Ministero della Salute – Direzione Generale della Ricerca Scientifica e Tecnologica

Rendiconto di assegnazione risorse 5 per mille ANNO 2018 Contributo percepito€ 1.578.095,22

Enti della Ricerca Sanitaria

ENTE*: IRCCS ISTITUTO CLINICO HUMANITAS - HUMANITAS MIRASOLE SPA

Num. Prog.	Titolo del progetto	Fondi 5 per mille assegnati al progetto	Costo complessivo del progetto	Data indicativa di inizio progetto	Durata prevista
1	Development of a Machine Learning Prognostic Model for AYA (Adolescents and Young Adults) patients with cancer	163,900.00	163,900.00	01-07-21	36 mesi
2	Artificial intelligence for locally advanced head neck cancer treated with multi-modality adaptive radiotherapy: machine learning-based radiomic prediction of outcome and toxicity	141,900.00	141,900.00	01-07-21	36 mesi
3	Artificial Intelligence to predict prognosis and response to chemo-radiotherapy in colorectal cancer through the analysis of cancer-associated fibroblasts' and tumour-associated macrophages' expression	148,500.00	148,500.00	01-07-21	36 mesi
4	Preoperative diagnosis of periprosthetic joint infection in patients undergoing total hip and knee arthroplasty revision: development and validation of a machine learning algorithm	46,200.00	46,200.00	01-07-21	36 mesi
5	Use of Artificial Intelligence for the Development and Validation of Prognostic Models of Poor Clinical and Functional Outcomes in Patients with Ischemic Stroke Treated with Fibrinolysis and Thrombectomy	69,300.00	69,300.00	01-07-21	36 mesi
6	ELEGANT AF Electrical signals and Genome data integration through Artificial Intelligence in patients undergoing Atrial Fibrillation Ablation	84,700.00	84,700.00	01-07-21	36 mesi
7	ADDRESSING PERSONALIZED PRECISION MEDICINE FOR PATIENTS AFFECTED BY IMMUNE DISEASES THROUGH AI-BASED ANALYSIS OF CORTICOSTEROIDS USERS AND PREDICTION ON BIOLOGIC DRUGS PRESCRIPTION	77,000.00	77,000.00	01-07-21	36 mesi
8	Microbiota, metabolome and nutrition: an 'artificially intelligent' way to personalized nutrition	38,500.00	38,500.00	01-07-21	36 mesi
9	Identification of Artificial Intelligence based biomarkers to predict HCC response to medical treatment.	375,628.00	375,628.00	01-07-21	36 mesi
10	Machine-learning ready Endoscopy-Data Bank for personalized and precision medicine	377,467.22	377,467.22	01-07-21	36 mesi
11	Assessing early predictors of mortality and treatment strategies for life-threatening, time-dependent diseases in the Emergency Department using the Artificial Intelligence approach	55,000.00	55,000.00	01-07-21	36 mesi
		1,578,095.22			

Data 31-07-21

Il Legale Rappresentante

DR. LUCIANO RAVERA

Si autorizza al trattamento dei dati ai sensi del d.lgs. 196/2003

Il Legale Rappresentante

DR. LUCIANO RAVERA

*Istituzione beneficiaria del contributo del 5 per mille.

Ministero della Salute – Direzione Generale della Ricerca e dell'Innovazione in Sanità

Rendiconto 5 per mille ANNO 2018
Contributo percepito € 1.578.095,22 In data 03/07/2020

Ente della Ricerca Sanitaria

Denominazione Ente: IRCCS Istituto Clinico Humanitas

Codice fiscale: 10125410158
Sede legale: via Manzoni 56 – 20089 – Rozzano (MI)
Indirizzo di posta elettronica dell'ente: amministratore.delegato@humanitas.it
HUMANITASMIRASOLE@LEGALMAIL.IT
Dati del rappresentante legale: Luciano Ravera

Titolo del progetto: Development of a Machine Learning Prognostic Model for AYA (Adolescents and Young Adults) patients with cancer

Data di inizio progetto: 01/07/2021	**Data di fine progetto**: 30/06/2024
Fondi 5 per mille assegnati al progetto: € 163.900,00	**Di cui:** **Quota sostenuta entro l'anno di rendicontazione: 0,00** **Quota accantonata, da sostenere, per progetti pluriennali (durata massima tre anni): 163.900,00**

VOCI DI SPESA	Quota sostenuta entro l'anno di rendicontazione	Quota accantonata, da sostenere, per progetti pluriennali (durata massima tre anni)
Personale di ricerca (borsista, a contratto e di ruolo in quota parte)		149.000,00

Apparecchiature (ammortamento, canone di locazione/leasing)		
Materiale d'uso destinato alla ricerca (per laboratori di ricerca, acquisto farmaci ecc.)		
Spese di organizzazione (manifestazioni e convegni, viaggi e missioni ecc.)		
Elaborazione dati		
Spese amministrative		14.900,00
Altro (indicare quali)		
TOTALE	**0,00**	**163.900,00**

Data, 30/07/2021

Il Responsabile del Progetto
Prof. Armando Santoro

Il Legale Rappresentante

Si autorizza al trattamento dei dati ai sensi del d.lgs. 196/2003

Il Legale Rappresentante

Ministero della Salute – Direzione Generale della Ricerca e dell'Innovazione in Sanità

Rendiconto 5 per mille ANNO 2018
Contributo percepito € 1.578.095,22 In data 03/07/2020

Ente della Ricerca Sanitaria
Denominazione Ente: IRCCS Istituto Clinico Humanitas
Codice fiscale: 10125410158
Sede legale: via Manzoni 56 – 20089 – Rozzano (MI)
Indirizzo di posta elettronica dell'ente: amministratore.delegato@humanitas.it
HUMANITASMIRASOLE@LEGALMAIL.IT
Dati del rappresentante legale: Luciano Ravera

Titolo del progetto: Development of a Machine Learning Prognostic Model for AYA (Adolescents and Young Adults) patients with cancer

Relazione illustrativa

Introduction

AYA (Adolescent and Young Adult) Oncology represents a unique program dedicated to oncologic patients between the ages of 16 and 39 patients. The incidence of cancer in this subgroup of patients and the lack of outcome improvement in the last 20 years has caused an international alert because of the patients themselves, the intrinsic nature of the disease, and the inadequate treatment, often in absence of clinical trials. Understanding the complex biological mechanisms of these patient's survival using biological, clinical, and radio-pathological data is vital, not only to develop new treatments for patients, but also to improve survival prediction. Deep learning techniques capture nonlinear relationships, from their input and a flexible model design. Our novel pathway-based neural network, for survival analysis by integrating high-dimensional clinical, histological, and radiological data may provide reliable individual survival information and allow formulating personalized treatment recommendations.

Design, setting, and participants

In this population-based cohort study, an AI-based pipeline will be developed and validated using retrospective consecutive cases of newly diagnosed cancer (sarcoma, thyroid, breast, and colorectal) in AYA patients in Humanitas research Hospital of Rozzano from January 2010 until 31st of December 2018.

Data collection and Image annotation

We will record patient demographics, clinical baseline characteristics, tumor characteristics, staging (TNM), risk factors, imaging and pathological data of primary tumor and treatment details according to the different tumors. Patients will be classified according to different four survival

classes from Class 1 (, Dead 5 years after for cancer) to Class 4 (, Dead 5 years after for other reasons).
Two radiologists using free-hand region of interest (ROI) on staging radiological examination will annotate tumors. Therefore, radiomics features extracted by imaging will be used as in-put features to feed the machine-learning model. Guided by the pathologist, we will extract tumor region-of-interest (ROIs) from whole slides images. All images will be labelled according to the different four survival classes.
Study End Point
The primary end point will be to assess the disease-free survival time, defined as the time from the date of diagnosis to the date of the first recorded evidence of clinical (local or regional) recurrence and/or distant metastasis as confirmed by imaging, histologic evidence, or death by any cause. According to different disease-free survival, patients will be classified into four class as mentioned above.
Model development
For the prediction of risk in AYA cancer patients, we will develop a ML-based pipeline able to perform multimodal analysis to explore, assess and define valid prognostic survival models for different types (Sarcoma, Thyroid, Breast and Colo-Rectal cancer) incorporating clinical, histological, and radiological information in AYA (Adolescents and Young Adults) cancer patients. More precisely, the model will consist, on one side, of specific radiological and histo-pathological convolutional layers (specifically designed to treat image data), coupled with specifically designed clinical and demography-specific feed forward layers. Each layer will extract content-specific information (a deep encoding or deep monogram) that will be aggregated in an-higher lever by a merging neural module targeted for computing the final prediction
Conclusion
A ML-based classifier for disease-free survival for AYA patients will classify AYA patients into four survival classes.

BACKGROUND
AYA (Adolescent and Young Adult) is a worldwide-recognized acronym to indicate a specific subgroup of population from 16 to 39 years of age.
In oncology, AYA patients represent a specific cohort with unique clinical and psychosocial needs. They are no children and no adults; they stand up in the so-called “no man’s land”, between the pediatrician and the adult world. The development of dedicated programs comes mainly from three warnings: the high incidence, the lack of an expected outcome improvement as observed in children and adult oncology and finally the need of proper follow-up in case of long-term sequelae. Compared to the survival improvement observed in the last 20 years in the adult and pediatric oncology, the outcome of AYA cancer remains poor in many cancers, especially in sarcoma [1]. Several reasons were reported to explain this gap and many studies have associated the course of the disease to the patient’s youth. The diagnostic delay due to the opposite feelings of fear and invincibility is recurrent in young people and seems to justify the high incidence of advanced disease at the diagnosis reducing the possibility of curative treatment.
The peculiar biological characteristics of the disease make rare even a common cancer as breast or colon cancer. The frequency of specific molecular features changes significantly with age but is still poorly understood and the prevalence of predisposing germline mutations in AYA cancer is unknown, especially in cancers typically found in older adults [2].
Finally, the poor availability of clinical trials in this subset of patients’ impact on the prognosis, especially in rare tumors where the access to orphan drugs offer less therapeutic options for treatment [3].
Rare cancers are characterized by late or incorrect diagnosis, a limited number of expertise centers, clinical trials, registries and tissue banks, the lack of access to appropriate therapies, and the consistent lack of interest in developing new therapies due to limited market.
In our Institute several specialists are working together to bridge the existing gap of AYA Oncology, including the diagnostic, therapeutic and follow-up phase.
The precise stratification of AYA cancer patients based on Tumor Node Metastasis [4] staging and according to survival outcomes represents a crucial step in the treatment program. It has been found that other independent prognostic factors including age, sex, histology, genetic, biological, radiological data, and treatment choices could significantly contribute to individualized predictions of survival.

An increasing body of research has focused on developing prognostic tools, to help practitioners predict the outcome in patients with cancer [5]. To improve the precision of cancer survival estimation, Cox proportional hazard models have gained popularity as a way of predicting events. For example, different types of patients' nomograms have been a reliable tool that has demonstrated the ability to quantify risk by combining and clarifying significant clinical characteristics for clinical oncology [6].
In some rare cancers, nomograms possess the ability to derive more precise risk predictions when incorporated with TNM staging [7,8]
However, these models have several limitations with respect to time-to-event prediction for the clinical management of patients with cancer, including the precise evaluation of overall survival and time to progression [9].
Survival analysis is a group of methods used for estimating survival distribution from data, in which the outcome is the survival time until the observation has an event of interest. The most prevalent approach for analyzing time-to-event data in clinical trials is the Cox Proportional Hazards (Cox-PH) regression model. However, the main obstacles in this model are analyzing high-dimension, low-sample size data; and handling the highly nonlinear relationship between covariates.
Artificial intelligence algorithms based on machine learning networks can learn the highly intricate and linear/nonlinear associations between demographic, clinical and radiological characteristics and outcome. Deep learning techniques capture nonlinear relationships, from their input and have a flexible model design. Several deep learning models, which incorporate a standard Cox-PH model as an output layer, have been proposed for predicting patient survival [10]. Artificial intelligence and deep learning have been used for the analysis of histological imaging and applied to recognition, classification and prediction tasks [11,12,13,14].
Growing evidence suggests that radiology examinations contain more prognostic information for the primary tumor beyond the tumor dimension and anatomic information for the clinical T categorization. Radiomics, defined as the quantification of imaging features from imaging studies, is a precious source of biophysical data and should be investigated for a better patient selection. Deep learning models have a distinct advantage over the radiomics approach because they enable automatic feature extraction in an incremental manner without human intervention [15].
In addition to accurate prediction of the patient's survival, supervised machine learning–based algorithms are also used for classification where input values (e.g. an image associated to clinical record) are assigned to an output class (e.g. survival within a given time period after diagnosis). Classifiers offer the possibility of predicting with high accuracy the class to which a group of patients belongs (e.g. time after diagnosis) compared to accurate prediction of the patient's survival methods that are less precise and works inefficiently.
Main questions.
By combining expertise in radiomics (C. Giannitto, L.Balzarini. A. Chiti), pathology (L Terracciano, M Roncalli, S. Renne, P. Colombo), medical oncology (A. Bertuzzi, A. Santoro), thyroid cancer (G. Mercante, G. Spriano), breast cancer (A. Santoro), sarcoma (V. Quagliuolo, A. Bertuzzi), colorectal cancer (A. Spinelli) and AI-based approaches (V. Savevski, A. Carlucci), we will:
1. investigate the clinical, radiomics, and pathological features that are informative of the disease-survival
2. develop and validate a multimodal machine learning method, which could serve as a prognostic model in AYA patients affected by sarcoma, thyroid cancer, breast cancer, and colorectal cancer, by classifying these patients into four classes of disease-survival according to the different cancer.

APPROACH AND METHODOLOGY
Aim 1 (month 6-12). Investigate clinical, pathological and radiomics features that are informative of disease survival in AYA patients

Data set and Data collection
We will perform a population-based, retrospective prognostic study of data from AYA patients (18-39 ys) diagnosed with sarcoma, breast cancer, colorectal and thyroid cancer in our Institute from the 1st of January 2010 to the 31st of December 2018.
By performing a query on our institutional data ware-house DWH we have already recorded 281 cases of breast cancer, 148 cases of thyroid cancer, 95 cases of sarcoma, and 49 cases of colorectal cancer from January 2015 to the 14th of December 2020. We will search for patients from January 2010 to January 2015, who have not been uploaded on DWH.

The patients enrolled will be divided into the training and validation cohort.
We will collect the baseline information of patients (sex, age, and Performance status), tumor characteristics (location, size, histologic grade, histologic type, TNM stage), risk factors, and treatment details according to the different tumors.
We will record all the information suggested by guidelines for the different cancers.
Patients will be classified according to different four survival classes
-Class 1: Dead after 5 years for cancer
-Class 2 : Alive after 5 years with recurrence
-Class 3 : Alive after 5 years without recurrence
-Class 4: Dead after 5 years for other reasons
Study End Point
The primary end point will be to assess disease-free survival time, defined as the time from the diagnosis to the date of the first recorded evidence of clinical (local or regional) recurrence and/or distant metastasis as confirmed by imaging, histologic evidence, or death by any cause. The time of censoring will be determined as the date of the last follow up in our hospital.

Image annotation
Two radiologists will annotate staging ultrasound (US) images for thyroid cancer (1-2 images), mammography for breast cancer (1-2 images), and magnetic resonance imaging (MRI) or computed tomography (CT) for soft tissue sarcoma (1-2 images major axis) and CT or MRI for colorectal cancer. Images, their annotation and their classification, defined according to the aforementioned four risk classes, will be used as input to feed the multimodal machine-learning model.

Pathological slide image annotation
Tumor specimens will be accessed from archival material stored at pathology service. Guided by the pathologist, we will extract several tumor region of interest (ROI) from whole slides images.
All images will be labelled according to the different four survival classes.

Aim 2 (month 12-24): Model development
The multimodal learning model will consist of 2D radiological and histopathological convolutional layers (specifically designed to treat image data), as well as a clinical and demography-specific layer, each of which will extract specific informative content, aggregated in an higher layer by a merging neural module that will compute the final prediction by integrating all the information from the different data sources.
The multimodal architecture with side outputs will exploit attention-based modules to simultaneously capture both the inter and intra data relationships.
For the prediction of risk in AYA cancer patients, we will develop a multimodal learning model, which will integrate the information from radiological, histopathological, clinical, and demographic data.
More precisely, the model will consist, on one side, of specific radiological and histopathological convolutional blocks (specifically designed to treat image data), coupled with specifically designed clinical and demography-specific feed forward blocks. Each block will extract content-specific information (a deep encoding or deep monogram) that will be aggregated in an-higher lever by a merging neural module targeted for computing the final prediction
The multimodal architecture will avoid vanishing gradient effects by side outputs and will exploit attention-based modules to simultaneously capture both the inter and intra-data source relationships.
The usage of side output will also allow monitoring the network classification process, to uncover the strategies learnt by the trained network and exploited for prediction. This will prove the interpretability of the underlying biological mechanism, and of the histopathological patterns associated with different cancer survival classes.

