Version 1.0 del 16.12.2022

TITLE: CANCER MULTI-OMICS AVATARS FOR INTEGRATED PRECISION MEDICINE

SHORT TITLE: LANTERN

PROMOTER: Fondazione Policlinico Universitario A. Gemelli, U.O.C. di Thoracic Surgery– I.R.C.C.S.

PRINCIPAL INVESTIGATOR: Filippo Lococo, U.O.C. di Thoracic Surgery – I.R.C.C.S. Fondazione Policlinico Universitario A. Gemelli

ABSTRACT

Lung cancer is one of the most commonly diagnosed forms of cancer worldwide, accounting for approximately 12.2% of total cancer diagnoses. Despite the recent advances in treatment, prognosis still remains poor with a high mortality rate.

The current therapeutic approaches have reached a significantly high level of complexity, due to the recent identification of mutations and other cancer heterogeneity drivers (i.e., actionable mutations or PD-L1 expression for immunotherapy indication), thereby complicating the process of decision-making by clinicians.

The integration of "omics" datasets in different domains (i.e., genomics, proteomics) may further support clinicians in identifying innovative variables to support a new kind of comprehensive cancer care, towards personalized medicine.

In the last decade, significant improvements have been observed in the development of omics technologies, resulting in an overwhelming number of variables to be considered while making decisions that can be hardly handled by a human operator. Interestingly, with the advent of advanced Artificial intelligence (AI) techniques, more information can be obtained via "Internet of Things" technologies and various omics datasets which could be used efficaciously in creating predictive models.

These models can be used can be used either individually or merged in a multi-omics patient avatars to assist in the clinical decision-making process and introduce new paradigms of care and research, evaluating unexpected links among variables that were previously unexplored.

Therefore, the primary aim of the LANTERN project is to develop accurate predictive models for lung cancer patients, using various omics-based variables and integrating well-established clinical factors with "big data" and advanced imaging features. With this, individual patient avatars can be created which will serve as powerful tools in guiding clinicians and researchers in making more accurate decisions.

AUTHORS

Principal Investigators

• **Prof. Filippo Lococo** - U.O.C. di Thoracic Surgery – Catholic University of Sacred Heart; I.R.C.C.S. Fondazione Policlinico Universitario A. Gemelli – filippo.lococo@policlinicogemelli.it

Sub-Investigators

- **Prof. Troost Esther** Chair of Department of Radiotherapy and Radiation Oncology of University Hospital and Faculty of Medicine Carl Gustav Carus of the Technische Universität Dresden, Dresden, Germany- esther.troost@uniklinikum-dresden.de;

- **Prof. Ádány Róza** Head of the MTA-DE Public Health Research Group, University of Debrecen (DEB) - adany.roza@med.unideb.hu;

- **Dr. Farré Núria** Radiation Oncologist, Consultant of Hospital de la Santa Creu i Sant Pau (HSCSP), Spain- nfarre@**santpau.cat;**

- Dr. Van Doorne Dominique President of Accademia del Paziente Esperto EUPATI – AdPEE, Italy - segreteria@accademiadeipazienti.it

- **Dr. Öztürk** Ece Asst. Prof. Of Koc University, Turkey- ozturkece@ku.edu.tr

FINANCING

Source

The proposal is funded with reference to the Joint Transnational Call for Proposals (JTC) 2022 on "PREVENTION IN PERSONALISED MEDICINE" within the framework of ERA PerMed.

RATIONALE OF THE STUDY

To date, the current strategy for lung cancer treatment is still based on traditional clinical variables (e.g. age, sex, TNM stage) that appear clearly inadequate to deal with such a complex disease. In the last years, the advent of genomics and precision medicine approaches have transformed the horizon, firstly for scientists and physicians, and later for lung cancer patients. Besides the classical wealth of histological, clinical and demographic information, today's technologies allow the easy capture of a wider range of data sources from genomics, proteomics, immunohistochemistry and imaging. All these additional sources need to be integrated by lung cancer physicians, in order to improve patient outcomes by stratifying or individualizing treatment decisions (personalised medicine approach). However, while this process is already available in the setting of controlled clinical trials, there is still a great limit to its application in daily clinical practice. Interestingly, with the introduction of Artificial Intelligence (AI) into clinical practice, physicians now potentially have the adequate technology to improve precision medicine approaches. Although, for AI applications to be successfully applied to the medical domain, an integrative machine learning (ML) system might be necessary, calling for the incorporation and fusion of heterogeneous datasets and diverse omics profiles.

