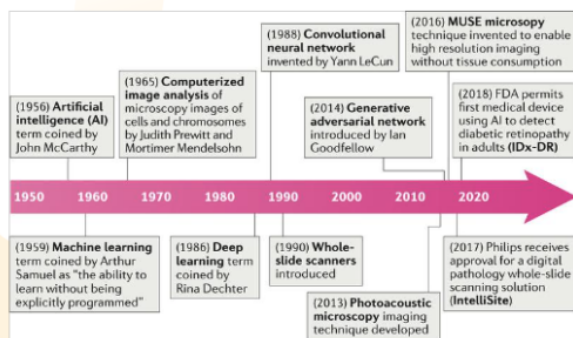


non-exclusive list

Big data and AI are changing the way in which biomarkers are discovered. In the traditional way, we start with R&D, define a question and biomarker, and then collect data (normally a small sample size). Assuming the biomarker succeeds in stratification, the biomarker is moved to Investigational Use Only (IUO) and can be used in clinical trials. When we leverage big data in machine learning, the process is different. Here too we start with defining the question after some R&D; then, we perform **feature generation**. FMI has sequenced over 600,000 patient samples, so we can generate features from them and use an appropriate AI algorithm to define the biomarker. Because the biomarker was defined on a larger dataset, the biomarker should be more robust when moved to the IUO stage. In certain cases, it is even possible to develop a biomarker based on big data and AI without clinical outcome data. Dr. Chen showed two real-world examples from work done at FMI on how we can utilise AI for biomarker discovery. The first one was a scar-based detection of **homologous recombination deficiency (HRD)**; the second was **variant classification without clinical annotation**. Why do we want to build a biomarker to detect HRD? Patients with HRD have a deficiency in the homologous recombination repair (HRR) pathway, which is responsible for repairing double strand breaks (DSB). There are studies showing that patients with HRD are more likely to benefit from a PARP inhibitor (PARPi) in multiple tumour types, including breast, ovarian, prostate, and pancreatic. Looking at data from the ARIEL3 trial, ovarian cancer patients have a better outcome when treated with a PARPi. However, a consistent way to call HRD has not been established, as it can be caused in various ways, by genomic alterations (e.g., in BRCA1/2) or by epigenetic means (e.g., BRCA1 promoter methylation). We know that patients with HRD accumulate genomic scars across their genome. We also know that HRD has a certain pattern in the copy number (CN) profile, so at FMI they used the supervised method to train a machine learning model. For supervised learning, we need the label and the feature set. For the label, they identified a subset of samples that almost always harbour an HRD phenotype, the biallelic BRCA1/2 mutation, to inform the ML of what HRD scarring looks like. For the input, a broad set of CN features was incorporated to allow the ML process to identify the nuanced phenotype of HRD. The copy number profile was transformed into numeric metric, for example, number of breakpoints per 10MB, distribution of copy number, distribution of segment size, and so on. Once they had the label and the training data, they could train the algorithm. They trained a subset of the 600,000 patients on the pan-tumour dataset, doing the standard training, testing, splitting, and cross-validation in a trained dataset to avoid overfitting. Comparing the new machine learning-based biomarker with two already built scar-based

HRD biomarkers - gLOH (genome-wide loss of heterozygosity) and GIS (genomic instability score) - the ML one outperformed the others. Looking into the biallelic BRCA sample to find out what percentage of patients are HRD biomarker-positive, we see high sensitivity in the biallelic BRCA sample mutated cohort overall, especially in ovarian cancer, where around 90% of patients with these alterations were HRD-positive. This scar-based biomarker detected HRD beyond BRCA mutations. Looking into the BRCA wild-type cohort, they showed that HRD-positivity is associated with biallelic alteration in some well-known HRR gene, like *PALB2*, *BARD1* and *RAD51C/D*. The model was trained on samples without clinical data, but still associated with the clinical outcomes. Using a clinico-genomic database containing the patient survival outcome, treatment history, and genomic set, ovarian and prostate patients treated with a PARPi whose tumours were HRD marker-positive had the best outcome. Likewise, metastatic pancreatic cancer patients treated with platinum-based chemotherapy, FOLFIRINOX, who had the HRD-positive signal, displayed increased survival benefit.

The AI-based application in healthcare is just getting started.