IMPACT ON THE HEALTH SYSTEM
Patients' prognosis is one of the core aspects of classification of diseases (malignant Vs benign, i.e. if a disease can kill a patient or not). It is possibly the first step to discern which patients deserves the most attention in a limited resource setting as medical research is.

Through a multimodal deep learning approach relying on clinical, radiological, and pathological data we will identify what are the characteristics that precisely allow patients' stratification, defining the patients that need more care producing an algorithm that can perform automatically this complex and multidimensional task.

Riferimenti bibliografici

[1] Chen I, Pasalic D, Fischer-Valuck B, et al. Disparity in outcomes for adolescent and young adult patients with pediatric solid tumour across 4 decades. Am J Clin Oncol 2018; 41:471-5;
[2] Cha S, Lee J, Shin J-Y, et al. Clinical application of genomic profiling to find druggable targets for adolescent and young adult (AYA) cancer patients with metastasis. BMC CAncer 2016; 16:170)
[3] (McVeigh TP, Sundar R, Diamantis N, et al. The role of genomic profiling in adolescents and young adults (AYAs) with advanced cancer participating in phase I clinical trials. Eur J Cancer 2018; 95:20-9)
[4] Amin M, Edge S, Greene F, et al. AJCC cancer staging manual. Eighth ed. New York: Springer; 2017.
[5].Rabin BA, Gaglio B, Sanders T, et al. Predicting cancer prognosis using interactive online tools: a systematic review and implications for cancer care providers. Cancer Epidemiol Biomarkers Prev. 2013; 22(10):1645-1656. doi: 10.1158/1055-9965.EPI-13-0513).
[6] LiangW, Zhang L, Jiang G, et al. Development and validation of a nomogram for predicting survival in patients with resected non-small-cell lung cancer. J Clin Oncol. 2015; 33(8):861-869. doi:10.1200/JCO.2014.56.6661)
[7] Gold JS, Gönen M, Gutiérrez A, et al. Development and validation of a prognostic nomogram for recurrence-free
survival after complete surgical resection of localised primary gastrointestinal stromal tumour: a retrospective analysis. Lancet Oncol. 2009;:1045-1052. doi:10.1016/S1470-2045(09)70242-6.
[8] Callegaro D, Miceli R, Bonvalot S, et al. Development and external validation of two nomograms to predict overall survival and occurrence of distant metastases in adults after surgical resection of localised soft-tissue sarcomas of the extremities: a retrospective analysis. Lancet Oncol. 2016;17(5):671-680. doi:10.1016/S1470-2045 (16)00010-3)
[9] Randall RL, Cable MG. Nominal nomograms and marginal margins: what is the law of the line? Lancet Oncol. 2016;17(5):554-556. doi:10.1016/S1470-2045(16)00072-3)
[10] Katzman JL, et al.DeepSurv: personalized treatment recommender system using a Cox proportional hazards deep neural network. BMC Med Res Methodol. 2018; 18(1):24. https://doi.org/10.1186/s12874-018-0482-1.
[11] Wei JW, et al (2019) Automated detection of celiac disease on duodenal biopsy slides: A deep learning approach. arXiv preprint arXiv:1901.11447. [13] Li Y, Wu J, Wu Q (2019) Classification of breast cancer histology images using multi-size and discriminative patches based on deep learning. IEEE Access 7:21400–21408.
[14] Khan S, et al (2019) A novel deep learning based framework for the detection and classification of breast Cancer using transfer learning.
[15] Pattern Recogn Let .Wei JW, Tafe LJ, Linnik YA, Vaickus LJ, Tomita N, Hassanpour S (2019) Pathologist-level classification of histologic patterns on resected lung adenocarcinoma slides with deep neural networks. Sci Rep 9(1):3358. 1) Gatta G, van der Zwan JM , Siesling Sabine, et al: Rare cancers are not so rare: The rare cancer burden in Europe European journal of cancer (1990), 2011, Vol.47 (17), p.2493-2511

Innovazione e trasferibilità

As a new analytic tool, the deep learning network model will likely become more widely applied to support clinical decision-making. The performance of deep learning models in improving treatment outcomes is a key question and requires solid validation in the real world.
The advantages of the deep learning network model the adaptability to variables with a nonlinear association, which includes real world clinical factor and the integration of the nonlinear riskfunctions associated with outcomes. Deep learning possesses flexibility in dealing with complex clinical factors, automatically learning of feature-representations from un-interpreted clinical data, and analysis of censored factors. For this reason, our model could play a big role in biomedical marker analyses. Our model can facilitate discussion of different potential treatment options.
The machine learning methods applied in our study can be translated into tools for clinical treatment decision-making in the AYA population. The visualization of outcomes produced in this

study will be implemented in a national research database, that could be used by the clinicians to analyze the survival of the patients administered at their hospital.
In the future, the results of this study can be extended to other public hospitals for tailored treatment reducing public health costs.

Data, 30/07/2021

Il Legale Rappresentante

Si autorizza al trattamento dei dati ai sensi del d.lgs. 196/2003

Il Legale Rappresentante

Ministero della Salute – Direzione Generale della Ricerca e dell'Innovazione in Sanità

Rendiconto 5 per mille ANNO 2018
Contributo percepito € 1.578.095,22 In data 03/07/2020

Ente della Ricerca Sanitaria

Denominazione Ente: IRCCS Istituto Clinico Humanitas

Codice fiscale: 10125410158
Sede legale: via Manzoni 56 – 20089 – Rozzano (MI)
Indirizzo di posta elettronica dell'ente: amministratore.delegato@humanitas.it
HUMANITASMIRASOLE@LEGALMAIL.IT
Dati del rappresentante legale: Luciano Ravera

Titolo del progetto: Artificial intelligence for locally advanced head neck cancer treated with multi-modality adaptive radiotherapy: machine learning-based radiomic prediction of outcome and toxicity

Data di inizio progetto: 01/07/2021	**Data di fine progetto**: 30/06/2024
Fondi 5 per mille assegnati al progetto: € 141.900,00	**Di cui:** **Quota sostenuta entro l'anno di rendicontazione: € 0,00** **Quota accantonata, da sostenere, per progetti pluriennali (durata massima tre anni): € 141.900,00**

VOCI DI SPESA	Quota sostenuta entro l'anno di rendicontazione	Quota accantonata, da sostenere, per progetti pluriennali (durata massima tre anni)
Personale di ricerca (borsista, a contratto e di ruolo in quota parte)		90.000,00

Apparecchiature (ammortamento, canone di locazione/leasing)		
Materiale d'uso destinato alla ricerca (per laboratori di ricerca, acquisto farmaci ecc.)		20.000,00
Spese di organizzazione (manifestazioni e convegni, viaggi e missioni ecc.)		
Elaborazione dati		4.000,00
Spese amministrative		12.900,00
Altro (servizi di storage)		15.000,00
TOTALE	**0,00**	**141.900,00**

Data, 30/07/2021

Il Responsabile del Progetto
Prof.ssa Marta Scorsetti

Il Legale Rappresentante

Si autorizza al trattamento dei dati ai sensi del d.lgs. 196/2003

Il Legale Rappresentante

Ministero della Salute – Direzione Generale della Ricerca e dell'Innovazione in Sanità

Rendiconto 5 per mille ANNO 2018
Contributo percepito € 1.578.095,22 In data 03/07/2020

Ente della Ricerca Sanitaria
Denominazione Ente: IRCCS Istituto Clinico Humanitas
Codice fiscale: 10125410158
Sede legale: via Manzoni 56 – 20089 – Rozzano (MI)
Indirizzo di posta elettronica dell'ente: amministratore.delegato@humanitas.it
HUMANITASMIRASOLE@LEGALMAIL.IT
Dati del rappresentante legale: Luciano Ravera

Titolo del progetto: Artificial intelligence for locally advanced head neck cancer treated with multi-modality adaptive radiotherapy: machine learning-based radiomic prediction of outcome and toxicity

Relazione Illustrativa

Radical radiotherapy (RT), exclusive or in combination with systemic therapy, represents an effective therapeutic option for Squamous cell carcinoma of the head and neck (HNSCC). Despite the recent technological advancements in the field of radiation therapy, about 30-50% of patients will develop locoregional failure and radiation induced toxicity still represents a relevant concern.
Advanced imaging modalities appear to play an essential role in the customization of the radiation treatment as shown through the use of Adaptive Radiotherapy (ART) and radiomic. The combination of both ART and radiomic analysis could potentially be considered a further advance in the personalization of oncological treatments. For this reason, we designed the present research project with the aim to prospectively evaluate a machine learning-based radiomic approach to predict outcome and toxicity of HNSCC patients treated with ART by mean of CT, MRI and PET-scan. We will include patients with histological proven squamous cell carcinoma of the pharynx, larynx or oral cavity, staged as T3-T4 or N1-3, and candidate to radical radiotherapy +/- chemotherapy as primary treatment.
All the patients will be treated with Volumetric modulated arc therapy (VMAT) technique in its RapidArc form, with simultaneous integrated boost (SIB) technique. After RT start, at week 3 patients will repeat contrast simulation CT, together with new MRI and FDG-PET scan in order to replan the adapted treatment. Patient will start with the new RT plan in week 4. The quantitative analysis and features' extraction of the tumor heterogeneity will be performed on the tumor volume, defined on the simulation CT, co-registered with MRI and PET images.
The combination of radiomic features and clinical characteristics will be analyzed in order to categorize patients according to the risk of locoregional relapse after ART.

Background: Squamous cell carcinoma of the head and neck (HNSCC) is characterized by an incidence in Europe of 140.000 new cases per year, with survival rates at 5 years ranging from 25 to 65%.
Current clinical management algorithms for HNSCC patients involve the use of surgery and / or radiotherapy depending on the stage ofthe disease at diagnosis. adical radiotherapy (RT), exclusive or in combination with systemic therapy, represents an effective therapeutic option according to the international guidelines. Despite the recent technological advancements in the field of radiation therapy, about 30-50% of patients will develop locoregional failure after primary treatment of head and neck cancer. Moreover, although the development of Intensity modulated radiation therapy (IMRT) and Volumetric modulated arc therapy (VMAT) techniques allowed a greater sparing of dose on healthy tissues, radiation induced toxicity still represents a relevant concern, impacting on quality of life of cancer patients even for long time after treatment. The continuous effort of personalized medicine has the goal of improving patient's outcome, in terms of both disease's control and pattern of toxicity. Advanced imaging modalities appear to play an essential role in the customization of the radiation treatment as shown through the use of Adaptive Radiotherapy (ART) and radiomic. With ART we mean the adaptation of tumor volumes and surrounding organs at risk (OARs) to the shrinkage and patient emaciation during RT treatment.
The recent literature showed that tumor shrinkage can reach 70% by the end of the RT treatment, and at the same time OARs, such as parotid glands, can reduce their size by 7 to 70%13. These alterations, if not taken into account, can lead to an unexpected delivery of lower dose on the tumor and higher dose of OARs compared to what planned. Adaptive radiotherapy (ART) includes techniques that allow knowledge of patient-specific anatomical variations informed by Image-guided radiotherapies (IGRTs) to feedback into the plan and dose-delivery optimization during the treatment course. The persisting weakest link in the treatment chain for radiotherapy remains clinician-led target identification. Compared to CT or CBCT, MRI offers superior soft-tissue definition with no associated radiation risk. MRI identifies targets larger than on CT because tumour that otherwise would have been missed is now seen; however, most commonly, targets are reported to be smaller when delineated on MRI. The resulting smaller MRI-derived target improves the therapeutic ratio so enabling dose escalation. The availability of 'functional' MRI sequences holds promise that geometric adaptation maybe complemented by biological adaptation.
Diffusion-weighted imaging (DWI) is a functional imaging technique dependent on the random motion of water molecules to generate image contrast. As tumours usually have greater cellularity than normal tissue, they demonstrate higher signal intensity, i.e., restricted diffusion on MRI. This is reflected in the low mean apparent diffusion coefficient (ADC) value. This has potential to provide both qualitative and quantitative information. Change in the ADC has been used to identify early treatment response, and to predict local recurrence. Therefore, on-board DWI could identify early non responders who may benefit from change in treatment approach.
Radiomic is the extraction of quantitative features from medical images to characterize tumor pathology or heterogeneity. Radiomic features extracted from medical images can be used as input features to create a machine learning model able to predict survival, and to guide treatment thanks to its predictive value in view of therapy personalization. We previously evaluated in a retrospective study the qualitative analysis of the radiomic characteristics of head and neck tumor tissues, in order to identify a predictive signature of the biological characteristics of the tumor. We stratified HNSCC patients according to the most significant radiomic features into high- and low-risk groups of relapse and survival after radio-chemotherapy.

Approccio e metodologia:
The combination of both ART and radiomic analysis could potentially be considered a further advance in the personalization of oncological treatments, and in particular for radiation treatments. For this reason, we designed the present research project with the aim to prospectively evaluate a machine learning-based radiomic approach to predict outcome and toxicity of HNSCC patients treated with ART by mean of CT, MRI and PET-scan.
This is an observational prospective study with the objectives: To analyze patients' outcome and toxicity when treated with adaptive radiotherapy.
To collect and analyze multiple image scans performed at diagnosis and during RT treatment.
To apply machine-based radiomic analysis on the collected imaging and to correlate the radiomic signatures to the clinical outcome.
Primary end-point:

- Locoregional recurrence free survival at 1 year.

Secondary end-points:

- Evaluation of toxicity
- Quality of life after treatment, at 6 months and 1 year
- Progression Free Survival at 1 year
- Overall Survival at 1 year

Inclusion criteria:

- Age > 18 years
- ECOG Performance status 0 to 2
- Life expectancy > 12 months
- Histological proven squamous cell carcinoma of the pharynx, larynx or oral cavity
- Locally advanced stage disease classified as T3-T4 or N1-3
- Radical radiotherapy +/- chemotherapy indicated as the primary treatment modality
- Visible disease at the primary site on imaging performed within 4 weeks of starting treatment
- Adequate liver function
- Adequate renal function for infusion of iv. contrast for CT-scan and MRI-scan
- Adequate bone marrow function
- Written informed consent
- No previous radiation therapy on head and neck region

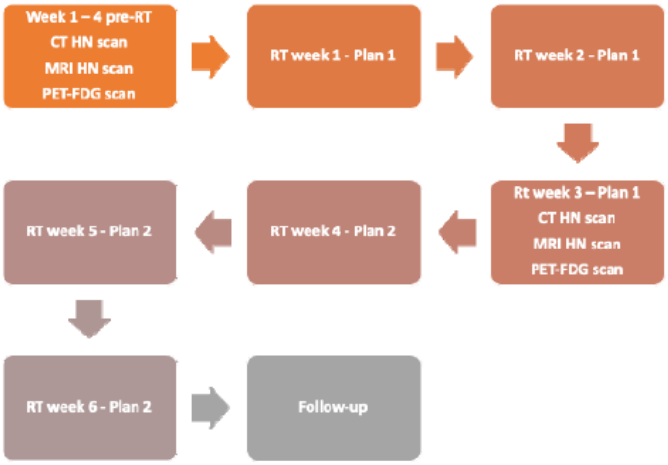
Exclusion criteria

- Inability to provide informed consent
- Presence of distant metastases
- Previous radiation therapy on head and neck region
- Pregnant or breastfeeding patients
- Prior malignancy within the last five years (except adequately treated basal cell carcinoma of the skin or in situ carcinoma of the skin or in situ carcinoma of the cervix, surgically cured, or localized prostate cancer without evidence of biochemical progression)
- Mental conditions rendering the patient incapable to understand the nature, scope, and consequences of the study
- Allergy or contraindication to contrast agents
- General contraindications to MRI
- ECOG PS >=3

Radiotherapy treatment and imaging.

All the patients will be treated with VMAT technique in its RapidArc form. A simultaneous integrated boost (SIB) technique will be used. The GTV will encompass the tumor delineated on CT scan, adjusted for MRI and PET scans. An expansion of 10 mm will be added to the GTV to create the CTV1. An expansion of 5 mm from GTV will be added to create the CTV2. Both CTVs will be adjusted for anatomical borders in which microscopic disease is unlikely to extend. Elective lymph node regions (CTV3) are defined based on tumour site and stage, delineation will be done according to published guidelines. Planning Target Volumes (PTV) will be constructed by the extension of the CTV with a 3 mm margin in all directions. Patients will be treated with a total dose of 66 Gy, 60 Gy and 54 Gy on PTV1, PTV2 and PTV3, respectively, delivered in 30 fractions, 5 fractions per week. Daily cone-beam CT scans will be made before every treatment session.