Preliminary results from Consortium members

Our consortium aims at fulfilling the objective of this year's call through the intersectional, multinational and interdisciplinary collaboration between University and clinical research teams and private organisations towards the implementation of personalised medicine, by combining clinical research with bio-informatics components as well as ELSA research or implementation research. **FPG** in Rome (https://gstep.policlinicogemelli.it), the coordinating site, has a sound background in translational research of advanced quantitative imaging analysis and AI based applications for personalized medicine and has been involved in several international trials on lung cancer. In a retrospective study, FPG researchers explored technical issues in personalized medicine through the prediction of NSCLC survival by quantitative image analysis, demonstrating the usefulness of density correction of volumetric CT data (Farchione et al., 2020). Recently, this research team also participated in a study involving a distributed learning performed on over 20,000 lung cancer patients by providing a privacy-by-design infrastructure (Personal Health Train), connecting FAIR (Findable, Accessible, Interoperable, Reusable) data sources from 5 different countries which enhanced distributed data analysis and ML (Deist et al., 2020). **OncoRay**, a national Centre for Radiation Research in Oncology, located in Dresden (**TUD**) (https://www.oncoray.de/center), is focused on applying ML/AI approaches to enable personalized radiation oncology based on big data, i.e. clinical imaging data (radiomics). In previous studies, their research team carried out a comparative study of machine learning methods for time-to-event survival data for radiomics risk modelling. From this study, they were able to identify a subset of algorithms which should be considered in future radiomics studies to develop stable and clinically relevant predictive models for time-to-event endpoints (Leger et al., 2017). **HSCSP** in Barcelona has a strong lung oncology research group, participating in numerous clinical trials in chemo-radiotherapy in lung cancer (Provencio et al., 2021), treating around 400 patients/year. In a recent retrospective study, **HSCSP** research team explored the effects of non-adherence to external beam radiation therapy on 1-year survival in cancer patients receiving treatment with a curative in Catalonia, Spain (Borras et al., 2020). **DEB**, is a renowned institution of higher education in Hungary, contributing immensely to the progress of clinical research with expertise in public health and play a major role in Ethico-legal aspects of personalised medicine research. In a Previous research project, **FPG** and **DEB** partnered with the PRECeDI consortium (http://www.precedi.eu/site), releasing a set of five recommendations for policy-makers, scientists and industry, including ethico-legal and policy recommendations, to foster the integration of PM approaches in the field of chronic disease prevention (Boccia et al., 2019). **KU** is a well recognized institute with over 150 research centres, forums, and laboratories that conduct high quality scientific research projects with expertise in Thoracic Surgery, Pathology, Bioinformatics and Network Modelling and Engineered Cancer & Organ Models. In Previous studies, KU researchers were involved in the development of a three-dimensional hydrogel model that can mimic the native glycosaminoglycans in tissues (Öztürk et al., 2016; Öztürk et al., 2017; Öztürket al., 2020). **AdPEE** is a patient organisation that ensures the active involvement of patients and caregivers in clinical research for constant and productive dialogue with decision-making bodies. (https://accademiadeipazienti.org/accademia-paziente-esperto-adpee/).

SUMMARY OF PROJECT RATIONALE AND WORKPLAN

The LANTERN project was developed, based on 4 Work Packages (WP):

i. **WP1: Patient enrollment and omics data collection [Months 1-30].**

Objective: To gather information from all the clinical and omics based data sources acknowledged as clinically significant or decision support variables for lung cancer comprehensive diagnosis and therapy workflow definition. A structured terminological system will be developed for prospective data collection through specific Case Report Forms (CRFs).

Methodology: The five main omics data tiers to be collected in this project, reflecting all the involved omics domains in the lung-cancer decision making pathway are: A. Demographic and physiological data; B. Medical history, clinic-functional, laboratory data and treatment information; C. Histo-pathological, immunological and genetic data; D. Radiomics and quantitative imaging data; E. Internet of Things derived data.

ii. **WP2: Omics data archiving and inter-actionability [Months 2-30].**

Objective: To ensure easy and effective omics data inter-actionability and to allow their complete integration into archiving systems and ad hoc digital platforms for data recording.

