Advancing Cancer Treatment by Strategically Integrating Multimodal Data: Challenges and Opportunities

Prof. Haridimos Kondylakis
Angelina Kouroubali
Dimitrios G. Katehakis

Hellenic Digital Health Cluster (HDHC)
Institute of Computer Science, Foundation for Research and Technology Hellas (FORTH-ICS)
Department of Computer Science, University of Crete (UOC)

Advancing Cancer Treatment a Top Priority in Europe

Cancer is one of the main priorities of the European Commission in the health domain.

In 2020, 2.7 million people in the European Union were diagnosed with cancer.

1.3 million people lost their lives to it, including over 2,000 young people.

The overall economic impact of cancer in Europe is €100 billion annually.

Evidence shows that 40% of cancers are preventable.

Unmet needs

01.
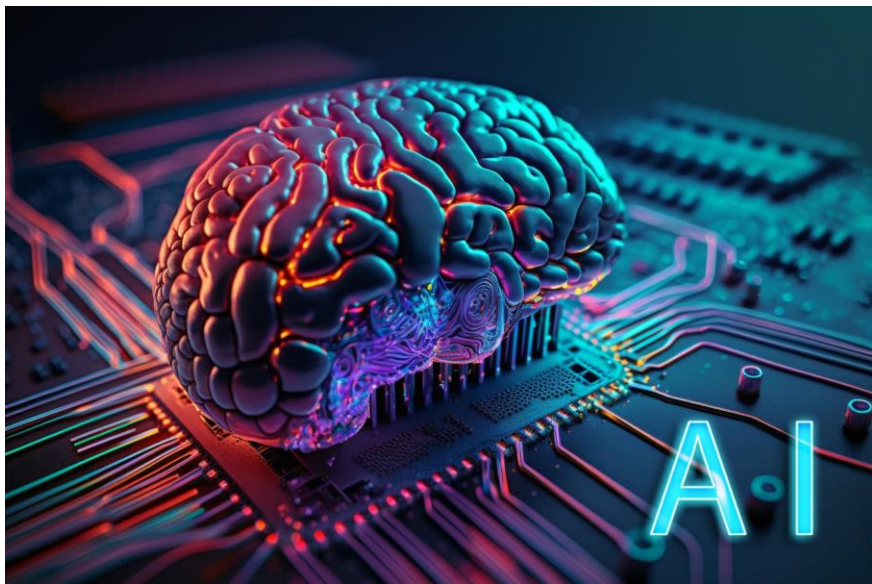Personalized Medicine & Cancer Prevention

02.
Early Detection & Diagnosis

03.
Effective Treatment & Reduced Treatment Side-effects

04.
Reduce relapse and improve QoL

The promise

AI is data hungry

Multiple types

AI models for personalized medicine require massive amounts of patient data, including genomics, medical histories, imaging, and lab results, to accurately predict individual responses to treatments.

Massive

Training AI systems to detect diseases early or recommend optimal therapies demands vast, diverse datasets, encompassing millions of cases, to ensure reliability and reduce bias.

Volume, Velocity, Variety, Veracity

The Economist
- Gordon Brown's pitch
- What went wrong at RBS
- Genetically modified crops blossom
- The EU woos Russia
- The right to eat cats and dogs

The data deluge
AND HOW TO HANDLE IT: A 14-PAGE SPECIAL REPORT

Challenges

Bloomberg
Data: The New Currency
#TheNewCurrency
Sponsored by:
ANZ

Data Heterogeneity

In both syntax and semantics. There is a need for better data curation, better tools to respect established standards, and common data models. Effective ETL tools.