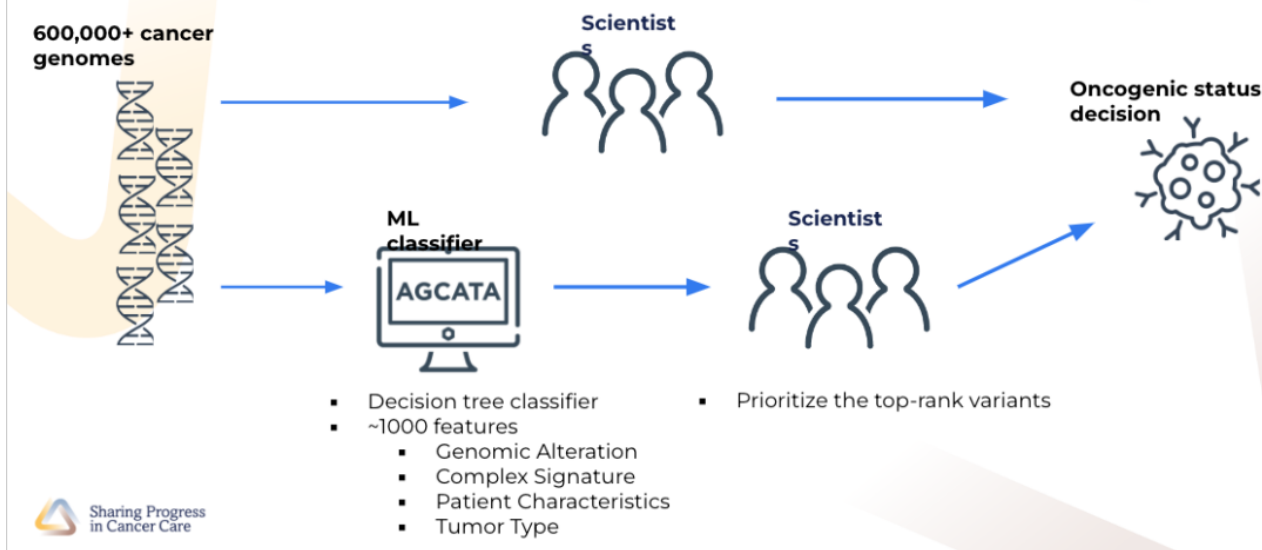


- The concept of AI is not breaking new
- Extraordinary Growth
 - High performance computing
 - Large annotated dataset

The second example shown by Dr. Chen was the **ML-based oncogenic variant classifier**. Oncogenic variants are those mutations that have potential to cause cancer. There are many mutations observed from over 600,000 patient samples and our goal is to report them correctly in order to learn more about cancer genomics and improve drug discovery. Most of the drivers are well-studied, but a lot of rare variants are not. How can you more accurately classify these? We need to leverage a lot of human resources. A group of scientists will do the literature research and data analysis to decide the oncogenic status. As a concrete example, Dr. Chen looked at a specific rare EGFR (epidermal growth factor receptor) mutation. It was observed only one time in all medical literature, but in the FMI database it was detected seven times. The feature is consistent with oncogenicity in terms of disease bias (lung), low TMB, domain hotspot, and mutual exclusivity with drivers. The idea now is to build a machine learning classifier incorporating relevant genomic features, including co-occurring alterations, complex signatures, patient characteristics, and tumour type, to train the model. This can flag likely oncogenic alterations, and then the scientists can prioritise the top-rank variants, in order to decide the oncogenic status. Big data can facilitate cancer research and improve patient care. In certain cases, big data and AI can allow for biomarker development without clinical outcomes.