Methodology: A standard unit of measurement will be used in order to record the single variables. Possible discrepancies originated from the use of different scales, will be solved through shared equivalence decision. All the laboratory data used in WP1 will undergo cross quality assurance (QA) procedures among the centres and all the images used in WP1 will have to adhere to the standard “dicom” format (.dcm).

iii. **WP3: Omics data modelling and avatar creation [Months 12-32].**

Objective: To set up dedicated multi-omics data models and create lung cancer Digital Human Avatars (DHA).

Methodology: An exploratory analysis across all collected datasets to enable the start of the biomarker identification process; followed by the set-up of multiple distinct modular multivariate models, trained though advanced ML and AI techniques; and validated through internal and external validation to test their robustness, transferability and generalizability.

iv. **WP4: DHA towards implementation in healthcare [Months 1-36].**

Objective: To enable feedback data loops for preventive healthcare strategies and Quality of life (QoI) and to ensure that all the potential stakeholders are involved in the process of development of DHA for a fast uptake into daily practice.

Methodology: All necessary factors (clinical, epidemiological, technical, economic, organizational, social, legal and ethical implications) will be considered for the proper implementation of the DHA in healthcare. To

this aim, a robust and rigorous scientific methodology will be applied by combining quantitative approaches (collection of scientific evidence and systematic literature reviews/meta-analysis; consultation of institutional flows and databases/hospital discharge forms/national price lists and data analyses; mathematical models; cost-utility/budget impact analyses) and qualitative research techniques (Consensus conference, Focus group, Delphi, Expert opinion).

The LANTERN project is configured as highly innovative, since for the first time in the field of lung cancer research, we intend to develop a predictive platform powered by the integrated reading of data from the main omics sciences. The project has the ambition to combine data from radiomics with clinical and biological data, derived from biological omics (through the analysis of the patient's biological samples) on a digital platform. This project though seems ambiguous, has a well-thought out plan and is highly feasible through the cooperative work efforts and experience of each consortium partner in their respective fields. With the commitment of each partnering centre in undertaking their respective delegated tasks as described in the project work plans and tasks, We believe the LANTERN project will be completed within a timeframe of 3 years which aligns with the expected time taken for the project completion as stated in the call document.

Personalised medicine represents an evolution from traditional “one size fits all” strategies, extending the concept of stratifying cancer treatments on well-known clinical and pathological characteristics to a more individual, patients specific, truly personalized level. However, the current approach typically relies on molecular and cellular data-sets (i.e. mutations) that are very limited and relative only to subsets of the complex web of cellular data that determine treatment outcomes. Due to various histological, genetic, immunological and imaging characteristics that define its highly variable oncological outcomes, lung cancer represents an ideal case to run AI based multidimensional personalized medicine studies, thereby fitting into the purpose of this project.

AIMS

Main objectives

The aim of LANTERN project is to develop accurate predictive models for lung cancer patients, using different omics-based variables, integrating well established clinical factors with big data and advanced imaging

features. This technological solution will be able to capture patients' conditions at multiple levels, supporting an effective disease prevention and further reducing the healthcare costs through personalized treatments.

More specifically, we will develop Digital Human Avatars (DHA) defined as computerized representations of individual patients and their characteristics, with a specific focus on advanced biological functions, as well as relevant pathological variables, which will be modelled, sourced and parametrized in a cutting-edge experimental "big data" setting.

Secondary objectives

Our main goals are:

i. To gather information from all the clinical and omics-based data sources acknowledged as clinically significant

ii. To ensure an easy and effective omics data inter-actionability and to allow their complete integration into archiving systems

iii. To set up dedicated multi-omics data models and create a comprehensive lung cancer DHA

iv. To enable feedback data loops for preventive healthcare strategies and quality of life (QoL) management

v. To ensure that all the potential stakeholders are involved in the DHA development process.

METHODOLOGY

<u>This is a prospective observational study</u>

WP1: Omics data collection

The five main omics data tiers that will be identified as the ontology of this project, reflecting all the involved omics domains in the lung-cancer decision making pathway are: A. Demographic and physiological data; B. Medical history, clinic-functional, laboratory data and treatment information; C. Histo-pathological, immunological and genetic data; D. Radiomics and quantitative imaging data; E. Internet of Things derived data.