Data Quality

Inconsistent, poor quality, and missing data: needs better data curation, and more rules when data are recorded. eCRF forms are more user-friendly and automatically completed
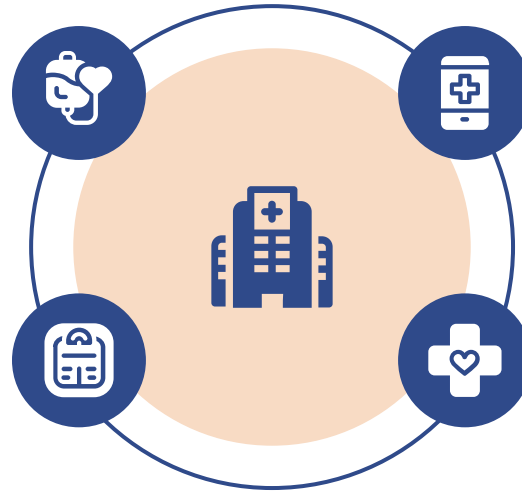
Hand-Written Reports

Clinicians use hand-writing notes: needs clinicians to use voice to text or write the text on a PC. Even better clinicians do not write any free text and only make selections.

Lack of Incentives

Clinicians in many cases believe they are the owners of the data and are reluctant to share them.

EU Initiatives

MyHealth@EU – Electronic cross-border health services in the EU
- Focus on **Patient Summary**, **ePrescription**

Joint Action Towards the European Health Data Space (TEHDAS):
- Prepares the ground for the **harmonized implementation of the secondary use of health data** in the European Health Data Space – EHDS

The VELES Excellence Hub
- Strengthening **smart health innovation in Southeast Europe**
- Creating a **sustainable, place-based innovation ecosystem**
- Development of a **Regional Smart Health Data Space**
- **Integrate and optimize health data across borders**, aiming to **improve outcomes in critical areas of public health**, including amongst others, **cancer treatment as one of the Regional Smart Health Data Space pilot** projects it focuses upon.

EU Initiatives

ProCAncer-I An AI Platform integrating imaging data and models, supporting precision care through prostate cancer's continuum

- Deliver a PCa AI platform featuring **a unique collection of PCa mpMRI images worldwide**, in terms of data quantity, quality, and diversity (> 17,000 Cases, >1,5 Million images)
- Develop **novel AI clinical tools** based on a three-stage ensemble modeling process for advancing the characterization of PCa lesions, assessment of the metastatic potential, and early detection of disease recurrence;

European Federation for CAncer IMages (EUCAIM)

- The cornerstone of the European Commission-initiated European Cancer Imaging Initiative, a flagship of Europe's Beating Cancer Plan (EBCP)
- Aims to foster innovation and deployment of digital technologies in cancer treatment and care to achieve more precise and faster clinical decision-making, diagnostics, treatment, and predictive medicine for cancer patients.

How?

Technological Innovation

- ➢ Through common data models and common terminologies
  - a. Ontologies vs Common Data Models (OMOP-CDM, FHIR, CDA, OpenEHR)
  - b. Terminologies (SNOMED-CT, LOINC, ICD)
  - c. Message exchange (FHIR, HL7)
  - d. Common implementation guides (IHE)

- ➢ Through methods and tools to enable the harmonization and integration of data

- ➢ AI developers with clinical understanding

- ➢ Methodologies for enabling Trustworthy AI

How?

Regulatory/Governance Innovation

Organizations are willing and capable to work together towards a collaborative medical data ecosystem, but:

1. appropriate infrastructures should ensure that **they remain, owners of their data,**
2. can **identify who is using their dataset and why**
3. have an **appropriate and sustainable pricing model**
4. they are **properly acknowledged in the developed AI models**

Key features of a data ecosystem

1. Clinicians to contribute their datasets, hosted on their own premises if needed or pushed to a central repository if there is the capability to do so.
2. All sources adopt the same common data model
3. A meta-data catalog will allow search on the meta-data of the available datasets and issue requests for access
4. Data owners can accept or reject data access requests
5. A pricing model
6. Upon data access is granted AI models can be sent to be trained in the datasets

> Shared, controlled vocabulary and clear semantic relationships among data from different sources can significantly enhance the integration process, leading to more effective and personalized cancer treatments, whereas governance and regulation problems are equally important and should not be neglected.

Thank you for your attention!

Prof. Haridimos Kondylakis
kondylakis@hdhc.gr