Oncogenic Variant classifier

- ML classifier can flag likely oncogenic alterations for additional review

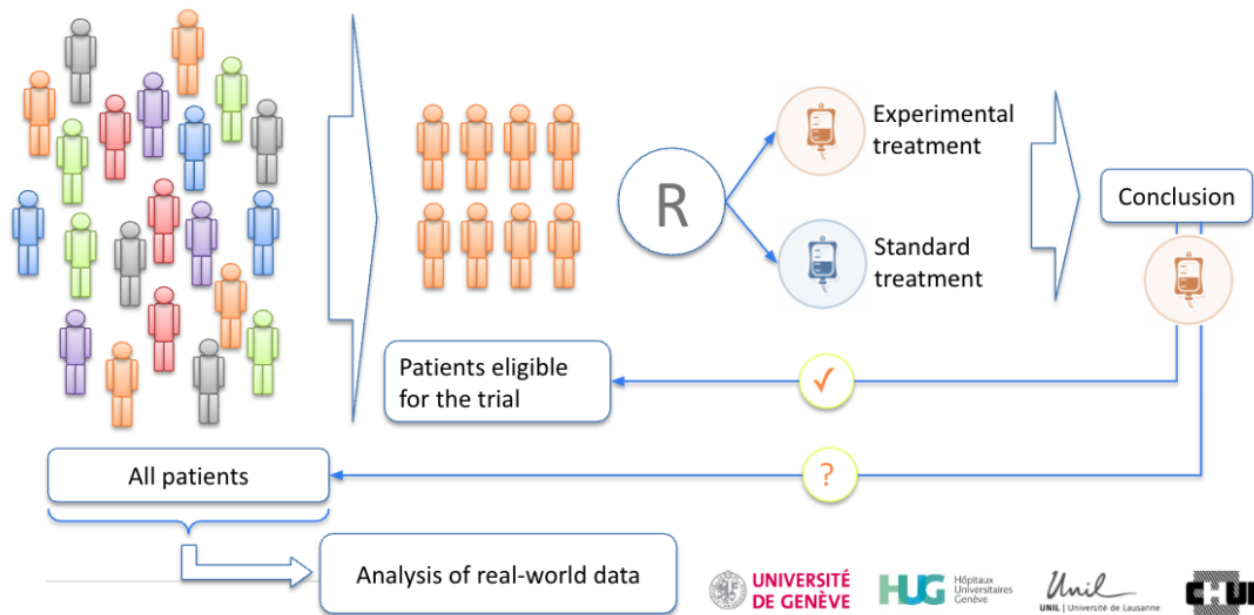


Using AI to mine clinical data

The second speaker, who also chaired this session was **Olivier Michielin**, Head of Oncology Department and Head of Precision Oncology at HUG, Genève; Head of the Precision Oncology Centre at CHUV, Lausanne, Switzerland.

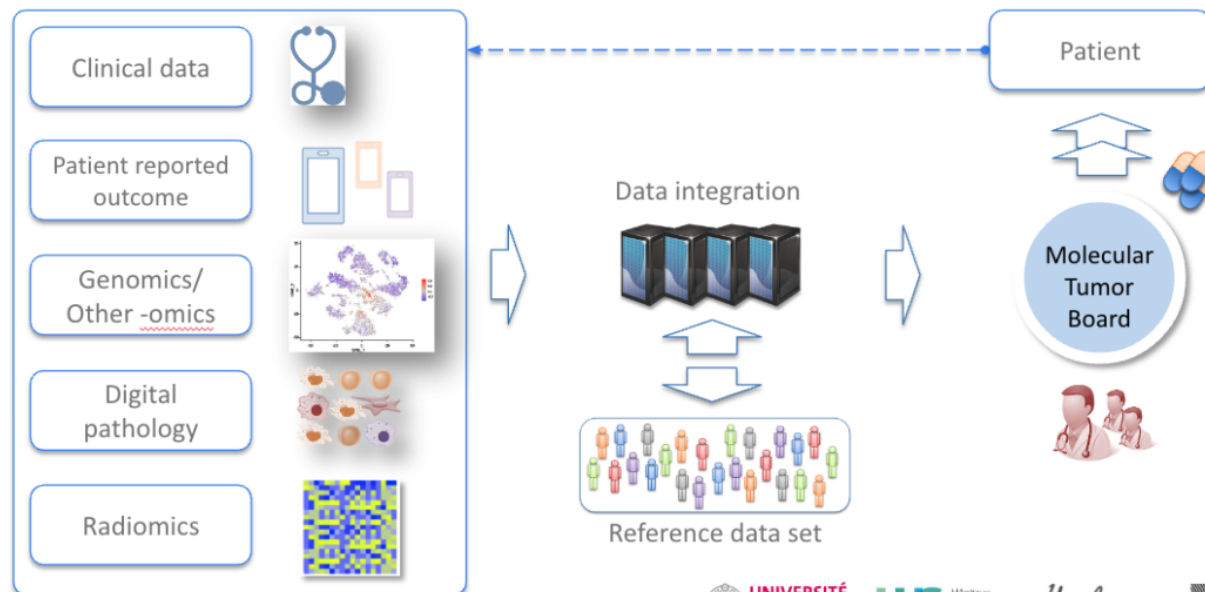
How can we use predictive biomarkers, digital markers, and AI to help us select a patient for therapy? In cancer, we often encounter the situation in which only a small fraction of patients is deriving benefit from a treatment. But to get this benefit, we expose the entire population. It is therefore important to identify those patients, isolate them and restrict the treatment to this population only. And of course, try to find a new biomarker for the rest of the population, so that the correct patient receives the correct treatment. The idea behind predictive biomarkers is to restrict exposure to a drug to only those patients who are likely to benefit from it. In order to do this, we need solid biomarkers, which are usually based on a deep interrogation of the tumour and its microenvironment; in other words, we need to mine a huge dataset. And in order to do this, we need AI to work through the data and associate them with clinical outcomes. This is also a very interesting way to utilise real-world data. Knowledge in medicine is usually developed through prospective clinical trials, that require us to reduce the variability in our sample. There are inclusion and exclusion criteria to define the trial population. Within that context, we are able to determine whether an experimental treatment is better than the standard one. At this point, we should bring the conclusion we have arrived at back to the population of the trial. Unfortunately, a lot of the conclusions are actually brought back to the general population, which is unselected. We have to do that because our patients are suffering, and we need to treat them. So, there is a gap between what we know from the trials and what we use in practice. Here, real-world data could be very useful, and AI would complement the traditional way of generating know-how through prospective trials.

