EAU leads innovative data initiatives

Provides support to evidence-based policy making

Prof. Dr. Monique Roobol
Dept. of Urology, Erasmus MC Rotterdam (NL)

m.roobol@erasmusmc.nl

In the past decade, there has been an explosion of healthcare-related data. With the digitalization of medical records, increasing affordability of molecular testing, advent of medical informatics and widespread use of wearables, the sheer volume of data available for analysis is staggering [1].

Big data in healthcare
"Big data" refers to large and complex datasets generated by a wide range of sources. These datasets are typically characterised by their large volume, high velocity, and extensive variety. They can be difficult to store, process, and analyse using traditional data management and analysis tools.

Big data can be structured, unstructured, or semi-structured. Structured data refers to data that is organised in a specific format such as tables, spreadsheets, and databases. Examples of structured data include clinical and financial data. Unstructured data refers to data that does not have a specific format such as text, images, and videos. Semi-structured data is a form of data that has some kind of structure but it is not as rigid as structured data.

There is a discussion that "big" is no longer the correct parameter, but rather how "smart" the data are, focusing on the insights that the volume of data can reasonably provide. This aspect is fundamental in the health sector. The potential of big data in improving health is enormous. However, its potential value is unlocked only when leveraged to drive decision making and enable such evidence-based decision making, it is necessary to have efficient processes to analyse and turn high volumes of data into meaningful insights [2].

Due to the complexity and diversity of the data, as well as, the computational power and storage required to handle it, the analysis of big data often requires specialised software, infrastructure, and expertise. How can we bridge the gap between the collected data, and our understanding and knowledge of human health? This is covered by "data science".

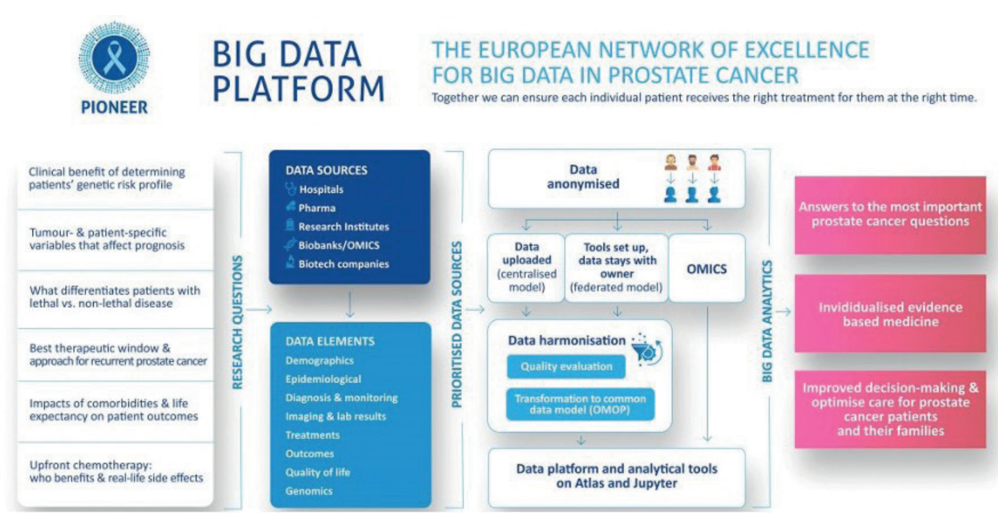
Data science has been defined by three distinct forms of analysis tasks: description, prediction, and counterfactual prediction [3]. Descriptives are useful for exploring and finding patterns in the data which may lead to testable hypotheses. Prediction analysis can improve diagnostics and prognostics, while counterfactual prediction constructs models to address flaws inherent to observational data for inferring causality [4].