At week 3 from RT start, patients will repeat contrast simulation CT with, and MRI and FDG-PET scan for treatment replanning. Patient will start with the new plan in week 4.



Radiomic analysis: The quantitative analysis of the tumor heterogeneity will be performed on the gross tumor volume (GTV) defined on the contrast enhanced planning CT co-registered MRI and PET images. The registered contrast-free CT datasets will be used for the extraction of the textural features. No specific artifact correction methods will be applied to the datasets. A total of 41 features will be derived from the contrast-free CT images and grouped according to intensity, shape, and second and higher order features: the gray level co-occurrence matrix (GLCM); the neighborhood, gray level different matrix (NGLDM); the gray level run length matrix (GLRLM) and the gray level zone length matrix (GLZLM). From the histogram of the gray level distribution in the volume, the following features will be extracted: the minimum, maximum, mean and standard deviation of the Hounsfield units (HU) distribution, as well as the skewness, the kurtosis, the entropy, and the energy derived from this distribution. The CT images were re-sampled to symmetrical voxels of 2mm corresponding to the lowest image resolution. A HU binning was applied resampling the images in intervals of 8 HU. In addition to textural features, a number of predictors related to the treatment or to the patients will be considered.
This is an exploratory research designed to evaluate the predictive value of radiomic features in head and neck cancer patients treated with ART.
As per previous trials, and considering the incidence of the HNSCC, a total of 50 patients enrolled is estimated to the aim of this exploratory pilot study. Two years of enrolment and 1 year of follow-up from the last enrolled patient will be planned.

Impatto sul SSN:
The study does not include exams and treatments outside the standard of care for locally advanced pancreatic cancer.

Bibliografia
1. Siegel RL, Miller KD JA. Cancer statistics, 2018. CA Cancer J Clin. 2017;67:7-30.
2. Gatta G, Botta L, Sánchez MJ, et al. Prognoses and improvement for head and neck cancers diagnosed in Europe in early 2000s: The EUROCARE-5 population-based study. Eur J Cancer. 2015;51(15):2130-2143. doi:10.1016/j.ejca.2015.07.043
3. https://www.nccn.org/professionals/physician_gls/pdf/headand- neck.pdf.
4. Jeong S, Yoo EJ, Kim JY, Han CW, Kim KJ, Kay CS. Reirradiation of unresectable recurrent head and neck cancer: Using Helical Tomotherapy as image-guided intensity-modulated radiotherapy. Radiat Oncol J. 2013;31(4):206-215. doi:10.3857/roj.2013.31.4.206
5. Franzese C, Fogliata A, Clerici E, et al. Toxicity profile and early clinical outcome for advanced head and neck cancer patients treated with simultaneous integrated boost and volumetric modulated arc therapy. Radiat Oncol. 2015;10(1). doi:10.1186/s13014-015- 0535-0

6. Olteanu LAM, Madani I, De Neve W, Vercauteren T, De Gersem W. Evaluation of Deformable Image Coregistration in Adaptive Dose Painting by Numbers for Head-and-Neck Cancer. Int J Radiat Oncol. 2012;83(2):696-703. doi:10.1016/j.ijrobp.2011.07.037
7. Capelle L, Mackenzie M, Field C, Parliament M, Ghosh S, Scrimger R. Adaptive Radiotherapy Using Helical Tomotherapy for Head and Neck Cancer in Definitive and Postoperative Settings: Initial Results. Clin Oncol. 2012;24(3):208-215. doi:10.1016/j.clon.2011.11.005
8. Sriram P, Syamkumar SA, Kumar JSD, Prabakar S, Dhanabalan R, Vivekanandan N. Adaptive volumetric modulated arc treatment planning for esophageal cancers using cone beam computed tomography. Phys Medica. 2012;28(4):327-332. doi:10.1016/j.ejmp.2011.10.006
9. van Kranen S, Mencarelli A, van Beek S, Rasch C, van Herk M, Sonke J-J. Adaptive radiotherapy with an average anatomy model: Evaluation and quantification of residual deformations in head and neck cancer patients. Radiother Oncol. 2013;109(3):463-468. doi:10.1016/j.radonc.2013.08.007
10. Schwartz DL, Garden AS, Thomas J, et al. Adaptive Radiotherapy for Head-and-Neck Cancer: Initial Clinical Outcomes From a Prospective Trial. Int J Radiat Oncol. 2012;83(3):986-993. doi:10.1016/j.ijrobp.2011.08.017
11. Brouwer CL, Steenbakkers RJHM, Langendijk JA, Sijtsema NM. Identifying patients who may benefit from adaptive radiotherapy: Does the literature on anatomic and dosimetric changes in head and neck organs at risk during radiotherapy provide information to help? Radiother Oncol. 2015;115(3):285-294. doi:10.1016/j.radonc.2015.05.018
12. Castelli J, Simon A, Louvel G, et al. Impact of head and neck cancer adaptive radiotherapy to spare the parotid glands and decrease the risk of xerostomia. Radiat Oncol. 2015;10(1):6. doi:10.1186/s13014-014-0318-z
13. Yao W-R, Xu S-P, Liu B, et al. Replanning Criteria and Timing Definition for Parotid Protection-Based Adaptive Radiation Therapy in Nasopharyngeal Carcinoma. Biomed Res Int. 2015;2015:1-8. doi:10.1155/2015/476383
14. Njeh CF. Tumor delineation: The weakest link in the search for accuracy in radiotherapy. J Med Phys. 2008;33(4):136-140. doi:10.4103/0971-6203.44472
15. Roques TW. Patient selection and radiotherapy volume definition - Can we improve the weakest links in the treatment chain? Clin Oncol. 2014;26(6):353-355. doi:10.1016/j.clon.2014.02.013
16. Noel CE, Parikh PJ, Spencer CR, et al. Comparison of on board low-field magnetic resonance imaging versus onboard computed tomography for anatomy visualization in radiotherapy. Acta Oncol (Madr). 2015;54(9):1474-1482. doi:10.3109/0284186X.2015.1062541
17. Lütgendorf-Caucig C, Fotina I, Stock M, Pötter R, Goldner G, Georg D. Feasibility of CBCT-based target and normal structure delineation in prostate cancer radiotherapy: Multi-observer and image multi-modality study. Radiother Oncol. 2011;98(2):154-161. doi:10.1016/j.radonc.2010.11.016
18. Oh S, Stewart J, Moseley J, et al. Hybrid adaptive radiotherapy with on-line MRI in cervix cancer IMRT. Radiother Oncol. 2014;110(2):323-328. doi:10.1016/j.radonc.2013.11.006
19. Ahmed M, Schmidt M, Sohaib A, et al. The value of magnetic resonance imaging in target volume delineation of base of tongue tumours - A study using flexible surface coils. Radiother Oncol. 2010;94(2):161-167. doi:10.1016/j.radonc.2009.12.021
20. Dubois DF, Prestidge BR, Hotchkiss LA, Prete JJ, Bice WS. Intraobserver and interobserver variability of MR imaging- and CTderived prostate volumes after transperineal interstitial permanent prostate brachytherapy. Radiology. 1998;207(3):785-789. doi:10.1148/radiology.207.3.9609905
21. Rasch C, Keus R, Pameijer FA, et al. The potential impact of CT-MRI matching on tumor volume delineation in advanced head and neck cancer. Int J Radiat Oncol Biol Phys. 1997;39(4):841-848. doi:10.1016/S0360-3016(97)00465-3
22. Vaugier L, Ferrer L, Mengue L, Jouglar E. Radiomics for radiation oncologists: are we ready to go? BJR|Open. 2020;2(1):20190046. doi:10.1259/bjro.20190046
23. Cozzi L, Franzese C, Fogliata A, et al. Predicting survival and local control after radiochemotherapy in locally advanced head and neck cancer by means of computed tomography based radiomics. Strahlentherapie und Onkol. 2019;195(9):805-818. doi:10.1007/s00066-019-01483-0

24. Biau J, Lapeyre M, Troussier I, et al. Selection of lymph node target volumes for definitive head and neck radiation therapy: a 2019 Update. Radiother Oncol. 2019;134:1-9. doi:10.1016/j.radonc.2019.01.018
25. Wahl RL, Jacene H, Kasamon Y, Lodge MA. From RECIST to PERCIST: Evolving Considerations for PET Response Criteria in Solid Tumors. J Nucl Med. 2009;50(Suppl_1):122S-150S. doi:10.2967/jnumed.108.057307
26. Common Terminology Criteria for Adverse Events (CTCAE) Version 4.0 Published: May 28, 2009 (v4.03: June 14, 2010) U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES National Institutes of Health National Cancer Institute.
27. Zhang P, Simon A, Rigaud B, et al. Optimal adaptive IMRT strategy to spare the parotid glands in oropharyngeal cancer. Radiother Oncol. 2016;120(1):41-47. doi:10.1016/j.radonc.2016.05.028
28. Hu Y-C, Tsai K-W, Lee C-C, et al. Which nasopharyngeal cancer patients need adaptive radiotherapy? BMC Cancer. 2018;18(1):1234. doi:10.1186/s12885-018-5159-y

Innovazione e trasferibilità

The project will aim to combine radiomic features extraction and clinical advantage of adaptive radiotherapy to improve outcome of head and neck cancer patients and categorized characteristics of disease at higher risk of locoregional relapse.

Data, 30/07/2021

Il Legale Rappresentante

Si autorizza al trattamento dei dati ai sensi del d.lgs. 196/2003

Il Legale Rappresentante

Ministero della Salute – Direzione Generale della Ricerca e dell'Innovazione in Sanità

Rendiconto 5 per mille ANNO 2018
Contributo percepito € 1.578.095,22 In data 03/07/2020

Ente della Ricerca Sanitaria

Denominazione Ente: IRCCS Istituto Clinico Humanitas

Codice fiscale: 10125410158
Sede legale: via Manzoni 56 – 20089 – Rozzano (MI)
Indirizzo di posta elettronica dell'ente: amministratore.delegato@humanitas.it
HUMANITASMIRASOLE@LEGALMAIL.IT
Dati del rappresentante legale: Luciano Ravera

Titolo del progetto: Artificial Intelligence to predict prognosis and response to chemo-radiotherapy in colorectal cancer through the analysis of cancer-associated fibroblasts' and tumour-associated macrophages' expression

Data di inizio progetto: 01/07/2021	**Data di fine progetto**: 30/06/2024
Fondi 5 per mille assegnati al progetto: € 148.500,00	**Di cui:** **Quota sostenuta entro l'anno di rendicontazione: € 0,00** **Quota accantonata, da sostenere, per progetti pluriennali (durata massima tre anni): € 148.500,00**

VOCI DI SPESA	Quota sostenuta entro l'anno di rendicontazione	Quota accantonata, da sostenere, per progetti pluriennali (durata massima tre anni)
Personale di ricerca (borsista, a contratto e di ruolo in quota parte)		51.400,00

Apparecchiature (ammortamento, canone di locazione/leasing)		
Materiale d'uso destinato alla ricerca (per laboratori di ricerca, acquisto farmaci ecc.)		68.600,00
Spese di organizzazione (manifestazioni e convegni, viaggi e missioni ecc.)		
Elaborazione dati		5.000,00
Spese amministrative		13.500,00
Altro (storage)		10.000,00
TOTALE	**0,00**	**148.500,00**

Data, 30/07/2021

Il Responsabile del Progetto
Prof. Antonino Spinelli

Il Legale Rappresentante

Si autorizza al trattamento dei dati ai sensi del d.lgs. 196/2003

Il Legale Rappresentante

Ministero della Salute – Direzione Generale della Ricerca e dell'Innovazione in Sanità

Rendiconto 5 per mille ANNO 2018
Contributo percepito € 1.578.095,22 In data 03/07/2020

Ente della Ricerca Sanitaria
Denominazione Ente: IRCCS Istituto Clinico Humanitas
Codice fiscale: 10125410158
Sede legale: via Manzoni 56 – 20089 – Rozzano (MI)
Indirizzo di posta elettronica dell'ente: amministratore.delegato@humanitas.it
HUMANITASMIRASOLE@LEGALMAIL.IT
Dati del rappresentante legale: Luciano Ravera

Titolo del progetto: Artificial Intelligence to predict prognosis and response to chemo-radiotherapy in colorectal cancer through the analysis of cancer-associated fibroblasts' and tumour-associated macrophages' expression

Relazione illustrativa

Colorectal cancer (CRC) is one of the most frequent tumours worldwide whose incidence is increasing in people aged < 50 (early-onset colorectal cancer: EOCRC). The molecular mechanisms underlying CRC and EOCRC growth, recurrence, and response to therapy are still unclear. Late-onset CRC and EOCRC share most of the molecular mechanisms driving the development and progression, but some pathways- such as those linked to the epithelial-mesenchymal transition- differ in EOCRC, suggesting that it may represent a distinct subtype.
The presence of cancer-associated fibroblasts (CAFs) and tumour-associated macrophages (TAMs) in CRC microenvironment correlates with tumour progression and drug resistance. The deficiency of the receptor for complement C3 anaphylatoxin C3a (C3aR) promotes anticipated tumorigenesis and localization in the last part of the colon (unpublished)-mimicking the mechanisms of EOCRC- and associates with increased expression of the permeability blood endothelial marker PV1.
We aim to assess the pattern of CAFs, TAMs, C3aR, and PV1 expression in CRC and EOCRC specimens to develop a biomarker of prognosis and response to chemo-radiotherapy.
We will scan the haematoxylin and eosin (H&E) stained section of the CRC patients operated from January 1st 2010 to December, 31st 2020 and analyse them with artificial neural networks. A subgroup analysis for age and stage will be performed, to explore possible differences in the cellular and molecular profiles in EOCRC.
We expect the neural networks to interfere with predictive output (local ordistant recurrence, impaired response to chemo-radiotherapy) once received adequate input (stained specimens and corresponding clinical and pathology data).

This study will contribute to identify specific CRC patterns and new potential targets to prevent cancer progression and overcome tumour multidrug resistance.

Background

Colorectal cancer (CRC) is one of the most frequent malignant tumours worldwide1 and is currently increasing in people aged < 50 (early-onset colorectal cancer: EOCRC)2. EOCRC represents the third leading cause of cancer mortality in the western young population2 and is now considered a specific and distinct subtype of CRC, according to biological, molecular, morphologic, and genetic features3. Late-onset CRC and EOCRC share most of the mechanisms underlying development and progression, but some of them are prevalent in young patients: epithelial-mesenchymal transition, in co-occurrence with immunosuppression, seems to specifically drive EOCRC4.

Although multiple efforts have contributed to a better characterization and stratification of the disease, the molecular mechanisms guiding CRC growth and recurrence are still unclear5: the reasons underlying the early recurrence of a Stage I-II CRC, or the wide differences in the individual response to chemotherapy regimens are still debated.

The presence of cancer-associated fibroblasts (CAFs) and tumour-associated macrophages (TAMs) in CRC microenvironment associates with tumour progression and drug resistance6-11. CAFs may play a pivotal role in EOCRC development, regulating the epithelial-mesenchymal transition of CRC cells through a complex paracrine crosstalk12, 13. CAFs are also enriched in the stromal component of chemo-treated CRC patients, suggesting a role in cancer drug resistance12, 13.

Several pathways contribute to recruit and activate CAFs and TAMs in the tumour microenvironment, including the complement system14.

The group of Maria Rescigno found that a deficiency of the receptor for complement C3 anaphylatoxin C3a (C3aR) promotes anticipated tumorigenesis and localization in the last part of the colon (unpublished). This process closely resembles the mechanisms of EOCRC development and associates with increased expression of the permeability blood endothelial marker PV1.

This study aims to assess the pattern of CAFs, TAMs, C3aR, and PV1 expression in CRC and EOCRC patients, by analysing scanned conventional haematoxylin and eosin (H&E) stained section using artificial intelligence networks. This will help to develop a biomarker of prognosis and response to neoadjuvant/adjuvant chemo-radiotherapy. To highlight possible differencesin the cellular and molecular expression profiles between CRC and EOCRC we will also perform subgroup analysis according to the age of onset and stage.

Approach and methodology

We will analyse H&E tiles and corresponding complete clinical information of patient operated for CRC from January 1st, 2010 to December 31st, 2020. These data will be gathered from the files of the Division of Colon and Rectal Surgery. Formalin-fixed, paraffin-embedded samples will be analysed for CAFs histology and immunophenotype established based on the simultaneous immunohistochemical expression of α-SMA, fibronectin, FSP1, HHF35, and vimentin. TAMs will be analysed using human anti-CD68, and anti-CD163, antibodies15.

C3aR downregulation and PV1 upregulation will be analysed on H&E slides using already set-up technologies. PV1 staining will be coupled with staining for the blood vessels (CD34 or CD31) to normalize the data on the number of vessels.