Complementarity between clinical trials and real-world data:



What kind of data do we need for precision oncology? There are several sources: clinical data, patient-reported outcomes (PROs), all the genomics (RNA, protein level data, and other omics); digital pathology and radiomics. Images are becoming increasingly interesting because of the deep learning and neural networks revolution, that allows us to tackle very complex datasets. The big question is, how do we integrate all of these data? We need a reference dataset to be able to understand what is normal or not in the sample of the patient we want to analyse. Then we bring all this information to a molecular tumour board composed of data scientists, clinicians, pathologists. So, we have a rich environment in which to reach a decision. And, to complete the loop, we need to return to the dataset. When we treat a patient with a drug that has been determined through this process, we want to make sure that the data for that patient is sent back to the dataset. So, we keep on adding information and, of course, need a learning system that allows us to progress at each step. To do this at a national scale in Switzerland, The SPO (Swiss Personalized Oncology) project was launched, co-led by Prof. Olivier Michielin and Prof. Mohamed Bentires-Al. The aim was to connect all the hospitals through their data warehouses. During the first 3 years, a norm was defined for how to encode clinical, molecular and images data, to make them exchangeable between the different hospitals in the country. Each hospital now adapts the cancer patient information flow to generate the Minimal Data Set (MDS) automatically. This can then be encoded into the RDF file format, which allows mutual interoperability. Not only can we use the structural data, but we can also use tools to query, for instance, the textual information. Prof. Michielin showed an example of mining radiology reports. His team used Natural Language Processing (NLP) to see whether it is possible to predict if a patient will be a responder or a non-responder to a treatment, by reading the report from the radiologist directly into the NLP system.

Precision oncology: integrating multiple data streams



UNIVERSITÉ DE GENÈVE

HUG Hôpitaux Universitaires Genève

Unil Université de Lausanne

CHUV

They have tested this on 100,000 reports and are beginning to have very good accuracy, in the region of 80-90%. They have now employed this strategy to start analysing real clinical data. A comparison was drawn between the results of a well-known metastatic melanoma trial and similar data queried using NLP processing capacity to evaluate the CT scans. The overall survival and progression-free survival statistics obtained were very similar to those from the trial. So, the real-world data appeared to be fairly similar to a prospective phase III trial. This is very interesting, because we have many more data-points for those patients, since they were treated by us, than what we would get from a regular clinical trial. In addition, we can perform any trial we believe might answer a question, even trials that cannot exist, say, for ethical reasons. We could, for instance, compare immunotherapy with targeted therapy, immunotherapy in combination with chemotherapy, and so on, even if these comparisons have never been done before. The idea now is to use a sort of big expert AI system to understand all these different trials and start to build a predictive system that can tell us what the best strategy would be for a given patient at a given time-point in their oncological history. An example of the type of discovery possible using these strategies: a retrospective analysis from the CheckMate 069 trial suggested that the usage of proton pump inhibitors (PPI) in melanoma patients treated with ICI could interfere with immunotherapy efficacy. To test this, scientists at CHUV performed a survival comparison of melanoma patients treated with anti-PD-1 and anti-CTLA-4 and receiving or not receiving Proton Pump inhibitors, using a Virtual Trial. They observed a clear difference between the PPI+ and PPI- cohorts, with a worse prognosis for patients with PPI medication. There seems to be an impact in taking these anti-acid pills. The mechanism could make sense as we know that acidity in the gut might change the microbiota, and the microbiota is directly linked to the responsiveness of the immune system. A larger cohort and a multivariate analysis would be needed to confirm this. Nonetheless, it is conceptually interesting to see that something we discover in real-world data can then be confirmed in prospective phase III trials.

Radiomics: application to precision oncology

- Radiomics is a quantitative approach to medical imaging that aims at
 - capturing (patho)physiological characteristics via quantitative image analysis in order to extract features to be associated with prognosis or therapeutic response in a specific clinical scenario.