A structured terminological system for prospective collection of anonymized data will be developed, creating a dedicated database for all omics domains considered. The data will be stored in a dedicated server and categorized according to the data reported in the WP1. Researchers will provide a structured terminological system for prospective collection of anonymized data in which all information will be recorded through specific Case Report Forms (CRFs), ensuring uniformity and consistency for all variables collected. Tumour

samples from all enrolling centres will be stored at the coordinating centre in a dedicated laboratory (**FPG**). In-house samples will be collected immediately after biopsy or surgery, while samples from other RUs will be delivered using an appropriate courier service following standard international rules and laws. A comprehensive molecular characterization and genomic profiling of tumour samples will be performed using the Illumina TruSightTM, which will allow the conversion of DNA and RNA extracted from formalin-fixed paraffin-embedded tissue samples into libraries enriched for cancer-related molecular targets. In our methodological approach, we will assess both DNA and RNA levels using a single high-throughput assay that evaluates a wide range of state-of-the-art clinical biomarkers. Specifically, we will analyze with high fidelity multiple somatic genomic alterations such as low-frequency somatic variants (single nucleotide variants, insertions and deletions, and splice variants) and copy number alterations (CNAs) in identified genes linked to cancer susceptibility and treatment.

WP2: omics data archiving and inter-actionability

The aim of this WP is to allow complete data integration into both existing and new archiving systems (Task 2.1: Data archiving QA), and to ensure an easy and effective omics data inter-actionability (Task 2.2: Inter-actionability QA). All the considered variables will be recorded according to a shared common ontology based on the variable-specific domains, in order to enhance their direct actionability. A standard unit of measurement will be used in order to record the single variables. Possible discrepancies originated from the use of different scales, will be solved through shared equivalence decision. All the laboratory data used in WP1 (Data A, B, C) will undergo cross quality assurance (QA) procedures among the involved centres prior to data upload in the shared platform. Patient data from **KU**, **DEB** and **HSCSP** will be transferred to **FPG** and **TUD** for the radiomic analysis. The shared general ontology will represent a structured terminological system for data archiving and analysis, granting uniformity and coherence for all the collected variables.

Proposed radiomic analysis will be as follows: first, gross tumour volume (GTV) delineation will be performed retrospectively by experienced lung malignancy Radiotherapy Oncologists on each patient's images (CT images in DICOM (.dcm) format). DICOM files containing not only the CT images but also the RT structure file will be exported and processed using MODDICOM, an R package developed to optimize automatic loading of DICOM images and radiomic analysis. Radiomic features will then be extracted automatically. The extracted radiomic features belong to different families: Intensity-based, morphological, textural and fractal. All analyzed image features will be based on the standardization criteria of the IBSI initiative.

WP3: omics data modelling, avatar creation and knowledge transfer

Developing accurate predictive models and creating Digital Human Avatars (DHA) - Omics features identification and selection: In the first step, an exploratory analysis across all collected datasets (An estimate of ≈ 240 NSCLC patients) will enable the start of the biomarker identification process and restrict the cast amount of information towards a more selected pool of potential biomarkers. This first phase will employ robust data analysis techniques in order to identify relevant variables (T3.1.) in a univariate setting, taking individual statistical distributions, feature-relevant correlations and general descriptive statistics into account.

Patient selection

INCLUSION CRITERIA:

- Patients with (suspected) NSCLC
- Age >18 yrs
- ECOG 0-3
- Written Informed Consent

EXCLUSION CRITERIA:

- ECOG 4
- Psychosocial, or emotional conditions controindicating participation to the study

STATISTICAL ANALYSIS METHODOLOGY

An estimate of 600 Patients (both male and female) with ≥18 years of age with pathologically proven diagnosis of non-small cell lung cancer (NSCLC) will be enrolled. Sample size was calculated on the basis of G-powerR program for "a priori" analysis, considering α=5% and power of 90%.

The Wilcoxon Mann-Whitney (WMW) test will be used to investigate the ability of these features to predict clinical outcomes on univariate analysis, and the feature showing the lowest p-value will be considered as the most predictive parameter. Receiver Operating Feature curve analysis (ROC) will also be performed on the most significant feature obtained at WMW analysis, calculating the area under the curve (AUC) with the corresponding 95% confidence interval (CI) according to the Clopper-Pearson method. Model processing will be performed through advanced ML and AI techniques first on a dedicated training dataset to fit the model parameters. The obtained hyperparameters will be tuned on a dedicated validation dataset and finally a test dataset will be employed as a dataset independent of the training dataset but following the same probability distribution.