H&E tiles will be scanned and collected. The following process of data preparation will include a pipeline for H&E tiles extraction in large amounts and a further dedicated pre-processing pipeline to make the H&E tiles Machine Learning ready.

We will create an image processing pipeline that will isolate the valid tissue inside each H&E tile and filter out any presence of artifacts stemming from the staining and preparation. The isolated tissue will be split into small manageable tiles to be fed to the Artificial Intelligence model.

Because we are framing this project as a classification task, ideally, we would require annotations at the pixel level: whether every part of the whole slide contains cancer or not. The diagnosis only gives global information at the slide level or suggest the position of malignant tissue with free text, which is not readily exploitable by convolutional neural networks. In case pixel-wise annotations would not be available, a weakly supervised learning paradigm will be a valid option to side-step the problem of not having per pixel annotations.

To explore additional possible correlations, we aim to test multiple neural networks including Convolutional Neural Networks (CNNs) and Deep Neural Networks (DNNs), widely used for inference-related tasks. The image features extracted, and the clinical data will then be linked and analysed to generate an AI-based prediction model. We expect the model to interfere with a predictive output (local or distant recurrence, impaired response to chemo-radiotherapy) once received adequate input (H&E tiles and corresponding clinical and pathology data).

Impact on National Healthcare:
This study will contribute to a better characterization of the molecular mechanisms driving CRC growth and recurrence.
We aim to identify specific patterns of expression of CAFs and TAMs and their association with local and distant recurrence, or impaired response to chemo-radiotherapy to contribute to the optimization of treatments and surveillance programs.
We aim also to identify potential new targets to overcome tumour multidrug resistance.

Riferimenti bibliografici

1.Dekker, E., Tanis, P. J., Vleugels, J. L. A., Kasi, P. M., & Wallace, M. B. (2019).Colorectal cancer. *Lancet (London, England)*, *394*(10207), 1467–1480.https://doi.org/10.1016/S0140-6736(19)32319-0
2.Mauri, G., Sartore-Bianchi, A., Russo, A.-G., Marsoni, S., Bardelli, A., & Siena, S.(2019). Early-onset colorectal cancer in young individuals. *Molecular Oncology*,*13*(2), 109–131. https://doi.org/10.1002/1878-0261.12417
3.Silla, I. O., Rueda, D., Rodríguez, Y., García, J. L., de la Cruz Vigo, F., & Perea, J.(2014). Early-onset colorectal cancer: a separate subset of colorectal cancer.*World Journal of Gastroenterology*, *20*(46), 17288–17296.https://doi.org/10.3748/wjg.v20.i46.17288
4.Nam, S., & Park, T. (2012). Pathway-based evaluation in early onset colorectalcancer suggests focal adhesion and immunosuppression along with epithelial-mesenchymal transition. *PloS One*, *7*(4), e31685.https://doi.org/10.1371/journal.pone.0031685
5.Guinney, J., Dienstmann, R., Wang, X., de Reyniès, A., Schlicker, A., Soneson, C.,…Tejpar, S. (2015). The consensus molecular subtypes of colorectal cancer.*Nature Medicine*, *21*(11), 1350–1356. https://doi.org/10.1038/nm.3967
6.Zhang, R., Qi, F., Zhao, F., Li, G., Shao, S., Zhang, X., … Feng, Y. (2019). Cancer-associated fibroblasts enhance tumor-associated macrophages enrichmentand suppress NK cells function in colorectal cancer. *Cell Death & Disease*, *10*(4),273. https://doi.org/10.1038/s41419-019-1435-2
7.Wei, C., Yang, C., Wang, S., Shi, D., Zhang, C., Lin, X., … Xiong, B. (2019). Crosstalkbetween cancer cells and tumor associated macrophages is requiredfor mesenchymal circulating tumor cell-mediated colorectal cancer metastasis.*Molecular Cancer*, *18*(1), 64. https://doi.org/10.1186/s12943-019-0976-4
8.Kramer, N., Schmöllerl, J., Unger, C., Nivarthi, H., Rudisch, A., Unterleuthner, D.,…Dolznig, H. (2017). Autocrine WNT2 signaling in fibroblasts promotescolorectal cancer progression. *Oncogene*, *36*(39), 5460–5472. https://doi.org/10.1038/onc.2017.144
9.Heichler, C., Scheibe, K., Schmied, A., Geppert, C. I., Schmid, B., Wirtz, S.,
Neufert, C. (2020). STAT3 activation through IL-6/IL-11 in cancer-associated fibroblasts promotes colorectal tumour development and correlates with poor prognosis. *Gut*, *69*(7), 1269–1282. https://doi.org/10.1136/gutjnl-2019-319200
10.Gordon, S. R., Maute, R. L., Dulken, B. W., Hutter, G., George, B. M., McCracken,M. N., … Weissman, I. L. (2017). PD-1 expression by tumour-associatedmacrophages inhibits phagocytosis and tumour immunity. *Nature*, *545*(7655),495–499. https://doi.org/10.1038/nature22396
11.Aizawa, T., Karasawa, H., Funayama, R., Shirota, M., Suzuki, T., Maeda, S., …Unno, M. (2019). Cancer-associated fibroblasts secrete Wnt2 to promote cancerprogression in colorectal cancer. *Cancer Medicine*, *8*(14), 6370–6382.https://doi.org/10.1002/cam4.2523
12.Xuefeng, X., Hou, M.-X., Yang, Z.-W., Agudamu, A., Wang, F., Su, X.-L., … Wu, X.-L. (2020). Epithelial-mesenchymal transition and metastasis of colon cancer cellsinduced by the FAK pathway in cancer-associated fibroblasts. *The Journal ofInternational Medical Research*, *48*(6),300060520931242. https://doi.org/10.1177/0300060520931242
13.Fiori, M. E., Di Franco, S., Villanova, L., Bianca, P., Stassi, G., & De Maria, R.(2019). Cancer-associated fibroblasts as abettors of tumor progression at thecrossroads of EMT and therapy resistance. *Molecular Cancer*, *18*(1), 70.https://doi.org/10.1186/s12943-019-0994-2

14.Zhang, R., Liu, Q., Li, T., Liao, Q., & Zhao, Y. (2019). Role of the complementsystem in the tumor microenvironment. *Cancer Cell International*, *19*, 300.https://doi.org/10.1186/s12935-019-1027-3
15.Nakagawa, T., Ohnishi, K., Kosaki, Y., Saito, Y., Horlad, H., Fujiwara, Y., …Komohara, Y. (2017). Optimum immunohistochemical procedures for analysis ofmacrophages in human and mouse formalin fixed paraffin-embedded tissuesamples. *Journal of Clinical and Experimental Hematopathology : JCEH*, *57*(1),31–36. https://doi.org/10.3960/jslrt.17017

Innovazione e trasferibilità

This is the first study that applies artificial intelligence technologies to immune-stained paraffin-embedded CRC specimens to assess the pattern of CAFs, TAMs, C3aR, and PV1 expression to evaluate their association with local and distant recurrence, or impaired response to chemo-radiotherapy. Our project converges innovative methods with established molecular analyses and will help automatizing and standardizing the therapeutic strategies decision according to specific patterns of cellular and molecular expression. If we can predict CRC prognosis and therapy response using artificial intelligence on biopsies/surgical specimens, the patients will benefit from an early and tailored therapeutic strategy.

Data, 30/07/2021

Il Legale Rappresentante

Si autorizza al trattamento dei dati ai sensi del d.lgs. 196/2003

Il Legale Rappresentante

Ministero della Salute – Direzione Generale della Ricerca e dell'Innovazione in Sanità

Rendiconto 5 per mille ANNO 2018
Contributo percepito € 1.578.095,22 In data 03/07/2020

Ente della Ricerca Sanitaria

Denominazione Ente: IRCCS Istituto Clinico Humanitas

Codice fiscale: 10125410158
Sede legale: via Manzoni 56 – 20089 – Rozzano (MI)
Indirizzo di posta elettronica dell'ente: amministratore.delegato@humanitas.it
HUMANITASMIRASOLE@LEGALMAIL.IT
Dati del rappresentante legale: Luciano Ravera

Titolo del progetto: Preoperative diagnosis of periprosthetic joint infection in patients undergoing total hip and knee arthroplasty revision: development and validation of a machine learning algorithm

Data di inizio progetto: 01/07/2021	**Data di fine progetto**: 30/06/2024
Fondi 5 per mille assegnati al progetto: € 46.200,00	**Di cui:** **Quota sostenuta entro l'anno di rendicontazione: € 0,00** **Quota accantonata, da sostenere, per progetti pluriennali (durata massima tre anni): € 46.200,00**

VOCI DI SPESA	Quota sostenuta entro l'anno di rendicontazione	Quota accantonata, da sostenere, per progetti pluriennali (durata massima tre anni)
Personale di ricerca (borsista, a contratto e di ruolo in quota parte)		30.000,00

Apparecchiature (ammortamento, canone di locazione/leasing)		
Materiale d'uso destinato alla ricerca (per laboratori di ricerca, acquisto farmaci ecc.)		5.000,00
Spese di organizzazione (manifestazioni e convegni, viaggi e missioni ecc.)		
Elaborazione dati		2.000,00
Spese amministrative		4.200,00
Altro (servizi di storage)		5.000,00
TOTALE	**0,00**	46.200,00

Data, 30/07/2021

Il Responsabile del Progetto
Dott. Mattia Loppini

Il Legale Rappresentante

Si autorizza al trattamento dei dati ai sensi del d.lgs. 196/2003

Il Legale Rappresentante

Ministero della Salute – Direzione Generale della Ricerca e dell'Innovazione in Sanità

Rendiconto 5 per mille ANNO 2018
Contributo percepito € 1.578.095,22 In data 03/07/2020

Ente della Ricerca Sanitaria

Denominazione Ente: IRCCS Istituto Clinico Humanitas

Codice fiscale: 10125410158
Sede legale: via Manzoni 56 – 20089 – Rozzano (MI)
Indirizzo di posta elettronica dell'ente: amministratore.delegato@humanitas.it
HUMANITASMIRASOLE@LEGALMAIL.IT
Dati del rappresentante legale: Luciano Ravera

Titolo del progetto: Preoperative diagnosis of periprosthetic joint infection in patients undergoing total hip and knee arthroplasty revision: development and validation of a machine learning algorithm

Relazione illustrativa

Periprosthetic joint infections (PJIs) is an important cuase of failure in THA and TKA, occurring in ~2% of cases. The diagnosis is challenging since several classifications exist but none of them has an adequate sensitivity and specificity, therefore the diagnosis is often confirmed postoperatively by microbiological analysis. But also the latter presents some diagnostic issues; indeed, it has 15% of false negatives, possibly due to the bacterial biofilm formation. Therefore, the current project aims to develop and validate a machine learning algorithm that based on preoperative clinical and radiological data can identify PJI in patients undergoing total hip and knee arthroplasty revision. This algoryhtm would increase the accuracy of the preoperative diagnosis of PJI, leading to a grater precision in the delivery of the proper antibiotic therapy and improving the surgical strategy indication. The current project will develop the above mentioned algoryhtm in a two-step approach. In the first part (will last one year), the algorithm will be developed by using clinical and radiological data extracted from the electronic folders of patients who underwent THA and TKA revision in Humanitas Research Hospital between January 2015 and December 2020. The variable of interest to build the algoryhtm are: demographic data, comorbidities, past medical history, blood tests and radiological data. In the second part (will last two years), the algoryhtm will be validated on 350 consecutive patients eligible for THA and TKA revision in Humanitas Research Hospital. In particular, in the same sample of patients, the algorithm will be applied in order to perform a comparison of its diagnostic ability versus EBJIS criteria, one of the most common classification systems developed to formulate the diagnosis of PJI.

BACKGROUND

Periprosthetic joint infections (PJIs) account for 25% of failures after total hip and knee arthroplasty (THA and TKA). The incidence of PJIs is ~2% of total primary replacement procedures performed (~150.000/year in Italy) and it is expected to rise due to the increasing number of arthroplasties performed [1,2].

The diagnosis of infection is challenging, especially when caused by organisms of low virulence or when the host is immune- compromised. Nowadays, a single, efficient tool for the preoperative diagnosis of PJI lacks, and the diagnosis is made on the basis of various preoperative and intraoperative diagnostic tests [3]. Overtime, several classification systems have been developed to formulate a proper diagnosis of PJI such as the Musculoskeletal Infection Society (MSIS), the International Consensus Meeting (ICM), the Infectious Diseases Society of America (IDSA) and the European Bone and Joint Infection Society (EBJIS) criteria [3]. However, confirmation of the diagnosis of infection requires the isolation of a causative organism; therefore, the diagnosis is often confirmed postoperatively by microbiological analysis. Indeed, in some cases the diagnosis is made on the basis of routine cultures of samples taken from joints considered not to be infected prior to the revision surgery [4]. For this reason, the rate of PJI is thought to be substantially higher than reported by joint registries, for which data are collected at the time of surgery. Moreover, the biofilm formed by bacteria on the surface of the prothesic implant complicates the detection of the pathogen resulting in false negative microbiological cultures in up to 15% of cases [5]. To overcome this problem, different microbiological methods have been developed as the sonication [6], synovial sampling and prolonged incubation period [7]. Indeed, these techniques represent the state of art in the field of PJI. If on one side these recent innovations in technology especially for what concerns the sequencing techniques, improved the sensitivity of this test, on the other hand they are for the time being not cost-effective, hence not applicable in the routine diagnosis of PJI [8].

Therefore microbiological tests alone either lack sensitivity due to the “culture negative infections” or are too expensive to be routinely employed in the daily practice, hence the need to develop new not- microbiological systems for the diagnosis of PJI. The machine learning approach has become increasingly applied to medicine and represents a natural extension of traditional statistical approaches [9,10]. In this respect, a machine learning model could have the potential to formulate a preoperative diagnosis of PJI regardless the microbiological examinations. For this reason, the present study aims to develop and validate a machine learning algorithm and perform multivariable logistic regression analysis using preoperative clinical and radiological data to identify a PJI in patients undergoing total hip and knee arthroplasty revision.

SPECIFIC AIMS

AIM 1: To develop and validate a machine learning algorithm and perform multivariable logistic regression analysis using preoperative clinical and radiological data from a hospital database to identify a PJI in patients who underwent revision of THA and TKA.

Expected duration: 12 months.

AIM 2: To validate and to investigate the diagnostic ability for PJI of the algorithm in a prospective cohort of patients eligible for THA and TKA revision surgery. Expected duration: 24 months.

METHODS

Experimental design – AIM 1:

The algorithm will be developed by using clinical and radiological data extracted from the electronic folders of patients who underwent THA and TKA revision in Humanitas Research Hospital between January 2015 and December 2020.

The following clinical data will be extracted as predictive features: age, gender, side and type of arthroplasty, previous surgery in the index joint, body mass index, history of smoking, diabetes, congestive heart failure, chronic obstructive pulmonary disease (COPD), American Society of Anesthesiologists (ASA) class, Charlson comorbidity score, steroid use, hypertension, preoperative blood tests, serum ESR, and serum CRP.

The following radiological data will be extracted as predictive features: prosthesis malalignement, periprosthetic bone reactions such as hypertrophy or atrophy, radiolucency

lines and osteolysis, prosthetic loosening, polyethylene wear, and heterotopic ossifications.
The varus/valgus femoral component malpositioning will be defined with an angle between the longitudinal stem axis and the longitudinal femoral axis greater than 3° in the AP view of the pelvis [11]. The malalignement of the femoral component will be defined in the AP knee radiographs by a placement <5° or >9° valgus relative to the long axis of the femur, while on the lateral view by a placement <87° relative to the long axis of the femur. The malalignement of the tibial component will be defined in the AP view by a placement <87° relative to the long axis of the tibia, while on the lateral view by a placement with more than 5° of posterior inclination relative to the long axis of the tibia [12].
For both THA and TKA, the radiolucent line will be defined as a 1-mm thick radiolucency adjacent to the implant developed during the radiographic follow-up after the replacement procedure. Osteolysis will be defined as a zone of radiolucency with a thickness >1 mm adjacent to the implant developed during the radiographic follow-up after the replacement procedure [12].
Loosening of the stem will be defined as a progressive axial radiolucency greater than 3 mm, or a varus/valgus deviation from the femoral shaft axis greater than 3° [13]. Loosening of the acetabular cup will be defined by a change greater than 2 mm in the horizontal and/or vertical position with an adjacent radiolucent zone, or a radiolucent zone greater than 3 mm [14]. For the TKA, the two radiographic hallmarks of component loosening will be a wide or progressive zone of radiolucency at the interfaces around the components (>2 mm lucency at cement-bone interface or any lucency at the metal-cement interface or metal-bone interface in uncemented prostheses) and an interval change in position of the components [12]. In particular, a loose tibial component typically shifts into varus alignment relative to the long axis of the tibia; while a loose femoral component often shifts into a flexed position on the lateral view relative to the long axis of the femur.
For the THA, the polyethylene wear will be defined by the eccentric position of the femoral head respect of the acetabular cup in AP and/or lateral view. For TKA, the polyethylene wear will be defined by a not equivalent joint space medially and laterally in the AP view. PJI will be considered as the outcome variable of interest. Patients that presented with a failed or painful total hip and knee arthroplasty that underwent a PJI workup will be included. If all predictive features and outcomes recorded will not available, the patient will be excluded. Patients will be divided into 2 groups: (1) PJI group and (2) non-PJI group. Patients will be included in the PJI group if they fulfilled the EBJIS criteria 3 to confirm the diagnosis. A PJI will be defined as one of the following criteria: preoperative or intraoperative sinus tract or visible purulence; preoperative or intraoperative joint aspirate >2000/µl leukocytes or >70% granulocytes; 2 positive intraoperative tissue cultures (only one for very virulent pathogens); intraoperative synovial culture positive; sonicate culture positive; histological evidence of inflammation on intraoperative tissue samples (2 granulocytes per 10 high-power fields).