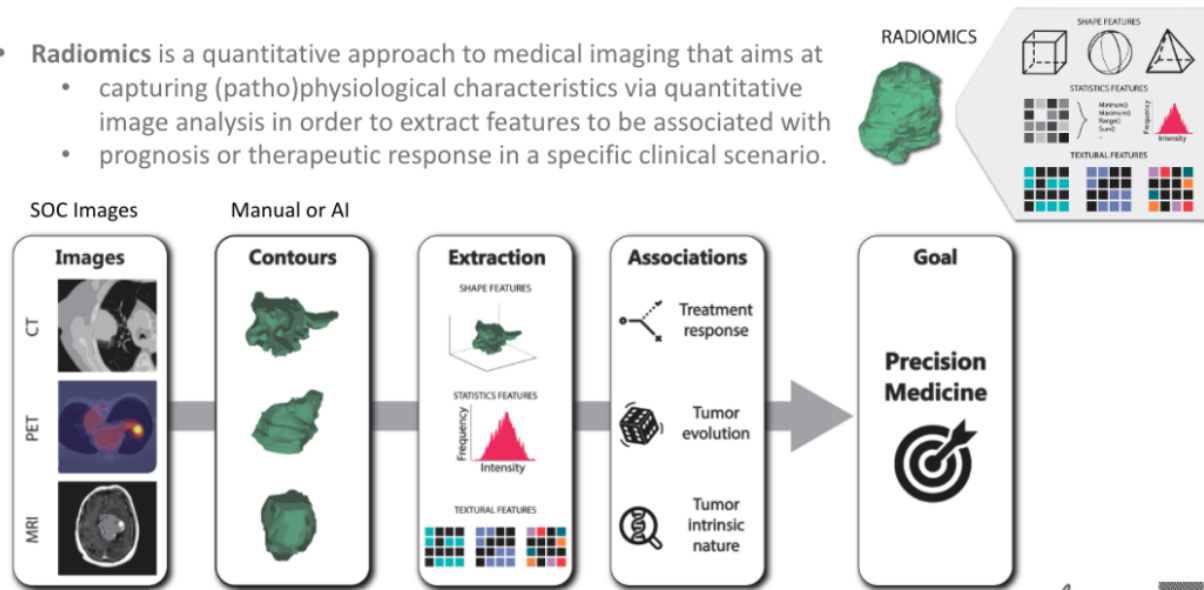
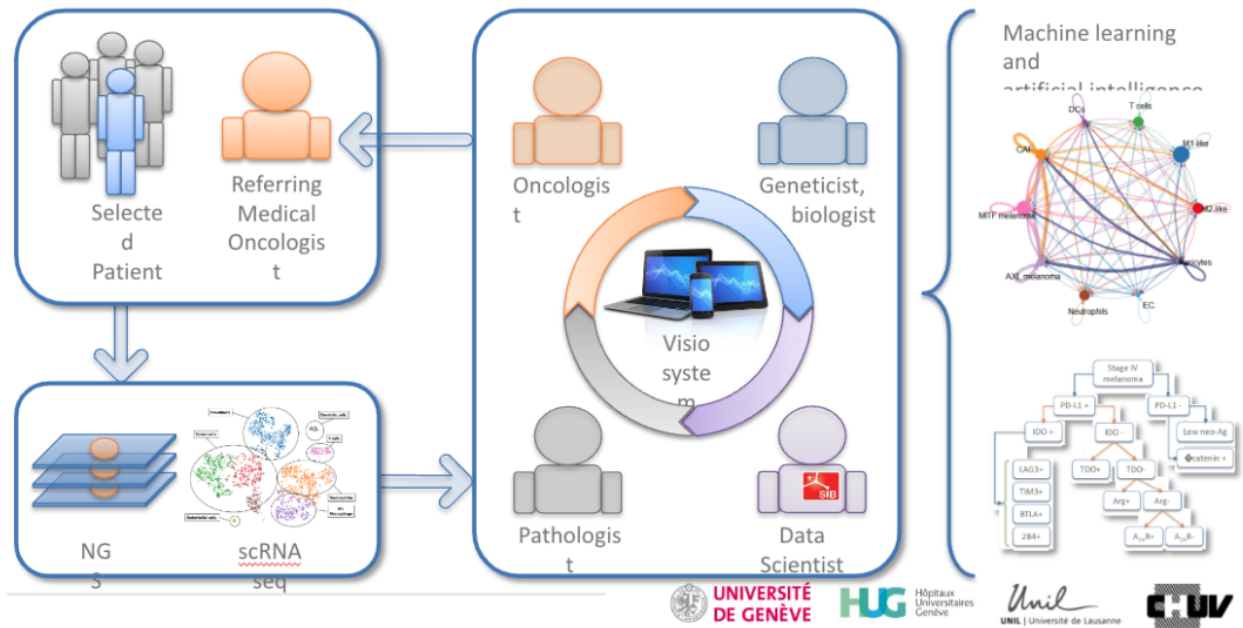