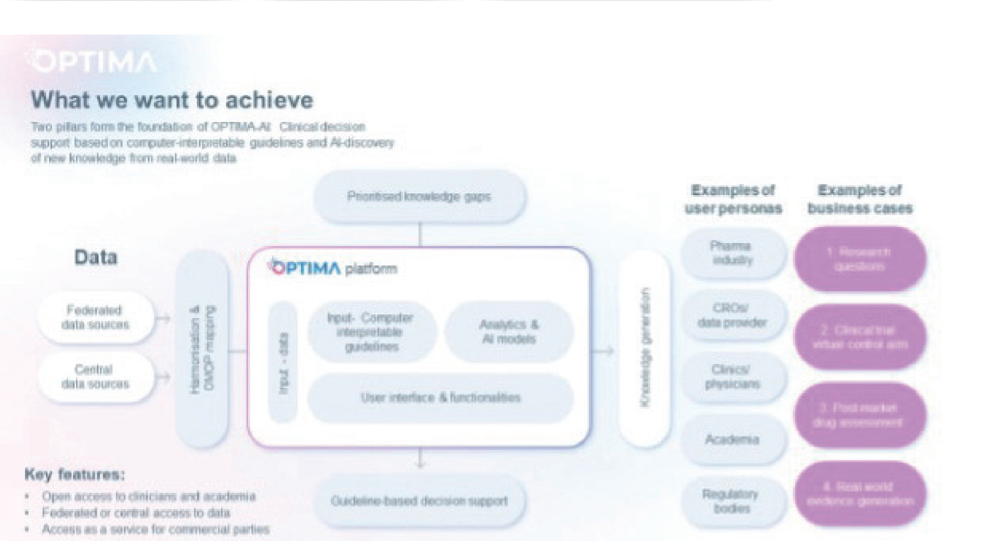
In summary, big data refers to large and complex datasets generated by a wide range of sources. Although big data can provide valuable insights, it is often difficult to process and analyse using traditional methods.

Big data and artificial intelligence
Since the capacity of the human brain to process information is limited, there is an urgent need to develop and implement alternative strategies to process big data. In addition to the increased availability of data, the augmentation of storage and computing power has boosted the development of data-processing techniques such as machine learning (ML) and artificial intelligence (AI), which are becoming increasingly important tools to tackle complex issues e.g. cancer care. A growing body of studies highlight AI as an emerging tool to help personalise cancer-care strategies by analysing available data. A recent study identified 97 registered clinical trials for AI in cancer diagnosis, most of them started after 2017 [5,6].

Big data and AI are both powerful tools that are being used to advance cancer research. A few examples are:

- Genomic data: The Cancer Genome Atlas (TCGA) and the International Cancer Genome Consortium (ICGC) have generated large amounts of genomic data that can be used to identify genetic mutations associated with different types of cancer.
- Drug discovery: AI can be used to analyse large amounts of data from preclinical and clinical studies to identify potential new therapies and inform drug development. For example, machine learning algorithms can be used to analyse data from electronic health records, genomic data, and clinical trials to identify new drug candidates and drug interactions.
- Radiology and pathology: Big data and AI can be used to analyse medical images such as CT (computerized tomography) scans, MRI (magnetic resonance imaging) scans, and pathology slides to identify patterns and biomarkers associated with cancer.
- Clinical decision support: AI can also be used to analyse large amounts of clinical data, such as electronic health records, to identify patterns and predict outcomes. For example, machine learning algorithms can be used to predict which patients are at high risk of progression, or to predict which patients are likely to respond well to a certain treatment.





????

Although these technologies are promising, more research is needed to fully realise their potential in cancer research. There is also a need for more data and big data infrastructure to handle and analyse the large amount of data. Additionally, ML models are only as good as the data they are trained on; data diversity and data governance are necessary in ensuring that the models are unbiased and generalizable to different patient populations.

Big data and policy makers
Big data analyses have the potential to provide valuable insights that can inform and guide policy decisions, but it is not a given that policy makers will automatically accept these insights. The use of big data in policy making is still a relatively new field. For it to be adopted more widely, there are a number of challenges that need to be addressed.

One challenge is the quality of the data. Policy makers need to have confidence that the data being used is accurate, reliable, and unbiased. This can be difficult to achieve, particularly with large and complex datasets, and requires a significant investment in data cleaning, integration, and quality control. Another challenge is the lack of understanding of the data science techniques used to analyse big data, which can make it difficult for them to assess the credibility and usefulness of the results. Lastly, the ethical and legal issues associated with the use of big data are still being resolved. This could be an obstacle for policy makers, who might be concerned about issues such as data privacy and data security.

Despite these challenges, there is considerable progress in healthcare. The use of big data analytics has been used to improve patient outcomes, reduce healthcare costs, and improve the efficiency of healthcare delivery systems.