Development of the algorithm
The data will randomly divided into a training set (70%) and testing set (30%). The training set data will be used to develop and train a machine learning model for predicting PJI, with the predictive features as inputs and the PJI diagnosis outcome as output. Then, an over-sampling method, the Synthetic Minority Over Sampling (SMOTE) technique, will be used to balance the minority class (PJI group) and 4 different algorithms will be trained and tested. In particular, a multivariate Logistic Regression and ensemble methods (Random Forest, Gradient Boosting and Extreme Gradient Boosting) will be chosen for the analysis. Hyper-parameters for all models will be chosen using a randomized search algorithm, with the aim of jointly maximize F1 macro, ROC AUC and balanced accuracy. After this, all models will be refitted to maximize F1 macro score, which is the harmonic mean of Precision and Recall. Finally, the predicted probabilities of being in PJI group will be recalibrated using isotonic Logistic Regression.

Experimental design – AIM 2:
In a period of two years, all the consecutive patients eligible for THA and TKA revision in Humanitas Research Hospital will be enrolled. All the patients will undergo a preoperative PJI workup according to the current good clinical practice. Moreover, they will be divided in aseptic and

septic patients according to the EBJIS criteria. In the same sample of patients, the algorithm will be applied in order to perform a comparison of the diagnostic ability of the algorithm versus EBJIS criteria. The management will be decided according to the preoperative and intraoperative findings and microbiological results basing on the EBJIS criteria. The prediction of the algorithm will not available to the surgeons, and it will not affect the decision making process for the management of the included patients.

Sample size

In a period of two years, around 400 patients eligible for THA and TKA revision are expected to be enrolled. In our tertiary referral center, around 25% of patients undergoing THA and TKA revision are affected by PJI. For this reason, around 100 patients with PJI are expected to be enrolled. Assuming that the machine learning algorithm would be an AUC not inferior to 0.90, the required number of patients to achieve a power of 80% with a marginal error of 0.05 with 95% CI is 73 per each group (PJI and non-PJI group).

Statistical analysis

The diagnostic ability for PJI of the algorithm (sensitivity, specificity, positive and negative predictive values, positive and negative likelihood ratios) will be calculated using the ROC (receiver operating characteristic) curve according to the Youden index (J). An area under the curve (AUC) value greater than 0.9 will be considered suggestive of an excellent diagnostic power of the test. The Youden index is function of both sensitivity and specificity and allows to identify the optimum cut-off point to set the threshold; graphically, this point corresponds to the shoulder of the ROC curve or the point of the curve closest to the point that corresponds to $x = 0$ and $y = 1$ (equal to 100% sensitivity and specificity).

Multivariable logistic regression will be used to determine the relationships between predictive features and PJI. Adjusted odds ratios and 95% confidence intervals will be calculated for all predictive features. Statistical significance will be set at $P < 0.05$. The overall proportion of exact agreement between the two diagnostic systems (machine learning algorithm and EBJIS criteria) will be assessed. The weighted Kappa Cohen and Kappa statistic will be calculated to evaluate the measure of agreement using the Landis and Koch benchmarks: ≤0.2 (poor agreement); 0.21 to 0.40 (fair agreement); 0.41 to 0.60 (moderate agreement); from 0.61 to 0.8 (substantial agreement); 0.81 (excellent agreement). These analyses will be performed using STATA software.

Riferimenti bibliografici

(1) Torre M, et al. Progetto Registro Italiano ArtroProtesi. Terzo Report. Controllo e qualità dei dati. Roma: Il Pensiero Scientifico Editore, 2016.

(2) Kurtz S, et al. Projections of primary and revision hip and knee arthroplasty in the United States from 2005 to 2030. J Bone Joint Surg Am. 2007;89:780-5.

(3) Izakovicova P, et al. Periprosthetic joint infection: current concepts and outlook. EFORT Open Rev. 2019;4(7):482-494.

(4) Jacobs A M E, et al. The unsuspected prosthetic joint infection: incidence and consequences of positive intra-operative cultures in presumed aseptic knee and hip revisions. Bone Joint J. 2017;99- B(11):1482-1489.

(5) Palan J, et al. Culture-negative periprosthetic joint infections. EFORT Open Rev. 2019 Oct 7;4(10):585-594.

6) Evangelopoulos D S, et al. Sonication: a valuable technique for diagnosis and treatment of periprosthetic joint infections. Scientific World Journal 2013; 2013:375140.

7) Talsma D T, et al. Time to positivity of acute and chronic periprosthetic joint infection cultures. Diagn Microbiol Infect Dis. 2021;99(1):115178.

8) Aggarwal D, et al. Clinical utility and cost-effectiveness of bacterial 16S rRNA and targeted PCR based diagnostic testing in a UK microbiology laboratory network. Sci Rep. 2020;10(1):7965.

9) Fu S, et Al. Automated Detection of Periprosthetic Joint Infections and Data Elements Using Natural Language Processing. J Arthroplasty. 2020 Aug 5;S0883-5403(20)30869-X.

10) Bingham JS et Al. An Evidence-Based Clinical Prediction Algorithm for the Musculoskeletal Infection Society Minor Criteria. J Arthroplasty. 2018 Sep;33(9):2993-2996.

11) Morales de Cano J, et al. Early clinical results of a new conservative hip stem. Eur J Orthop Surg Traumatol. 2014;24(3):359-63.

Innovazione e trasferibilità

The differential diagnosis between PJIs and aseptic failure is particularly challenging in septic cases associated with low virulence and biofilm-forming pathogens. However, the correct preoperative diagnosis is of crucial importance in order to choose the proper antibiotic therapy and surgical strategy. Indeed, undetected and mistreated PJIs at the time of revision can result in persistence of the infection, failure of the revised implant, longer hospital stay, multiple surgeries, prolonged immobilization and rehabilitation, together with increased overall costs. Moreover, also the opposite case is problematic, since a patient without PJI treated for it will receive an unnecessary antibiotic therapy and a longer and more complicated surgery than needed, accounting for higher costs, higher distress health risks for the patient. This study is expected to improve the preoperative diagnostic ability of PJIs by a machine learning algorithm based on preoperative clinical and radiological patient- specific features.

Data, 30/07/2021

Il Legale Rappresentante

Si autorizza al trattamento dei dati ai sensi del d.lgs. 196/2003

Il Legale Rappresentante

Ministero della Salute – Direzione Generale della Ricerca e dell'Innovazione in Sanità

Rendiconto 5 per mille ANNO 2018
Contributo percepito € 1.578.095,22 In data 03/07/2020

Ente della Ricerca Sanitaria

Denominazione Ente: IRCCS Istituto Clinico Humanitas

Codice fiscale: 10125410158
Sede legale: via Manzoni 56 – 20089 – Rozzano (MI)
Indirizzo di posta elettronica dell'ente: amministratore.delegato@humanitas.it
HUMANITASMIRASOLE@LEGALMAIL.IT
Dati del rappresentante legale: Luciano Ravera

Titolo del progetto: Use of Artificial Intelligence for the Development and Validation of Prognostic Models of Poor Clinical and Functional Outcomes in Patients with Ischemic Stroke Treated with Fibrinolysis and Thrombectomy

Data di inizio progetto: 01/07/2021	**Data di fine progetto**: 30/06/2024
Fondi 5 per mille assegnati al progetto: € 69.300,00	**Di cui:** **Quota sostenuta entro l'anno di rendicontazione: € 0,00** **Quota accantonata, da sostenere, per progetti pluriennali (durata massima tre anni): € 69.300,00**

VOCI DI SPESA	Quota sostenuta entro l'anno di rendicontazione	Quota accantonata, da sostenere, per progetti pluriennali (durata massima tre anni)
Personale di ricerca (borsista, a contratto e di ruolo in quota parte)		53.000,00

Apparecchiature (ammortamento, canone di locazione/leasing)		
Materiale d'uso destinato alla ricerca (per laboratori di ricerca, acquisto farmaci ecc.)		
Spese di organizzazione (manifestazioni e convegni, viaggi e missioni ecc.)		
Elaborazione dati		
Spese amministrative		6.300,00
Altro (cloud)		10.000,00
TOTALE	**0,00**	**69.300,00**

Data, 30/07/2021

Il Responsabile del Progetto
Dott.ssa Simona Marcheselli

Il Legale Rappresentante

Si autorizza al trattamento dei dati ai sensi del d.lgs. 196/2003

Il Legale Rappresentante

Ministero della Salute – Direzione Generale della Ricerca e dell'Innovazione in Sanità

Rendiconto 5 per mille ANNO 2018
Contributo percepito € 1.578.095,22 In data 03/07/2020

Ente della Ricerca Sanitaria
Denominazione Ente: IRCCS Istituto Clinico Humanitas
Codice fiscale: 10125410158
Sede legale: via Manzoni 56 – 20089 – Rozzano (MI)
Indirizzo di posta elettronica dell'ente: amministratore.delegato@humanitas.it
HUMANITASMIRASOLE@LEGALMAIL.IT
Dati del rappresentante legale: Luciano Ravera

Titolo del progetto: Use of Artificial Intelligence for the Development and Validation of Prognostic Models of Poor Clinical and Functional Outcomes in Patients with Ischemic Stroke Treated with Fibrinolysis and Thrombectomy

Relazione illustrativa

Acute ischemic stroke (AIS) is a serious condition accounting for substantial premature death and long-term disability worldwide. This represents a major global public health issue being stroke the third most costly health condition in developed countries.
Urgent reperfusion of the ischemic brain is the most effective treatment either by means of intravenous administration of recombinant tissue plasminogen activator (thrombolysis, tPA) or by endovascular interventional techniques (thrombectomy, EIT). Both these treatments lead to a higher survival rate and more favourable outcomes. Although the advantages of these treatments are clear, the individual treatment responses are variable. Only a part of treated patients experiences a major neurologic improvement and has a more favorable long-term outcome. Individualized prediction of the outcome of AIS is challenging but necessary to assist the clinicians in deciding treatment and rehabilitation strategies.
The clinically complex characteristics of AIS patients limit an accurate estimation by conventional prediction models and scoring systems. Advanced machine learning algorithms based on the analysis of a great number of variables and data may offer a novel option for improving the outcome prediction.
The aim of our study is to take advantage from the Artificial Intelligence (AI) tools to identify and validate a prognostic model for good or poor outcome prediction in AIS patients centred on the current treatments based on intravenous (tPA) and endovascular (EIT) thrombolytic therapies administered individually or in combination but taking into account all possible variables extractable from the electronic health record (EHR). We foresee to obtain results, which will allow to improve prospectively the classification of patients by offering possibly a higher personalization of the cure.

Background

Acute Ischemic Stroke (AIS) is a serious condition affecting in Italy more than 200,000 people every year, of which 80% represent new cases and 20% recurrencies, representing a prevalence of 6.5% in the elderly population (65-84 years). In Italy it is the third cause of death, after heart diseases and cancer and about one third of the stroke survivors have a high degree of disability at 1 year from the acute event (913.000 in Italy) so much so that they can be defined as totally dependent [1]. This is due to neurodegeneration occurring following the ischemic hit, which exposes brain cells to a complex cascade of events caused by deprivation of oxygen, glucose and other essential nutrients. AIS represents a major global public health issue being stroke the third most costly health condition in developed country. To date urgent reperfusion of the ischemic brain territory either by means of intravenous administration of recombinant tissue plasminogen activator (thrombolysis, tPA) or by endovascular interventional techniques (thrombectomy, EIT) is the most effective treatment. Both these treatments lead to a higher survival rate and more favourable outcomes. Although the advantages of these treatments are clear, the individual treatment responses are variable. Only a part of treated patients experiences a major neurologic improvement and has a more favorable long-term outcome. Individualized prediction of the outcome of AIS is challenging but necessary to assist the clinicians in deciding treatment and rehabilitation strategies.

Several efforts to predict the outcome after AIS have been made and published [2-8], however these studies, either foreseeing major neurological improvements or other long-term outcomes, have demonstrated only moderately acceptable prediction capability. The clinically complex characteristics of AIS patients limit therefore an accurate estimation by conventional prediction models and scoring systems.

Advanced machine learning algorithms based on the analysis of a great number of variables and data may offer a novel option for improving the outcome prediction.

The aims of our study are:

1. To derive and validate, using artificial intelligence (AI) algorithms, some predictive models of poor outcomes and complications (e.g., hemorrhagic events, infections occurrence) for patient risk stratification and improved medical decision making.
2. To identify and construct a set of indicators and measures useful to track the complete clinical pathway of stroke patients in our hospital, from the hyperacute phase of the event, managed in the Emergency Department (ED) and Stroke Unit (SU) to the post-acute phase, concluded in the Neurological Rehabilitation Unit (NRU).

Study design, approach and methodology

A retrospective analysis of prospectively collected “real world” data of AIS patients treated in IRCCS Humanitas Research and Clinical Institute from January 2015 to December 2020 (estimated 500 patients) who underwent either intravenous (tPA) and/or endovascular (EIT) thrombolytic therapies. Patients with AIS (ICD-9-CM code 433.x, 434.x,) as a first event or recurrence will be included in the study.

Data Sources

A review of the ED, SU and NRU electronic health records (EHR) will be conducted in order to collect and extract structured data and measures useful for assessing stroke severity and the efficiency and effectiveness of diagnostic and therapeutic procedures in the continuum of care. The HER of the Neuro Rehabilitation Unit is already organized around a broad set of standardized binary indicators and validated scores useful to measure the complexity of care processes and functional outcomes of the patient.

Outcomes

The analysis will focus on short and medium-term outcomes. Short-term (up to 8 days after stroke): neurological and general complications and functional status measured as dependence in basic mobility and communicative disability.

Medium term (up to 30 and 60 days after stroke), measured as dependence in basic mobility and in basic activities of daily living, and performance in walking, upper limb dexterity, and specific neuropsychological impairments (e.g., aphasia, neglect). It should be noted that compared to the relevant literature, an important strength of the present study is that the assessment of medium-term outcomes is done in the continuum of the clinical track of stroke patients, with the possibility to control and measure also the influence of post-acute rehabilitation treatments.

Indicators and Measures (Potential Predictive Variables)

Figures 1, 2, and 3 show a map of indicators and possible measures, drawn from the literature and clinical experience, grouped according to the stroke patient-focused workflow in the ED, SU, and NRU, respectively. The items reported are all potential predictors of intermediate (e.g., early or late complications) or final health outcomes of the entire clinical track (e.g., functional outcomes). For example, if the outcomes of interest are bleeding or malignant edema (neurologic complications) at 1 day after fibrinolysis, all indicators and measures collected before the endpoints can be treated as potential predictors. Similarly, if the outcome of interest is dependence in basic mobility at day 8 after stroke, all indicators and measures recorded before the end point can be treated as potential predictors. Finally, if the outcome is independent walking at 60 days after stroke in patients admitted to rehabilitation, all measurement items collected before the end point should be treated as potential predictors.
And so on.
The data collection and harmonization phases (technically the preprocessing phase) are particularly important for predictive outcome modeling and there is a need for simple, reproducible, and consistent measures. The NRU has recently developed a simple risk score of poor functional outcome that is placed at a critical step in the clinical stroke pathway, namely that of rehabilitation triage in the Stroke Unit for the selection of patients and for planning of interventions on the basis of the risk [9].
The EHR already includes validated tools (e.g., the NIHSS score for stroke severity) and reliable measures (e.g., laboratory tests), but much of the information that clinicians currently use to make decisions does not have a standardized form that can be easily transformed into measurable units.
To this regard, for example, the interpretation of neuroradiological findings can be quite discordant among clinicians and assigning to neuroimaging data standardized scores may improve the reliability and accuracy of the data usage.
The dataset construction will be composed of all possible structured data extractable from the EHR. In addition, also data from the hospital management control (duration of hospital stay, costs of interventions) will be collected. It will be also considered to retrieve non structured data from the EHR by means of algorithms that have been developed by the AI team.
We aim to build an Artificial Intelligence pipeline able to perform risk stratification for stroke affected patients. The analysis will be conducted as follows:
• Relevant clinical data of included patients will be retrieved from DWH and laboratory data will be queried from relevant HER accordingly;
• An exploratory data analysis (EDA) will be performed on the available structured data to assess the informative content of the dataset, then a principal component analysis (PCA) will be executed to retrieve overall characteristics of the dataset before proceeding with final data curation;
• Predictive features will be extracted through a multivariate analysis of the structured dataset;
• Unsupervised learning methods will be applied to discover eventual outliers and generate optimal clusters arrangement;
• A set of neural networks will be tested to perform classification tasks. The classifier with the best performances will be tested and validated and finally the classifier will undergo a hyperparameters fine tuning process to optimize prediction performances.
Once the analysis will be performed and the predictive algorithms developed it will be possible to perform prospective validation analyses on independent cohorts.