Figure from Coroller, T. P. (2017). Combining data science and medical imaging: Advancing cancer precision medicine with radiomics. *DataWysse / Universitaire Pers Maastricht*. <https://doi.org/10.26481/dis.20171214tc>

What can we do with additional digital data? Radiomics is an interesting example. When we follow cancer patients, we usually do imaging every three months, CT or PET-CT scans. We measure the lesions and then we call it a progression if the diameter increases more than a certain amount, or a response if there is a shrinking of the disease. But there is so much more we could do with this highly dimensional information-rich dataset. We can use these new AI tools to extract the tumour. We do segmentation from PET-CT, CTs, or MRI using AI algorithms. Once we have the 3D isolation of the tumour, we can use the whole toolbox of data science to analyse texture, entropy, all the aspects of the tumour signal, then try to extract features from these datasets and associate them to any clinical parameter that we judge useful – precision medicine being the final goal. As an illustration, Prof. Michielin showed a marker that his team is trying to develop on PET-CTs. The algorithm was able to detect a feature that seems to separate patients who are going to respond to immunotherapy, that is, progression-free survival, and those who will not. This feature is associated with the heterogeneity, the entropy of the PET signal on that lesion, which could make sense, because then we are measuring the heterogeneity of the tumour, which might be a negative predictive factor for response, as the more diverse the tumour is, the more likely there will be an escape mechanism in its behaviour. Prof. Michielin’s group is now trying to confirm that this signal is also observed in an orthogonal dataset they are building with colleagues from another university. So, this was an example on how images can be used to build a predictive programme. All the information gathered is then brought to a molecular tumour-board made of oncologists, pathologists, and data scientists where we can do the omics for the patient, use AI, and then refer a proposition to the medical oncologist for the next line of therapy. This programme started at HUG and CHUV five years ago, and now they have treated more than 2500 patients, which is a very high volume, and the algorithms continue to be trained on it. Prof. Michielin concluded with an example of a clinical outcome, where the trajectory of a patient’s sarcoma could be changed thanks to the discovery of a PD-L1 amplification. They treated the patient with a PD-1 monoclonal antibody, and were able to put him into deep remission, whereas on chemotherapy he was still progressing. This was just an example on how these rare findings may drive important changes in the disease course of some patients. Of course, we need to look at all the statistics and learn from this, but hopefully, step-by-step, we will increase the performance of all these new strategies.

Molecular tumor-board and AI

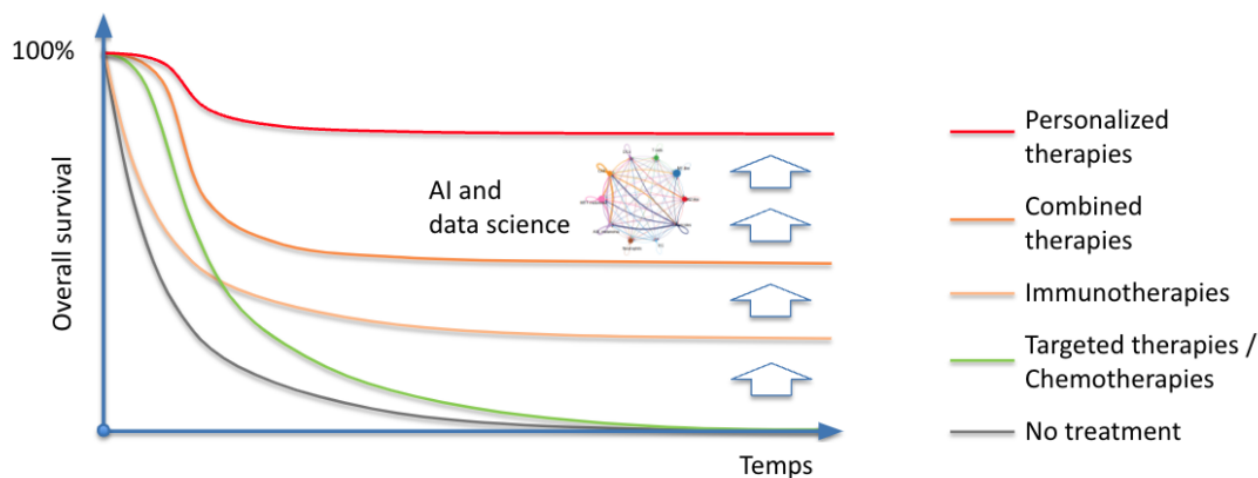


Integrating Genomic Data and Clinical Data in Practice with AI tools

The final presentation was by **Dimitris Kalogeropoulos**, Independent Global Health Innovation Adviser and Expert World Bank, the World Health Organization, the European Commission, and UNICEF's Digital Health Centre of Excellence; Health Executive in Residence at the UCL Global Business School for Health, London. Out of the 5 Ps of medicine, omics are personalisation and biomarkers precision, and patient-reported outcomes are a very good way to lend structure to real-world data. If we find a way to link those together through telehealth or virtual care, then we have the 5 Ps together with Predict, Prevent, and Participate.