Predictive model development

The objective of the second phase is creating multiple distinct but modular multivariate models, trained through advanced ML and AI techniques and segmented into specific modular areas of interest (T.3.2).

Different supervised models will be developed including logistic regression, decision tree, support vector machine, random forest, XGBoost classifier, and artificial neural networks. The k-fold cross-validation will be used for hyperparameters tuning. The final step in the proposed methodology is the statistical significance comparison of the performance of the ML models. This will be done evaluating predictive performances in terms of different metrics: the first evaluation will be based on accuracy (number of subjects correctly classified on the total number of patients) while the other will be on precision (true positive on total test positive), recall (sensitivity), F1 score (2*precision*recall/(precision+recall)) and AUC-ROC.

We estimate about 300 NSCLC cases with complete data to start developing the models. Specific care will be taken to always ensure both user friendliness and model explainability as primary pillars of the model development strategies. Easily interpretable values such as SHAP (SHapley Additive exPlanations) values will be attached to each model in order to avoid any black-box approaches that might render model outputs, difficult to explain to the patients during their interactions with the clinicians.

Predictive model validation

Both the developed model and the comprehensive DHA will be validated in order to test their robustness, transferability and generalizability. Two consecutive validation strategies will be employed respectively: the internal and external validation techniques. We estimate a total number of ≈ 420 NSCLC cases to start the validation process. This process will include both internal and external validation. The internal validation step will be focused mainly on techniques such as K-fold cross validation. This will allow to achieve a Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD) level 2b. At this TRIPOD level, data are non-randomly split (e.g., by location or time) into 2 groups: one to develop the prediction model and the remaining dataset to evaluate its predictive performance. In the case of suboptimal performances, all the failing models will be reverted back to the previous stage and the models will be retrained with new specifications in order to increase their robustness. At this point, the models can be validated using external datasets, reaching TRIPOD 3 and 4 levels (External validation). Finally, when the models reach a TRIPOD 4 level, data transferability and full actionability will be granted, allowing the successful and safe application of the model in the clinical Routine (T3.3).

ETHICAL CONSIDERATIONS

Considering the technicality and complexity of the LATERN Project, consortium partners are well aware of the various ethical and legal questions relating to the project. The main ethical issues related to this project include:

the respect for persons and for human dignity; fair distribution of benefits and burden; the rights and interests of the participants; the need to ensure participants' informed consent; processing of personal data; concerns involving the development, deployment and/or use of artificial intelligence (AI)-based systems or techniques.

Therefore, LATERN partners will identify and implement a comprehensive set of ethical and legal requirements in accordance with Declaration of Helsinki principles to ensure compliance for all activities of the project. Appropriate ethical monitoring systems will also be set up.

To this effect, LATERN will implement the following activities:

Participation in this study must be entirely voluntary and fully informed consents of the participants must be obtained via project-specific **Informed consent forms** which must be documented before their participation in this project.

- Obtaining ethical approvals by the relevant competent national/local ethics authorities in relation with project protocol (i.e., For using or storing human cells or tissues for genetic testing), according to each country's specific ethics requirements and legislation
- Personal data obtained in this study must be processed in an pseudomised format and Information about the data processing operations and the contact details of the data protection officer must be provided to the participants in accordance to applicable EU regulations (art 13/art 14 GDPR).
- The development and deployment of AI-models will be based on specific ethical- approaches which must guarantee privacy and data protection throughout the system's lifecycle. The adopted approach will be built upon key prerequisites that align with "Assessment List for Trustworthy Artificial Intelligence"(ALTAI) in order to establish procedures to detect, assess the level and address potential risks for the production of ethically sound AI models.
- Monitoring compliance with set of ethical and legal requirements defined at EU and national level.

BIBLIOGRAPHY

- Molina JR, et al. AA. Non–small cell lung cancer: epidemiology, risk factors, treatment, and survivorship. Mayo Clin Proc. 2008;83:584-594
- Pastorino U, et al. Prolonged lung cancer screening reduced 10-year mortality in the MILD trial: new confirmation of lung cancer screening efficacy. Ann Oncol. 2019 Oct 1;30(10):1672. doi: 10.1093/annonc/mdz169. Erratum for: Ann Oncol. 2019 Jul 1;30(7):1162-1169. PMID: 31168572 [3] Holzinger A, et al. Knowledge discovery and data mining in biomedical informatics: the future is in integrative,

interactive machine learning solutions. Interactive knowledge discovery and data mining in biomedical informatics: state-of-the-art and future challenges. Lecture notes in computer science LNCS 8401. In: Holzinger A and Jurisica I, editors. Springer; 2014. p. 1–18.