Riferimenti bibliografici

1. Sanità, I.S.d. Progetto Cuore Istituto Superiore di Sanità, Epidemiologia e prevenzione delle malattie cerebro e cardiovascolari. Available from: http://www.cuore.iss.it/).
2. Abedi V, Goyal N, Tsivgoulis G, Hosseinichimeh N, Hontecillas R, Bassaganya-Riera J, Elijovich L, Metter JE, Alexandrov AW, Liebeskind DS, Alexandrov AV, Zand R.Abedi V, et al. Novel Screening Tool for Stroke Using Artificial Neural Network Stroke. 2017 48(6):1678-1681.
3. Chung CC, Chan L, Bamodu OA, Hong CT, Chiu HW.Chung CC, et al. Artificial neural network based prediction of postthrombolysis intracerebral hemorrhage and death. Sci Rep. 2020 25;10(1):20501
4. Dharmasaroja P, Dharmasaroja PA.Dharmasaroja P, et al. Prediction of intracerebral hemorrhage following thrombolytic therapy for acute ischemic stroke using multiple artificial neural networks. Neurol Res. 2012 Mar;34(2):120-8.
5. Feng R, Badgeley M, Mocco J, Oermann EK.Feng R. Deep learning guided stroke management: a review of clinical applications. J Neurointerv Surg. 2018 Apr;10(4):358-362.

6. van Os HJA, Ramos LA, Hilbert A, van Leeuwen M, van Walderveen MAA, Kruyt ND, Dippel DWJ, Steyerberg EW, van der Schaaf IC, Lingsma HF, Schonewille WJ, Majoie CBLM, Olabarriaga SD, Zwinderman KH, Venema E, Marquering HA, Wermer MJH; MR CLEAN Registry Investigators., Predicting Outcome of Endovascular Treatment for Acute Ischemic Stroke: Potential Value of Machine Learning Algorithms. Front Neurol. 2018 25;9:784.
7. Ramos LA, et al. Predicting Poor Outcome Before Endovascular Treatment in Patients With Acute Ischemic Stroke. Front Neurol. 2020.
8. Sirsat MS, Fermé E, Câmara J.Sirsat MS, et al. Machine Learning for Brain Stroke: A Review. J Stroke Cerebrovasc Dis. 2020 Oct;29(10):105162. Epub 2020 Jul 28.
9. Bernardini B, Cerina G, Colantonio V, Corsini C, Ghirmai S, Pagani M, Malosio M L, Marcheselli S, Corato M, Straffi L, Spinelli M C. Clinical complexity markers improve the prediction of functional outcome in patients whith ischemic stroke compared to disease-specific severity-of-illness measure. (Manuscript in preparation)

Innovazione e trasferibilità

Validation of outcome and complication predictors could help the clinical decision-making and therapeutic process in the acute phase and could provide a personalized clinical and rehabilitation track for patients with acute ischemic stroke. We foresee to obtain results, which will allow to improve prospectively the classification of patients by offering possibly a higher personalization of the cure. Also, in the management control field we foresee to obtain potentially interesting results.

Data, 30/07/2021

Il Legale Rappresentante

Si autorizza al trattamento dei dati ai sensi del d.lgs. 196/2003

Il Legale Rappresentante

Ministero della Salute – Direzione Generale della Ricerca e dell'Innovazione in Sanità

Rendiconto 5 per mille ANNO 2018
Contributo percepito € 1.578.095,22 In data 03/07/2020

Ente della Ricerca Sanitaria

Denominazione Ente: IRCCS Istituto Clinico Humanitas

Codice fiscale: 10125410158
Sede legale: via Manzoni 56 – 20089 – Rozzano (MI)
Indirizzo di posta elettronica dell'ente: amministratore.delegato@humanitas.it
HUMANITASMIRASOLE@LEGALMAIL.IT
Dati del rappresentante legale: Luciano Ravera

Titolo del progetto: ELEGANT AF Electrical signals and Genome data integration through Artificial Intelligence in patients undergoing Atrial Fibrillation Ablation

Data di inizio progetto: 01/07/2021	**Data di fine progetto**: 30/06/2024
Fondi 5 per mille assegnati al progetto: € 84.700,00	**Di cui:** **Quota sostenuta entro l'anno di rendicontazione: € 0,00** **Quota accantonata, da sostenere, per progetti pluriennali (durata massima tre anni): € 84.700,00**

VOCI DI SPESA	Quota sostenuta entro l'anno di rendicontazione	Quota accantonata, da sostenere, per progetti pluriennali (durata massima tre anni)
Personale di ricerca (borsista, a contratto e di ruolo in quota parte)		77.000,00

Apparecchiature (ammortamento, canone di locazione/leasing)		
Materiale d'uso destinato alla ricerca (per laboratori di ricerca, acquisto farmaci ecc.)		
Spese di organizzazione (manifestazioni e convegni, viaggi e missioni ecc.)		
Elaborazione dati		
Spese amministrative		7.700,00
Altro (Workstation dedicata e Classificatori)		
TOTALE	**0,00**	**84.700,00**

Data, 30/07/2021

Il Responsabile del Progetto
Dott.ssa Ilaria My

Il Legale Rappresentante

Si autorizza al trattamento dei dati ai sensi del d.lgs. 196/2003

Il Legale Rappresentante

Ministero della Salute – Direzione Generale della Ricerca e dell'Innovazione in Sanità

Rendiconto 5 per mille ANNO 2018
Contributo percepito € 1.578.095,22 In data 03/07/2020

Ente della Ricerca Sanitaria
Denominazione Ente: IRCCS Istituto Clinico Humanitas
Codice fiscale: 10125410158
Sede legale: via Manzoni 56 – 20089 – Rozzano (MI)
Indirizzo di posta elettronica dell'ente: amministratore.delegato@humanitas.it
HUMANITASMIRASOLE@LEGALMAIL.IT
Dati del rappresentante legale: Luciano Ravera

Titolo del progetto: ELEGANT AF Electrical signals and Genome data integration through Artificial Intelligence in patients undergoing Atrial Fibrillation Ablation

Relazione illustrativa

Background:

Atrial fibrillation

Atrial fibrillation (AF) is the most common cardiac arrhythmia, with a prevalence in Europe estimated approximately 1-2%. Though AF itself does not represent a lethal condition, it can increase risks of morbidity and mortality due to AF-related complications, such as heart failure and atrial thrombosis-related stroke. The incidence of AF increases with age and in our aging society it is estimated a dramatic rise in the coming decades in the number of patients that will require healthcare assistance due to AF. Clinically, AF can be classified into the following types depending on the episode duration: paroxysmal, persistent and permanent AF (Hindricks G., Potpara T.S., et al. 2020).

Arrhythmogenic activity initiating AF usually originates in the muscle sleeves of the pulmonary veins (PVs), where pacemaker activity from these cells is thought to result in the formation of ectopic beats that initiate AF. PVs are also thought to be important in the maintenance of the arrhythmia. The chaotic architecture and electrophysiological properties of these vessels provides an environment where AF can be perpetuated (Jais P. et al. 2002).

Therefore, pulmonary vein isolation (PVI) through radiofrequency energy is the cornerstone of the current ablation management of AF (Haïssaguerre M. et al. 1998).

PVI was demonstrated to be effective in paroxysmal AF, but outcomes of ablation are modest in persistent AF, and hampered by a high rate of AF recurrence. After an initially successful procedure, the rate of freedom from AF at one year is only 60–80% and many patients need to undergo a re-ablation procedure. In studies with follow-up to 3-5 years success rates are even lower at 30-50% (Ouyang F. et al. 2010).

A very large body of literature in the recent years has examined the value of alternative ablation techniques in addition to PVI in order to maximize the effectiveness of the procedures, but none of these approaches has been consistently shown to improve long-term success rates (Buch E. et al. 2016, Verma A. et al. 2015). Several genetic variants were also associated with AF risk development using both genome-wide association studies and candidate gene screening approaches (Ellinor P.T. et al. 2012). But none of them, so far, directly corelated patient-specific genomic data, e.g., ionic channels variants, to electrical activity directly found in PVs.
Therefore, deeper characterization of the disease in patients selected to undergo the PVI procedure is sought, in order to increase our understanding and improve management and treatment-choices.

Artificial intelligence in electrophysiology

Artificial intelligence (AI) has been heralded in the family of "disruptive" technology especially for healthcare (Jiang F. et al. 2017). AI tools use advanced algorithms to detect clinically applicable information hidden in the massive amount of data that can be applied to clinical decision making, driven by relevant clinical questions and suggestions (Murdoch and Detsky 2013, Dilsizian and Siegel 2014, Kolker, Ozdemir et al. 2016). Moreover, AI systems are aimed to reduce diagnostic and therapeutic errors that are inevitable in the routinely clinical practice (Lee, Nagy et al. 2013).
AI refers to machine-based data processing to achieve objectives that typically require human cognitive function. In the modern era, AI has mined dense data and provided the potential to classify complex patterns and novel representations of data beyond direct human interpretation. Machine Learning (ML) is a subdiscipline of AI and employs algorithms to learn patterns empirically from data, in a supervised or unsupervised manner. Supervised ML algorithms are capable of learning linear and non-linear relationship from labeled data. For this reason, ML extends the range of traditional statistics because it is able to identify nonlinear relationships and high-order interactions between multiple variables that may be challenging for traditional statistics (Feeny AK et al. 2020). Electrocardiogram signals (ECGs) are particularly suitable for ML algorithms and were the first applications of AI in cardiology (Zhao Q et al. 2005). In particular, AI applied to ECG interpretation enabled the rapid, reliable, and automated determination of ECG diagnosis.
Recently, in this field, AI interpretation of single- and 12-lead ECG proved to be equal and even superior to the ability of board-certified cardiologists to detect most common cardiac arrhythmias (Hannun A.Y. et al. 2019) and was proven able to also confer information about disease characteristics not typically diagnosed on ECG, e.g., serum potassium or left ventricular systolic function (Galloway C.D. et al. 2019, Attia Z.I. et al. 2019 a and b). These two capabilities represent complementary application of AI to medicine – scaling our current workflow and insuring quality but also adding value to a routine medical test (Feeny A.K. et al. 2020). ML was also adopted to diagnose AF (Attia ZI et al. 2019c) and identify potential sites for ablation, including termination sites of persistent AF (Alhusseini M et al. 2019). However, so far only a few small studies investigated though AI and ML endocardial electrograms (EGM) signals, recorded during AF ablation procedures. Atrial electrograms provide a wealth of information, which is rarely fully utilized in the current clinical practice and could reflect the heterogeneity of the disease. ML may improve phenotypic classification of AF to aid clinical evaluation and management. Cluster phenotyping not only enables better detection of AF but may also guide therapy, through characterization of risks and outcomes, inherent to certain cluster phenotypes that can be targeted with specific pharmacological, behavioral, or catheter-based therapies. Beyond clinical management, cluster phenotyping may also facilitate clinical investigation. For example, when designing a clinical trial for a novel investigational therapy, targeting a specific cluster phenotype may be more relevant than targeting a broad population (Feeny A.K. et al. 2020).
The data science revolution in the form of sophisticated ML algorithms and increasing availability of computing power opens up possibilities to manage this data overload, both in terms of learning from the data, integrating them with other datasets and inferring model parameters able to make increasingly accurate predictions. This is particularly applicable in the electrophysiology field, where ML will be increasingly able to merge several datasets, ranging from bioelectrical signals to imaging and even genomic data, changing clinical practice and research.
The aim of our study is to use supervised ML tools, to extract features of AF electrical signals, both from ECGs and EGMs, and integrate them with next generation sequencing data, in order to predict patient outcomes after PVI procedures and guide personalized treatment and follow-up strategies (***Figure 1***).

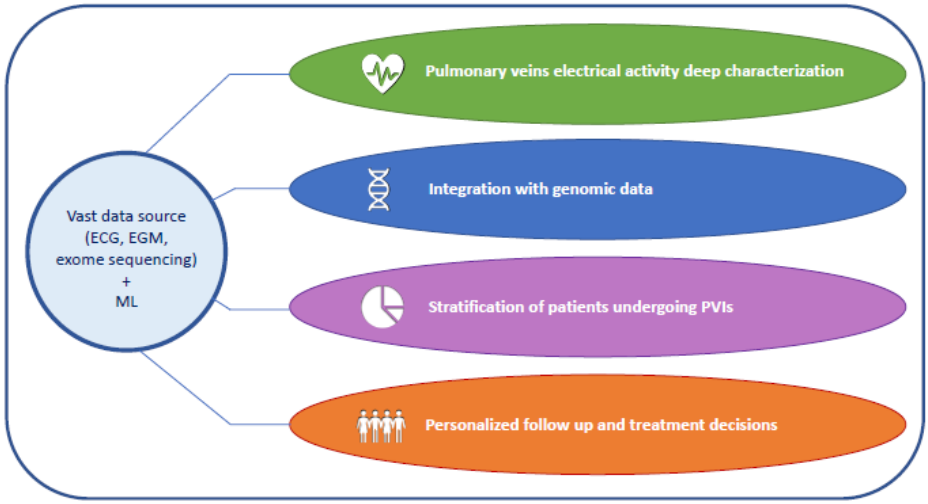


Figure 1. *Aim of the project*

Methods:

The following data will be collected from patients (n=60) undergoing PVI procedures:

1. 12-lead surface ECGs in sinus rhythm and AF, recorded before the ablation procedure.
2. Endocardial EGMs in sinus rhythm and AF (if available), recorded through Lasso catheter immediately before the ablation procedure.
3. Blood sample for Whole Exome sequencing.

We will develop, with the help of Prof. Miragoli at University of Parma and Prof. Fassina at University of Pavia, supervised ML classifiers in MATLAB® program language (Mathworks). The classifiers obtained will help to find a correlation between electrical signals, genomic data and patient outcomes after 1-year follow-up.

The matrix will be processed by MATLAB®'s Classification Learner Application to train different models such as:

• optimizable KNN (*k*-nearest neighbor classifier), via the "fitcknn" function with more than 100 optimization iterations and with no standardization of the input features.
• optimizable SVM (support vector machine classifier), via the "fitcsvm" function with more than 100 optimization iterations and with no standardization of the input features.
• discriminant analysis (Linear or Quadratic) via the "fitcdiscr" function that generate linear or non-linear boundaries between classes. All classifiers will be optimized via the minimization of the classification error and will be 10-fold cross-validated at least in order to avoid the over-fitting, which is an undesired memorization of the training data.

In addition, when the optimized training ends, we will obtain the confusion matrix and the preceding confusion matrices to compute the corresponding true positive rate (TPR) and false negative rate (FNR).

To test the robustness of our models, the classifiers will be not only 10-fold cross-validated but also applied on at least 20 new patients referring for PVI procedures (***Figure 2***).

We will validate the trained classifiers thanks to the decision surfaces, where every point of the Cartesian plane is inputted into the trained models and consequently classified as AF recurrence susceptible or not-susceptible.