Looking at the clinical side of AI rather than the research side, two important institutions with large care research samples and surveys point out that the clinical side of artificial intelligence is still problematic. Stanford Institute for Human-Centered AI says that the carbon footprint of AI is an issue, we are using a lot of computing power to produce relatively limited results. And reasoning is still a frontier for artificial intelligence, which is a very interesting observation. The Council of Europe Steering Committee for Human Rights in Biomedicine & Health says that AI deployment in care remains nascent; clinical efficacy demonstrations are lacking in relation to research; and performance generalisation from trials to clinical practice is still largely unproven. This is fairly recent data, as the reports were published last summer. Why is this happening, and what can we do about it? Before attempting to answer this question, we must remember that there are many types of artificial intelligence, and there are primary versus secondary uses. Primary is research, and secondary is clinical practice. Clinical AI is possibly under-performing because knowledge is lacking. AI increasingly covers the entire spectrum of translational medicine and learning healthcare, from research (rAI) to clinical trials (ctAI), to clinical practice and remote patient management. In ctAI, AI helps in many different ways to locate patients, create cohorts, and so on. Clinical AI (cAI) is a term used to refer to the artificial intelligence which directly benefits clinical practice, including prediction and education, for example real-world 3D models for pre-surgical preparation, nowadays very common in surgery. If we could collect all these types of artificial intelligence in one space, we would call that translational AI. And this is envisaged too as a future for artificial intelligence, where connected and collaborating artificial intelligence agents (collective AI) deliver learning healthcare in real time.

Expected benefit from personalized strategies



UNIVERSITÉ DE GENÈVE

HUG Hôpitaux Universitaires Genève

Unil Université de Lausanne

CHUV

What are the primary challenges in clinical AI? CAI is an integral part of clinical practice; hence it shares with it the same data misfortunes, including the risk of over and under-diagnosis and treatment. This is a circular policy effect, because there are numerous studies showing that if we introduce AI with bias into clinical practice, we run the risk of downstreaming the over-utilisation of healthcare services. This is also true with cancer, but to a lesser extent, because cancer treatment personalisation is a very important part of oncology. Then, we have sensitivity to changes in the clinical environment. If we transfer an algorithm to a new clinical environment, with new data, it will be liable to performance decay in efficacy and effectiveness; we have sensitivity to quality of data, (its sources), thus liability to ethics compromises such as explainability, reproducibility and accountability. There is also the scale predicament: we need to have cohorts that are wide enough, and which cover primary, secondary, and tertiary follow-up care. We need to define bias. We have limits for sensitivity and specificity, and many limits for clinical endpoints in clinical trials, but we have not paid enough attention to how we define bias along with acceptable limits. We must decide how many biomarkers and how many reference points we need to validate our results. A study was conducted by Johns Hopkins University School of Medicine with other institutions in the US on the clinical robustness of digital health companies. The score was based on clinical trials applications registered on ClinicalTrials.Gov and FDA filings. Only 9% of the companies and start-ups that were sampled had a robustness score of 8 to 10+. 70% of those companies had 0 to 3 robustness scores. Clearly, we have an issue with the introduction into clinical practice of artificial intelligence beyond the research environment.

Based on experience from the industry, Dr. Kalogeropoulos suggests using omics as an AI strategy. If we combine omics with biomarkers and AI, we can make an extremely powerful translational precision medicine tool which can be used in a circular pattern, feeding three different validation processes. One of them is framing systems medicine knowledge (SMK), as in phenotyping. The other one is discovering SMK, so nuancing endotypes, specialising the phenotypes, the omics with biomarkers, and finally using SMK to conduct digital phenotyping. An example of digital phenotyping is when a smartphone detects changes in the patient's voice, and from that it flags a virus, or other health issues. Translational Omics Apps enable the collection of reliable and auto-validated data with which cAI can be deployed effectively. Omics and biomarkers offer structure to current clinical data practices, leading to systems medicine. Data origin validation is enabled

through the use of smart mobile devices, telehealth, mHealth, or digital phenotyping apps, because they can collect patient reported outcome measures, etc. Omics as an AI deployment strategy coupled with smart devices provide a new pathway to radically different structures in delivery models, reduces physician and nurse workload, improves outreach, engagement, and prevention at scale, all while collecting structured data.