- Salehi-Rad R, et al. The Biology of Lung Cancer: Development of More Effective Methods for Prevention, Diagnosis, and Treatment. Clin Chest Med. 2020 Mar;41(1):25-38. doi:
- 10.1016/j.ccm.2019.10.003. PMID: 32008627.
- Adam T, et al. The Multiple Potential Biomarkers for Predicting Immunotherapy ResponseFinding the Needle in the Haystack. Cancers (Basel). 2021 Jan 13;13(2):E277. doi: 10.3390/cancers13020277. PMID: 33451015.
- Canale M, et al. Role of liquid biopsy in oncogene-addicted non-small cell lung cancer. Transl Lung Cancer Res. 2019 Nov;8(Suppl 3):S265-S279. doi: 10.21037/tlcr.2019.09.15. PMID: 31857950; PMCID: PMC6894991.
- Farchione A, et al. Exploring technical issues in personalized medicine: NSCLC survival prediction by quantitative image analysis-usefulness of density correction of volumetric CT data. Radiol Med. 2020 Jul;125(7):625-635. doi: 10.1007/s11547-020-01157-3. PMID: 32125637
- Akinci D'Antonoli T, CT Radiomics Signature of Tumor and Peritumoral Lung Parenchyma to Predict Nonsmall Cell Lung Cancer Postsurgical Recurrence Risk.
- Aerts H, et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. Nat Commun 5, 4006 (2014).
- Chaddad A, et al. Predicting survival time of lung cancer patients using radiomic analysis. Oncotarget 8, 104393–104407 (2017).
- Chen B, et al. Development and clinical application of radiomics in lung cancer. Radiat Oncol 12, 154 (2017).
- [12] Dinapoli N, et al. Moddicom: a complete and easily accessible library for prognostic evaluations relying on image Features. Conf Proc IEEE Eng Med Biol Soc. 2015:771-4 (2015)
- Zwanenburg A, et al. The Image Biomarker Standardization Initiative: Standardized Quantitative Radiomics for High-Throughput Image-based Phenotyping. Radiology. 10:191145 (2020)
- Björnsson B, et al. Digital twins to personalize medicine. Genome Med. 2019 Dec 31;12(1):4. doi: 10.1186/s13073-019-0701-3.
- Croatti A, et al. On the Integration of Agents and Digital Twins in Healthcare. J Med Syst. 2020 Aug 4;44(9):161. doi: 10.1007/s10916-020-01623-5.
- Kato S, et al. An avatar for precision cancer therapy. Nat Biotechnol. 2018 Nov 9;36(11):10531055. doi: 10.1038/nbt.4293.
- Luengo-Fernandez R, et al. Economic burden of cancer across the European Union: a population-based cost analysis. Lancet Oncol. 2013 Nov;14(12):1165-74. doi: 10.1016/S14702045(13)70442-X. Epub 2013 Oct 14. PMID: 24131614.

• Frandsen BR, et al. Care fragmentation, quality, and costs among chronically ill patients. Am J Manag Care. 2015 May;21(5):355-62. PMID: 26167702.

• https://www.forbes.com/sites/gilpress/2017/01/20/6-predictions-for-the-203-billion-bigdata-analytics-market/#13d52dc12083

• Cesario A, D'Oria M, Calvani R, Picca A, Pietragalla A, Lorusso D, Daniele G, Lohmeyer FM, Boldrini L, Valentini V, Bernabei R, Auffray C, Scambia G. The Role of Artificial Intelligence in Managing Multimorbidity and Cancer. J Pers Med. 2021 Apr 19;11(4):314. doi: 10.3390/jpm11040314. PMID: 33921621; PMCID: PMC8074144.

• Kyriazakos S, Pnevmatikakis A, Cesario A, Kostopoulou K, Boldrini L, Valentini V, Scambia G. Discovering Composite Lifestyle Biomarkers With Artificial Intelligence From Clinical Studies to Enable Smart eHealth and Digital Therapeutic Services. Front Digit Health. 2021 Sep 6;3:648190. doi: 10.3389/fdgth.2021.648190. PMID: 34713118; PMCID: PMC8521973.