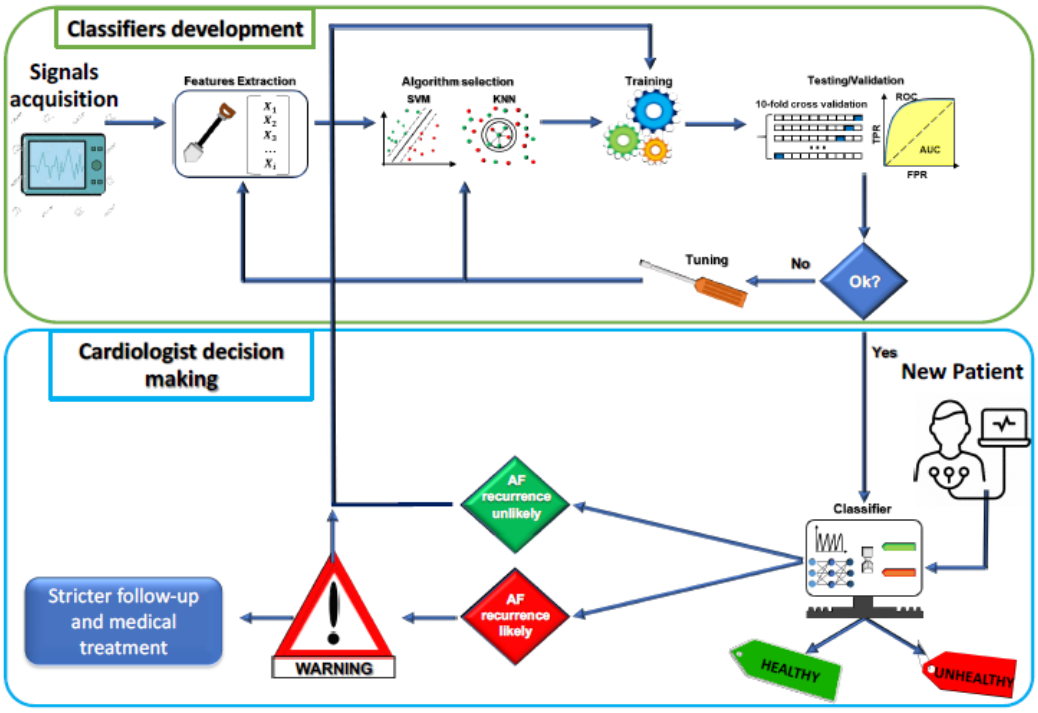


***Figure 2**. Supervised classifiers development (top) and robustness test (bottom)*

Project timeline:
1st year: collection of the above data at the time of PVI procedures, Next Generation Sequencing and analysis of Exome-Seq data.
2nd year: collection of AF recurrences at 1-year follow-up visit and training of the ML classifiers.
3rd year: validation of the algorithm on new patients referring for PVI.

Impact on the national health system:
AI showed promise both to perform expert-level tasks and to extend capabilities beyond healthcare professional cognition. Its applications to the electrophysiology field range from disease detection and diagnosis to novel characterization of the disease and prediction of outcomes, influencing the management of a large number of patients. The present project offers the opportunity to optimize the follow-up of patients undergoing PVI and will inevitably impact on the national health system in terms of workflow and costs, providing the possibility to choose the most effective and patient-specific disease oriented treatment. Moreover, the employment of AI in this project will inevitably reduce the number of observation necessary to achieve the desired endpoints (Deo R.C. 2015, Fralick and Colak et al 2019, Mukherjee and Cintra et al. 2020).

Innovazione e Trasferibilità
So far, there are no available data in literature integrating endocardial electrograms and sequencing data to the outcome of AF patients. This project represents an exploratory research aimed at investigating the possibility to macth electrical activity in the pulmonary veins with genetic variants and connect them with the success of PVI isolation procedures.
It will enable to set up the system that has the potential, after validation in bigger cohorts, of opening the way to new understandings in the field of cardiac electrophysiology, providing high amount of data that could implement the growth of several satellite projects.

Data, 30/07/2021

Il Legale Rappresentante

Si autorizza al trattamento dei dati ai sensi del d.lgs. 196/2003

Il Legale Rappresentante

Ministero della Salute – Direzione Generale della Ricerca e dell'Innovazione in Sanità

Rendiconto 5 per mille ANNO 2018
Contributo percepito € 1.578.095,22 In data 03/07/2020

Ente della Ricerca Sanitaria

Denominazione Ente: IRCCS Istituto Clinico Humanitas

Codice fiscale: 10125410158
Sede legale: via Manzoni 56 – 20089 – Rozzano (MI)
Indirizzo di posta elettronica dell'ente: amministratore.delegato@humanitas.it
HUMANITASMIRASOLE@LEGALMAIL.IT
Dati del rappresentante legale: Luciano Ravera

Titolo del progetto: ADDRESSING PERSONALIZED PRECISION MEDICINE FOR PATIENTS AFFECTED BY IMMUNE DISEASES THROUGH AI-BASED ANALYSIS OF CORTICOSTEROIDS USERS AND PREDICTION ON BIOLOGIC DRUGS PRESCRIPTION

Data di inizio progetto: 01/07/2021	**Data di fine progetto**: 30/06/2024
Fondi 5 per mille assegnati al progetto: € 77.000,00	**Di cui:** **Quota sostenuta entro l'anno di rendicontazione: € 0,00** **Quota accantonata, da sostenere, per progetti pluriennali (durata massima tre anni): € 77.000,00**

VOCI DI SPESA	Quota sostenuta entro l'anno di rendicontazione	Quota accantonata, da sostenere, per progetti pluriennali (durata massima tre anni)
Personale di ricerca (borsista, a contratto e di ruolo in quota parte)		50.000,00

Apparecchiature (ammortamento, canone di locazione/leasing)		
Materiale d'uso destinato alla ricerca (per laboratori di ricerca, acquisto farmaci ecc.)		10.000,00
Spese di organizzazione (manifestazioni e convegni, viaggi e missioni ecc.)		
Elaborazione dati		10.000,00
Spese amministrative		7.000,00
Altro (indicare quali)		
TOTALE	**0,00**	**77.000,00**

Data, 30/07/2021

Il Responsabile del Progetto
Prof. Carlo Francesco Selmi

Il Legale Rappresentante

Si autorizza al trattamento dei dati ai sensi del d.lgs. 196/2003

Il Legale Rappresentante

Ministero della Salute – Direzione Generale della Ricerca e dell'Innovazione in Sanità

Rendiconto 5 per mille ANNO 2018
Contributo percepito € 1.578.095,22 In data 03/07/2020

Ente della Ricerca Sanitaria
Denominazione Ente: IRCCS Istituto Clinico Humanitas
Codice fiscale: 10125410158
Sede legale: via Manzoni 56 – 20089 – Rozzano (MI)
Indirizzo di posta elettronica dell'ente: amministratore.delegato@humanitas.it
HUMANITASMIRASOLE@LEGALMAIL.IT
Dati del rappresentante legale: Luciano Ravera

Titolo del progetto: ADDRESSING PERSONALIZED PRECISION MEDICINE FOR PATIENTS AFFECTED BY IMMUNE DISEASES THROUGH AI-BASED ANALYSIS OF CORTICOSTEROIDS USERS AND PREDICTION ON BIOLOGIC DRUGS PRESCRIPTION

Relazione illustrativa

La medicina moderna e sostenibile implica un approccio proattivo e personalizzato nella gestione terapeutica della patologia infiammatoria cronica del paziente Immunocenter. In particolare la Medicina di precisione necessita dell'individuazione precoce di biomarkers di patologia e di risposta terapeutica. L'introduzione dei farmaci biologici nella pratica clinica richiede la ricerca parallela di biomarcatori per la selezione dei pazienti e la previsione dei risultati e monitoraggio, per consentire una scelta terapeutica adeguata anche in termine di cost effectiveness.
Mandatorio a livello clinico e di impatto socioeconomico diventa la riduzione del trattamento con corticosteroidi sistemici nei pazienti eleggibili a trattamento biologico elettivo.

BACKGROUND

Le attuali stime indicano che il 60% della popolazione mondiale soffre di multimorbidità ed è quindi colpito da due o più patologie croniche. Inoltre, si stima che la multimorbidità colpisca l'80% di coloro che hanno più di 85 anni; ciò suggerisce una correlazione positiva con l'aumento dell'età ed è invece inversamente correlata con la mortalità; tali tendenze si riscontrano principalmente nei paesi occidentali, ma sono ancora poco comprese.
Sebbene spesso usati come sinonimi, tale termine non deve essere confuso con la comorbidità, che dà la priorità a una condizione patologica a cui è affetto il paziente. Il concetto di comorbidità è comune in patologie autoimmuni o infiammatorie, come: l'artrite reumatoide, la broncopatia cronica ostruttiva (BPCO), l'asma, la psoriasi, l'artrite psoriasica e le malattie infiammatorie croniche intestinali (IBD). La medicina moderna e sostenibile implica un approccio proattivo e personalizzato nella gestione terapeutica della patologia infiammatoria cronica del paziente

Immunocenter. In particolare la Medicina di precisione necessita l'individuazione precoce di biomarkers di patologia e di risposta terapeutica. L'introduzione dei farmaci biologici nella pratica clinica richiede la ricerca parallela di biomarcatori per la selezione dei pazienti e la previsione dei risultati e monitoraggio, per consentire una scelta terapeutica adeguata anche in termine di efficientamento costi.
L'assegnazione di risorse associate al trattamento di pazienti affetti da multimorbidità ha dimostrato di essere una sfida importante nell'ambito dell'assistenza sanitaria. In particolare, la valutazione di come la multimorbidità interferisca con i trattamenti e i farmaci sulla condizione patologica principale fornirebbe una comprensione più profonda della multimorbidità in se e, allo stesso tempo, faciliterebbe la scelta di una migliore strategia di trattamento.
I principali enti governativi regolatori quali FDA, EMA e AIFA richiedono alle aziende ospedaliere di identificare i biomarcatori predittivi per sviluppare nuovi farmaci per le condizioni autoimmuni e infiammatorie croniche. Nell'ultimo decennio molta letteratura ha evidenziato inoltre i costi del mancato utilizzo di trattamenti biologici personalizzati a fronte della prescrizione di trattamenti corticosteroidei immunosoppressivi con aumento di patologie correlate all'utilizzo dei corticosteroidi stessi. In una recente pubblicazione dei dati del registro SANI (Severe Asthma Network Italy) ad esempio si è evidenziato che il costo annuo totale relativo agli eventi avversi correlati agli OCS sia stimabile 242,7 milioni di Euro per gli asmatici gravi. Sono stati dimostrati infatti spese incrementali di circa 110,6 milioni di Euro e 75,2 milioni di Euro rispetto a popolazione non asmatica e moderata,

POPOLAZIONE E CRITERI DI SELEZIONE DEI PAZIENTI
Utilizzeremo la Immuno DataBank sviluppata durante il 2020, che comprende i dati di circa 50.000 pazienti afferenti ai reparti di Allergologia, Reumatologia, Malattie infiammatorie dell'intestino e Dermatologia dall'inizio del 2017 al momento dell'analisi.
L'Immuno Databank permette un'analisi approfondita dei profili dei pazienti in termini di anamnesi, tipologia di prestazioni erogate ed eventuali prescrizioni farmacologiche. Di questo primo insieme saranno considerati esclusivamente i pazienti per i quali è stata erogata almeno una prestazione tra quelle considerate rilevanti ai fini dello studio. Inoltre sono considerati i pazienti a cui sono stati prescritti uno o più farmaci corticosteroidei in terapia. Tutti i dati afferenti alla popolazione considerata saranno anonimizzati prima di procedere con l'analisi e non sono previste validazioni individuali dei dati estratti mediante strumenti di Natural Language Processing (NLP).

DURATA E DISEGNO DELLO STUDIO
Lo studio sarà di tipo retrospettivo, monocentrico osservazionale. Si vogliono caratterizzare i diversi cluster di pazienti ai quali sono prescritti farmaci corticosteroidei per multiple combinazioni di stati di multimorbidità. Per la caratterizzazione si utilizzano i dati clinici presenti nei referti dei pazienti nel repository ICH, estratti tramite utilizzo di espressioni regolari (RegEx), incrociandoli con i loro trattamenti in uso per valutare le correlazioni tra farmaci utilizzati, differenti outcomes e condizioni patologiche di coesistenza. Particolare attenzione verrà indirizzata verso i cluster di pazienti a cui vengono prescritti un maggior numero di farmaci corticosteroidei e per i quali esiste una possibile sostituzione con farmaci biologici.
Ancora una volta sarà utile analizzare gli indici di multimorbidità già esistenti, come l'indice di Charlson e il punteggio di Elixhauser Comorbidity [2], per quantificarne la prevalenza all'interno dei dati disponibili. Sarà eseguita un'analisi esplorativa dei dati (EDA) per identificare i principali tratti dei gruppi considerati e permettere una clusterizzazione più efficace. Selezioneremo i cluster di pazienti che presentano prescrizione di corticosteroidi e che abbiano profili di multimorbidità simili. L'estrazione del dato, nell'eventualità non fosse già strutturato, sarà effettuata per mezzo di analisi testuale basata su tecniche di Natural Language Processing.
La prima fase si concluderà con l'identificazione dei cluster di interesse e la definizione dei percorsi terapeutici basati su corticosteroidi per ogni cluster identificato. Qualora emergessero interruzioni o modifiche nella terapia più comune, si indagherà sulle cause alla base delle stesse.
Per l'indagine si utilizzeranno diversi modelli di clustering, a partire da K-means, metodi gerarchici e bi-clustering. I cluster saranno poi utilizzati per la rappresentazione dei dati, al fine di migliorare l'identificazione di pattern rilevanti con l'utilizzo di metodi di feature reduction, come t-distributed stochastic neighbor embedding (T-SNE) o Principal Component Analysis (PCA).

La seconda fase dello studio invece, consisterà in un'analisi longitudinale delle condizioni cliniche e terapeutiche dei pazienti nei clusters. A questo proposito i dati saranno riorganizzati ed analizzati secondo un ordine temporale al fine di ottenere una modellizzazione del ciclo di cura del paziente.
L'analisi sarà volta a individuare dei pattern temporali condivisi da un numero rilevante di pazienti e permetterà di esplorare modelli di cura ottimali per la gestione di determinati percorsi terapeutici caratterizzati da set di comorbidità ad oggi trattati con corticosteroidi.
Per tale scopo verranno impiegati diversi modelli di differente complessità basati su metodologie di apprendimento supervisionato (ad esempio regressioni logistiche, alberi di classificazione).

OBIETTIVI PRINCIPALI

- Identificare e caratterizzare i clusters di pazienti afferenti all'Immunocenter trattati con farmaci corticosteroidei.
- Esplorare e definire le caratteristiche temporali che caratterizzano il trattamento dei pazienti appartenenti ai cluster identificati. Definire i trattamenti che presentano opportunità di miglioramento per mezzo di prescrizione di farmaci biologici nei pazienti con multimorbidità.

METODI STATISTICI ANALISI DEI DATI

L'infrastruttura di base sarà l'Immuno Databank, una struttura in cui sono contenuti informazioni strutturate estratte dai testi liberi scritti in seguito alle visite mediche dei pazienti esterni. Le informazioni qui contenute evidenziano la presenza di determinate patologie nei paragrafi di anamnesi e conclusioni, e la presenza di prescrizioni di farmaci corticosteroidei nel paragrafo dedicato alla terapia finale. Questa infrastruttura sarà arricchita con nuovi marker che saranno utilizzati per identificare la presenza di prescrizioni di farmaci biologici. La produzione di questi nuovi marker sarà effettuata con la stessa tecnica utilizzata per i precedenti, ossia tramite l'utilizzo di espressioni regolari. Tramite questi marker sarà individuata la coorte di interesse per l'analisi, composta da quelle visite in cui i pazienti che dopo una passata prescrizione di farmaci corticosteroidei hanno tratto o meno un beneficio da un cambio di terapia verso farmaci biologici. Identificate queste famiglie di pazienti si andranno ad utilizzare i marker di comorbidità e di diagnosi finale per identificare la situazione clinica del paziente. Queste informazioni, integrate con i parametri di laboratorio disponibili, saranno utilizzate per effettuare predizioni su un futuro giovamento da parte del paziente verso un cambio di terapia con farmaci biologici.
Per le previsioni saranno applicate metodologie di apprendimento supervisionato utilizzanti modelli di complessità crescente come regressione logistica, alberi di classificazioni e reti neurali. Questi modelli saranno sottoposti a diversi stage tra cui selezione di variabili più predittive, regolarizzazione e valutazione delle prestazioni. Per la selezione delle variabili sarà fatto affidamento sulla misura P-value, seguita da una regressione L1 (LASSO) per la conferma dei risultati. Il passo di regolarizzazione verrà effettuato tramite cross-validation in uno spazio di iperparametri randomizzati. Per la valutazione delle prestazioni verranno presi in considerazione le misure di sensibilità, specificità ed F-score.