AI is 100% aligned with biomarker uses as laid out by the Institute of Medicine (now National Academy of Medicine): discovery, early product development, surrogate endpoints for claim and product approvals, clinical practice, clinical practice guidelines, comparative efficacy and safety, and finally, public health practice. If we were to fit omics and biomarkers in those uses and build a digital phenotyping machine, a connected intelligence that actually implements agents in those fields, then, we could have translational AI. Natural language programming AI tools are used for biomarker discovery. A scoping review published recently identified 78 articles reporting the extraction of data frames for cancer diagnosis, description, procedures, treatment, and management. These articles had more or less the same data frames in terms of biomarker discovery. PROMs are a great longitudinal care validator. Real-world data are collected through a variety of sources and tools as part of routine care or virtual care as digitally enabled continuity add-ons. They give RWD a structure. A mobile phone that patients use, for example, in survivorship support, to register their current symptoms, etc., provides a validated temporal/longitudinal clinical context that we can then superimpose on our omics and biomarkers. Ideally PROMs turn qualitative symptoms into numerical scores, making them actionable for triage decisions. PROMs provided by patients using digital sign-on devices also give a validated snapshot in time. PROMs enable telehealth as a new P5 medicine paradigm: for patient outreach, connectivity, and engagement. There is a huge market which is building on PROs and smart devices, and also a number of new biomarker technologies that are connected to AI, for example, multi-omics phenotyping, digital pathology, digital twins, including P3 medicine (Predictive, Preventive, Participatory), P3 risk assessments, and P5 treatments. We have omics endotyping with biomarker discovery and survivorship support. This type of application for PROMs in cancer is gaining a lot of traction and provides outcome improvement beyond the clinical outcome. There are many applications out there in all of these three groups. The idea is that we somehow bring them together through sharing knowledge and data. For example, genome sequencing provides a wealth of genomic mutation pattern-trajectory data. The challenge is to link these with clinical phenotypes and endotypes for clinical interpretation of omics and biomarkers. Biomarker endotyping depends on published studies. 200,000 articles related to biomarker endotyping came out in 2019 alone. So, this cannot be done manually, and AI needs to be employed for data mining. Multimodal ML AI can be used for complex diseases such as cancer. We need to bring together omics, biomarkers, and clinical data.

The omics paradigm can be extended with NLP for P5 treatment. A study was conducted in 2019 on temporal phenotyping by mining healthcare data to derive lines of therapy for cancer. We can discover different lines of therapy and their impact on disease progression and therapy interventions. Processing speed notwithstanding, we can do this in real-time, and this could affect our decision with regard to treatment, based on morphological tumour staging, biochemical, physiological, and behavioural data that AI can discover. Provided, of course, that the phenotypes and biomarkers are ready. Digital pathology can be used, for example, for feature extraction and detection of cancer in WSI based on tumour-infiltrating lymphocyte maps. A study showed that this feature extraction with AI discovered prognostic capacity for 13 different cancer types, including breast, lung, and colorectal. Digital twins for P5 oncology are also increasingly used in assisting decisions for treatment options versus expected disease trajectories. For survivorship, AI offers state-of-the-art PROM-driven P5 oncology; symptom management and monitoring; and integrated follow-up care, to provide improved outcomes. Of course, this has to be subsequently clinically validated.

In this complex context of translational AI in the use of phenotyping, biomarkers, and PROs, what are the **secondary challenges**? First of all, we have regulatory complexity, gaps, and inconsistencies. There is the issue of regulatory policies because we need to move AI away from the medical device or product liability, which is currently pursued primarily in Europe, to the actual medical liability frameworks. This is a direction the US is currently investigating. We have issues with regulatory practice incorporating RWE. This is going to be a big challenge for future development in clinical and research AI integration. We must improve delivery vehicles, paradigms, and policies. We must target social innovation endpoints. When we deploy clinical AI, we want to change the delivery channels for healthcare, change the models with which a patient is managed in the health system. We do not want to just change the clinical endpoints. We need to support a transition to systems medicine. Germany, for example, is now considering a huge investment programme to support systems medicine nationally. And finally, data recycling and circular economies in health and care are something we should seriously consider.

What are the opportunities? We can recombine clinical correspondence reasoning with clinical coherence reasoning, the old way of doing things with the new way, where it is effective and where it is necessary. We can enable translational policymaking - policy research and industry in 360° views of sustainable health innovation. Currently, policymaking is not translational. Policy is based on old evidence collection methods and old evidence quality expectations. We can deliver data recycling, using health and care data as assets. And this is also very important, because at the current rate, we are simply consuming too many resources to produce relatively simple AI experiments, which eventually do not scale because of how expensive they are in terms of data requirements. And finally, we can enable the transition to net-zero health systems with AI in service, for example, of quaternary prevention.