• Farchione A, Larici AR, Masciocchi C, Cicchetti G, Congedo MT, Franchi P, Gatta R, Lo Cicero S, Valentini V, Bonomo L, Manfredi R. Radiol Med. 2020 Jul;125(7):625-635. doi: 10.1007/s11547-020-01157-3. Epub 2020 Mar 3. PMID: 32125637.

• Deist TM, Dankers FJWM, Ojha P, Scott Marshall M, Janssen T, Faivre-Finn C, Masciocchi C, Valentini V, Wang J, Chen J, Zhang Z, Spezi E, Button M, Jan Nuyttens J, Vernhout R, van Soest J, Jochems A, Monshouwer R, Bussink J, Price G, Lambin P, Dekker A. Distributed learning on 20 000+ lung cancer patients - The Personal Health Train. Radiother Oncol. 2020 Mar;144:189-200. doi: 10.1016/j.radonc.2019.11.019. Epub 2020 Jan 3. PMID: 31911366.

• Leger S, Zwanenburg A, Pilz K, Lohaus F, Linge A, Zöphel K, Kotzerke J, Schreiber A, Tinhofer I, Budach V, Sak A, Stuschke M, Balermpas P, Rödel C, Ganswindt U, Belka C, Pigorsch S, Combs SE, Mönnich D, Zips D, Krause M, Baumann M, Troost EGC, Löck S, Richter C. A comparative study of machine learning methods for time-to-event survival data for radiomics risk modelling. Sci Rep. 2017 Oct 16;7(1):13206. doi: 10.1038/s41598-017-13448-3. PMID: 29038455.

• Provencio M, Majem M, Guirado M, Massuti B, de Las Peñas R, Ortega AL, Dómine M, Marsé R, Sala MÁ, Paredes A, Morán T, Vázquez S, Coves J, Larriba JLG, Sánchez JM, Vicente D, Farré N, Fornos LF, Zapata I, Franco F, Serna-Blasco R, Romero A, Isla D. Phase II clinical trial with metronomic oral vinorelbine and tri-weekly cisplatin as induction therapy, subsequently concomitant with radiotherapy (RT) in patients with locally advanced, unresectable, non-small cell lung cancer (NSCLC). Analysis of survival and value of ctDNA for patient selection. Lung Cancer. 2021 Mar;153:25-34. doi: 10.1016/j.lungcan.2021.01.005. Epub 2021 Jan 10. PMID: 33453470.

• Borras JM, Font R, Solà J, Macia M, Tuset V, Arenas M, Eraso A, Verges R, Farré N, Pedro A, Mollà M, Algara M, Solé JM, Mira M, Espinàs JA. Impact of non-adherence to radiotherapy on 1-year survival in cancer patients in Catalonia, Spain. Radiother Oncol. 2020 Oct;151:200-205. doi: 10.1016/j.radonc.2020.08.002. Epub 2020 Aug 6. PMID: 32771615.

• Boccia S, Pastorino R, Ricciardi W, Ádány R, Barnhoorn F, Boffetta P, Cornel MC, De Vito C, Gray M, Jani A, Lang M, Roldan J, Rosso A, Sánchez JM, Van Dujin CM, Van El CG, Villari P, Zawati MH. How to Integrate Personalized Medicine into Prevention? Recommendations from the Personalized Prevention of Chronic

Diseases (PRECeDI) Consortium. Public Health Genomics. 2019;22(5-6):208-214. doi: 10.1159/000504652. Epub 2019 Dec 5. PMID: 31805565.

- Öztürk, E., Arlov, Ø., Aksel, S., Li, L., Ornitz, D.M., Skjåk-Braek, G., and Zenobi-Wong, M. (2016). Sulfated Hydrogel Matrices Direct Mitogenicity and Maintenance of Chondrocyte Phenotype through Activation of FGF Signaling. Adv. Funct. Mater. 26, 3649–3662. https://doi.org/10.1002/adfm.201600092.

- Jahn, Ingeborg & Börnhorst, Claudia & Günther, Frauke & Brand, Tilman. (2017). Examples of sex/gender sensitivity in epidemiological research: Results of an evaluation of original articles published in JECH 2006-2014. Health Research Policy and Systems. 15. 10.1186/s12961-017-0174-z.

[29] Barrera-Rodriguez, R., & Morales-Fuentes, J. (2012). Lung cancer in women. Lung Cancer (Auckland, N.Z.), 3, 79–89. https://doi.org/10.2147/LCTT.S37319.