Big data projects in urology
With the vision of establishing a connected data infrastructure and an ecosystem for big data analysis, the EAU already started aggregating data regarding urological conditions (primarily urology cancers) in 2016. This is essential as urology faces many barriers to potential improvements in better patient outcomes.

The Innovative Health Initiative (IHI) is a public-private partnership between the European Union and the European life science industries. IHI's goals are to translate health research and innovation into tangible benefits for patients and society to ensure that Europe remains at the cutting edge of interdisciplinary, sustainable, patient-centric health research.

IHI funded the projects PIONEER and OPTIMA. PIONEER focuses on predictions regarding particular outcomes concerning particular interventions. OPTIMA broadens the spectrum of prostate, breast and lung cancer research to explore commonalities across cancer patients.

PIONEER, the European network of excellence for big data in prostate cancer (PCa), is led by the EAU and Bayer (figure 1). PIONEER consists of 38 partners across nine countries and aims to unlock the potential of big data and big data analytics. Currently, 95 PCa data sources have been identified and 23 datasets are ready for analyses. A total of 56 research questions have been identified among all relevant stake holders (i.e. clinicians, industry and patients). Data have been analysed, also in the form of a so-called study-a-thon, and 11 manuscripts have been published (see https://prostate-pioneer.eu/outcomes/newsletters) [7]. PIONEER will continue to collect PCa data and address knowledge gaps in care to improve PCa-related outcomes, health system efficiency and quality of health and social care across Europe. Please visit https://prostate-pioneer.eu/ or contact pioneer.info@uroweb.org for more information.

OPTIMA, tackling cancer through real world data and AI, is led by the EAU and Pfizer. OPTIMA aims to implement a vast federated and centralized network of European data providers to help answer the highest priority research questions in prostate, breast, and lung cancer.

OPTIMA brings together 38 partners from across 13 countries. It consists of private and public stakeholders in the clinical, academic, patient, regulatory, data sciences, legal and ethical and pharmaceutical fields. A special focus is on research questions where the existing evidence underpinning clinical practice guidelines is weak or lacking. In parallel, the consortium will develop an application in clinical settings of comprehensive dynamic computer-interpretable guidelines (potentially refined with AI) to better support shared decision making. Please visit https://www.optima-oncology.eu/ or contact communication@optima-oncology.eu for more information.

EAU UroEvidenceHub
The challenge in medicine is that clinical trials – the mainstay of clinical guidelines – are often conducted in small groups; information on long-term (adverse) effects and safety are missing, and key subgroups are frequently not represented. These imply that evidence gaps are there and cannot be filled fast enough with high-quality randomised clinical trials. Therefore, the EAU has started an ambitious new data innovation programme, UroEvidenceHub, which is parallel to and following up the current data innovation initiatives PIONEER and OPTIMA.

To kickstart this programme, the EAU is developing a new state-of-the-art technology platform called the Data Haven. The Data Haven aims to become the largest urology database by collecting large volumes of Real World Data and generating reliable Real World Evidence. The Data Haven will be interoperable with current/future data hubs and other relevant urology data initiatives. Aside from data providers and users, stakeholders from academia, data science and IT companies, patient organisations, regulatory bodies and industry will be invited as well to become involved leading to a tight-knit ecosystem of excellence.

The EAU UroEvidenceHub will elevate the current urological guidelines and will offer a new way of interaction between clinicians and patients, to make shared treatment decisions on a truly personalised basis in the most optimal way possible. Please contact datainnovation@uroweb.org for more information.

Time is against us if we do not progress. Urology has always been at the forefront of innovation, embracing new technology and data. Together, we are able to shift in how medicine will be done in the future and how care will be provided. However, a new approach to knowledge exchange within the medical community is required to succeed.

With more effort to improve data collection, sharing and quality, data governance, education and training of policy makers, we are likely to see big data analyses increasingly accepted and used in healthcare policy making in the future. These data innovations will help urologists improve and individualize the EAU Guidelines and ultimately will provide more personalized care for all urology patients.

The list of references are available upon request.

Sunday 12 March 14:30 - 14:40
The road to evidence-based European policy on early detection of prostate cancer
Yellow Area, eURO Auditorium 1