Innovazione e trasferibilità

1. Definire i trattamenti che presentano opportunità di miglioramento per mezzo di prescrizione di farmaci biologici nei pazienti con multimorbidità.
2. Ridurre l'utilizzo dei trattamenti corticosteroidei sistemici al fine di l'impatto clinico diretto e indiretto socio conomico (sviluppo comorbidità metasteroidee) nei pazienti afferenti all'Immunocenter Humanitas affetta da patologie infiammatorie croniche.
3. Individuare biomarkers digitali predittivi di risposta terapeutica e clinica

Data, 30/07/2021

Il Legale Rappresentante

Si autorizza al trattamento dei dati ai sensi del d.lgs. 196/2003

Il Legale Rappresentante

Ministero della Salute – Direzione Generale della Ricerca e dell'Innovazione in Sanità

Rendiconto 5 per mille ANNO 2018
Contributo percepito € 1.578.095,22 In data 03/07/2020

Ente della Ricerca Sanitaria

Denominazione Ente: IRCCS Istituto Clinico Humanitas

Codice fiscale: 10125410158
Sede legale: via Manzoni 56 – 20089 – Rozzano (MI)
Indirizzo di posta elettronica dell'ente: amministratore.delegato@humanitas.it
HUMANITASMIRASOLE@LEGALMAIL.IT
Dati del rappresentante legale: Luciano Ravera

Titolo del progetto: Microbiota, metabolome and nutrition: an 'artificially intelligent' way to personalized nutrition

Data di inizio progetto: 01/07/2021	**Data di fine progetto**: 30/06/2024
Fondi 5 per mille assegnati al progetto: € 38.500,00	**Di cui:** **Quota sostenuta entro l'anno di rendicontazione: € 0,00** **Quota accantonata, da sostenere, per progetti pluriennali (durata massima tre anni): € 38.500,00**

VOCI DI SPESA	Quota sostenuta entro l'anno di rendicontazione	Quota accantonata, da sostenere, per progetti pluriennali (durata massima tre anni)
Personale di ricerca (borsista, a contratto e di ruolo in quota parte)		

Apparecchiature (ammortamento, canone di locazione/leasing)		
Materiale d'uso destinato alla ricerca (per laboratori di ricerca, acquisto farmaci ecc.)		35.000,00
Spese di organizzazione (manifestazioni e convegni, viaggi e missioni ecc.)		
Elaborazione dati		
Spese amministrative		3.500,00
Altro (indicare quali)		
TOTALE	**0,00**	**38.500,00**

Data, 30/07/2021

Il Responsabile del Progetto
Dott. Gionata Fiorino

Il Legale Rappresentante

Si autorizza al trattamento dei dati ai sensi del d.lgs. 196/2003

Il Legale Rappresentante

Ministero della Salute – Direzione Generale della Ricerca e dell'Innovazione in Sanità

Rendiconto 5 per mille ANNO 2018
Contributo percepito € 1.578.095,22 In data 03/07/2020

Ente della Ricerca Sanitaria
Denominazione Ente: IRCCS Istituto Clinico Humanitas
Codice fiscale: 10125410158
Sede legale: via Manzoni 56 – 20089 – Rozzano (MI)
Indirizzo di posta elettronica dell'ente: amministratore.delegato@humanitas.it
HUMANITASMIRASOLE@LEGALMAIL.IT
Dati del rappresentante legale: Luciano Ravera

Titolo del progetto: Microbiota, metabolome and nutrition: an 'artificially intelligent' way to personalized nutrition

Relazione illustrativa

Irritable bowel syndrome (IBS) is a food-related disorder that affects 10-20% of the population and is expected to increase(*1*). Patients with IBS present with dysbiosis, i.e. a change in microbiota composition, but it is not clear whether this is the consequence or the effect of a modified diet which allows patients to cope with the disease.
We propose a clinical trial with a dual intent:
1. To compare the microbiome, metabolome and energy conversion in IBS patients versus healthy individuals and to create datasets of integrated data 2. To reestablish the correct microbiome using an inclusive diet based on fermented products.
We will generate an algorithm that will allow us to identify through Artificial Intelligence (AI) the classes of microbes and metabolites that characterize the IBS population. We will also evaluate whether a diet based on fermented products (milk and beans) can restore the microbiota.

Background

The microbiota modulates the pathogenesis, progression and treatment of diseases, ranging from metabolic disorders to neurological diseases (*2-4*). Reshaping host-microbiota interactions through personalized nutrition is a new therapeutic avenue for both disease control and prevention. The composition of the microbiota depends on the diet(*5*). The more varied is the diet, the more varied is the microbiota(*6*). In several disorders the composition of the microbiota is modified, even though a clear identity of a 'healthy' microbiota is yet to be defined(*7*) and how to re-establish a correct microbiota composition is unknown. Furthermore, different microbiota compositions can lead to similar metabolic output. Therefore, it might be incorrect to analyze only the microbiota, and their metabolic output should also be considered. In this regard, we have recently published a study highlighting how the metabolome, differently from the microbiome, can clearly cluster and

distinguish babies with different nutrition regimens (breast feeding, standard formula and fermented formula) (*8*).
IBS is not life threatening, but it definitely has an impact on the quality of life with symptoms such as abdominal pain and cramping, diarrhea or constipation, changes in bowel movements, gas and bloating, food intolerance, fatigue and difficulty sleeping. Patients that suffer from IBS avoid some food items (Fermentable Oligo-, Diand Mono-Saccharides & Polyols, FODMAP) that make them feel uncomfortable. These food items include a series of products that deliver several important nutrients. FODMAP rich food items (primarily legumes and dairy products) ferment within the host causing bloating and intestinal tension. As a consequence, the individual can experience diarrhea and or constipation. To overcome these effects, IBS patients have a very poor diet (low in FODMAP) that excludes those nutrients that ferment in the gut during digestion (legumes, milk, certain fruits). This leads to nutrient deficiencies, including iron, calcium, vitamins. A low FODMAP nutrition leads to a vicious cycle (**Figure 1**).

The problem

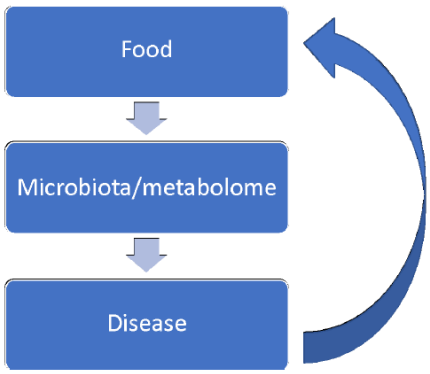


Figure 1: IBS Vicious Cycle: avoiding food that one is intolerant to drives the modification of the microbiota and its metabolome. As we become more and more intolerant to the food that we avoid or we eat (eg. rich in fat), the more the microbiota changes. This further impacts on the type of food we ingest and on the exacerbation of the disorder.

Here, we want to foster the notion of personalized nutrition focusing on the regeneration of a correct microbiota (*9*). Reconstitution of missing bacterial species will end the vicious cycle caused by dysbiosis. In this project, we will analyze the individual from a very wide angle including, the microbiome, the metabolome, the energy harvest potential and the capacity to reintroduce the correct food to reconstitute the microbiota. We do not want to offer simply a photography of the health status of the individual, but we also want to solve their nutrition and even more so, microbiome problems through supplementation with already fermented FODMAP food. Milk and beans will be fermented with Lactobacillus paracasei CNCM I-1390. This strain, besides allowing the degradation of the cellulose part forming the cuticle of the beans, which is responsible for the development of gas, also releases anti-inflammatory metabolites within the fermented product (*10*) which may participate in restoring intestinal immune homeostasis.

Study design
Patient population: IBS patients (D, C and M types) and healthy sex and matched controls.
Inclusion criteria
1. Patients of both sex and any race
2. Age ≥18 and ≤70.
3. IBS confirmed by Rome III diagnostic criteria
4. Colonoscopy within the past 10 years to rule out inflammatory bowel disease.
5. Willing to adhere to the proposed diet.
Exclusion Criteria:
1. Patient age <18 and > 70.

2. Diabetes (Type 1 or 2).
3. Lactose intolerance.
4. Pregnant or planning to become pregnant or is lactating.
5. History of HIV or hepatitis B or C.
6. Participation in investigational study within past 30 days.
7. Unstable cardiovascular or pulmonary disease, with change in treatment in last 30 days due to worsening disease condition.

Primary end point:
To compare the microbiome, metabolome and energy conversion in IBS patients versus healthy individuals and to create datasets of integrated data.
Secondary end point:
To reestablish the correct microbiome using an inclusive diet based on fermented food.
Enrollment:
Day 0 - All subjects will be exposed to a controlled diet for one day to avoid diet related microbiome and metabolome variability before sample collection.
Day 1 - A sample of the feces will be collected (baseline) and thoroughly analyzed for microbiome and metabolome composition. Physical characteristics and metabolic energy harvest and expenditure will also be evaluated (**Figure 2**).

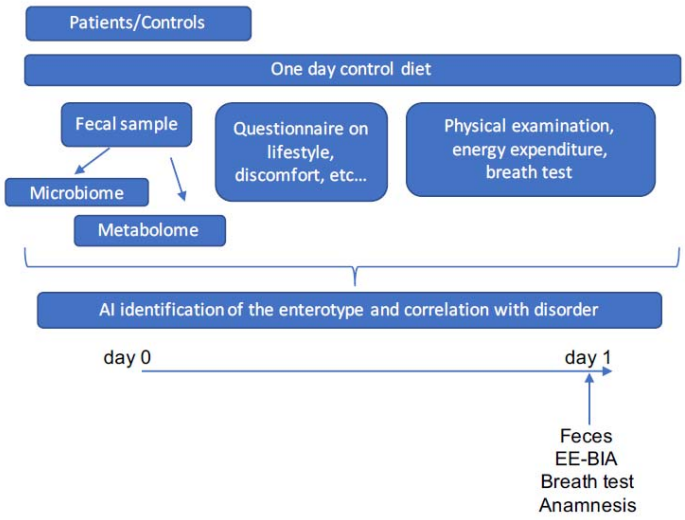


Figure 2: Enrolment to set the baseline.

Intervention:
Week 1 - Both IBS patients and healthy controls will receive a low FODMAP diet, which is recommended for IBS patients, for one week. This will allow to eliminate differences related to the low FODMAP diet between IBS and controls.
Weeks 2-4 - Starting from the second week both controls and IBS patients will receive low FODMAP diet supplemented with fermented beans (two weeks) and then with fermented milk. Patients' microbiome, metabolome, physical characteristics, metabolic energy harvest and expenditure will be evaluated in a longitudinal study for 5 weeks (**Figure 3**).

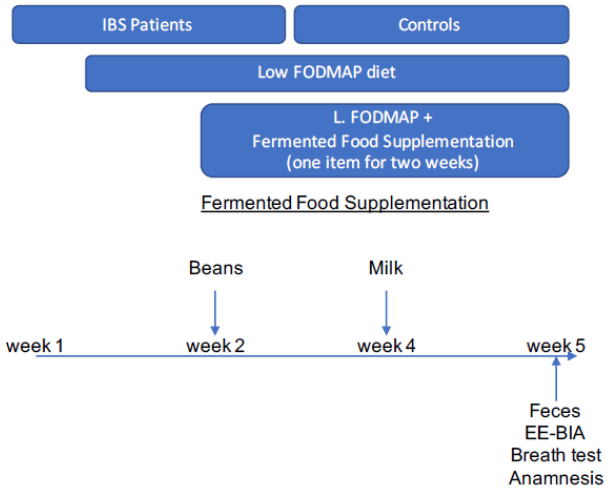


Figure 3: Outcome of the intervention.

The supplementation of two food items belonging to the category of FODMAP aims at reconstituting the microbiota and providing essential nutrients for the patients. Indeed, as previously mentioned, IBS patients have a very poor diet (low in FODMAP) that excludes those nutrients that ferment in the gut during digestion (legumes, milk, certain fruits). This leads to nutrient deficiencies, including iron, calcium, vitamins.

The novelty that we propose, and that is based on our experience, focuses on providing a fermented foods-based diet. By providing already fermented food, we will avoid food fermentation within the intestines of the patients thus reducing the IBS symptoms. Hence, the patients will be able to re-introduce food items that are normally avoided allowing rebalance of their microbiota and improvement of nutrient deficiency. In addition, L. paracasei fermented products will also deliver anti-inflammatory metabolites that help re-establish intestinal immune homeostasis.
Through this longitudinal study we will be able to evaluate the effect of the fermented food on the patient's microbiota, metabolome and grade of disorder, starting from the initial enterotype.

Impact on SSN
IBS has a significant impact on health care use and costs, which is heightened by the variability of IBS symptoms and the imprecise nature of diagnosing and treating IBS. In Italy, the estimated mean annual cost per patient for the national health system was €1761; while the mean annual indirect cost, due to productivity loss, was estimated to be €4905 per subject(*11*). Given the substantial economic burden for the patients, healthcare systems and society, IBS should be included among the priorities of the public health agenda. Furthermore, IBS is often linked to other disorders which appear later on in life, such as anxiety, depression and Alzheimer, making this intervention far more outreaching(*12*). We believe that this study could be a platform extendable to several other disorders like the metabolic syndrome, and neurodegenerative disorders.
Altogether, this project will allow us to:
• Identify complex enterotypes that have never been evaluated (metabolome, microbiome and metabolic parameters)
• Generate integrated datasets of IBS versus healthy individuals
• Offer a scheme of personalized nutrition
• Propose a revolutionary nutrition approach based on fermented food.

Data evaluation and generation of the algorithm/s
• Identification of a metabolome/microbiota associated to the different IBS types
• Identification of the effect that fermented food administration has on IBS disease progression and on patient's microbiome/metabolome

• Identification of a tool to propose a personalized nutrition

Method and Materials

We expect to retrieve metabolome, microbiome and metabolic related data from a pool of 40 patients and 40 healthy controls. Particularly data coming from 40/40 individuals' data (IBS patients/control) will be used to set the baseline (identification of enterotype and correlation with IBS disorder). Principal bottleneck of the study would be represented by a limited number of recruited enrolled subjects. To partially overcome this problem, we evaluate the possibility to enlarge our dataset by retrieving data from public databases and should this data provide interesting but not conclusive data we will design a new study to reach the objectives. Hence, it would be possible to properly train dedicated Neural Networks to recognize correlations among metabolome, microbiome and metabolic composition.

Preparation of fermented products Fermented milk and beans will be obtained using Lactobacillus paracasei CNCM I-1390 and following a methodology established in the laboratory (unpublished).

Fecal Microbiota Identification: Fecal samples will be stored at −80 °C until the DNA extraction with G NOME DNA isolation kit (MP Biomedicals) following the previously described protocol (*13*). Partial 16S rRNA gene sequences will be amplified from extracted DNA using primer pair Uni/Rev, which target the V3 region of the 16S rRNA gene sequence(*14*). 16S rRNA gene sequencing will be performed using a MiSeq (Illumina) within Humanitas Genomic Unit. Paired-end reads pairs will be assembled to reconstruct the complete amplicons. In order to calculate alpha diversity measures (Unifrac analysis), 16S rRNA Operational Taxonomic Units (OTUs) will be defined at ≥ 99 % sequence homology using uclust(*15*) and OTUs with less than 10 sequences will be filtered. All reads will be classified to the lowest possible taxonomic rank using QIIME(*16*) and a reference dataset from the SILVA database. The bacterial profile at phylum, family and genus will be reported as relative abundance. In detail, only taxa with relative abundance >0.5% (or 1 %) will be shown.

Metabolomic Analysis:

Metabolomics data analysis of fecal samples will be performed considered the partial least square discriminant (PLS-DA)(*17*) on normalized Internal Standard peak area(*18*). Classes separation will be performed by PLS-DA regression performed using the plsr function included in the R pls package(*19*). Using Metaboanalyst3.0, a weighted sum of squares from PLS will be calculated as Variable Importance in Projection (VIP). Variations with a VIP score of at least 1.5 were considered relevant. Fecal metabolite abundances will be tested for association to 16S levels using Speraman rank correlation. Pathway Impact and Metabolite Sets enrichment will be performed using Metaboanalyst3.0 (http://www.metaboanalyst.ca).

AI analysis will be performed as follows:

• stratification of subjects by means of hierarchical biclustering approaches(*20,21*) and identification of patient's clusters of relevance.

• individuation of dummy features through Principal Component Analysis (PCA), in order to describe data structure and represent the informationemerged from data after feature aggregations.

• By means of dedicated and properly trained NN (Supervised Learning), classify the patients (by IBS type) starting from metabolome/microbiota information and perform predictive evaluations on patients' conditions.

• Build an AI based tool able to propose a personalized nutrition to patients based on linear models (parametrized regression based) and non-linear (ie. multilayer perceptron (MLP) feedforward artificial neural network models.

Bibliografia

1. W. D. Chey, J. Kurlander, S. Eswaran, Irritable bowel syndrome: a clinical review. *JAMA* **313**, 949-958 (2015).

2. I. Cho, M. J. Blaser, The human microbiome: at the interface of health and disease. *Nat Rev Genet* **13**, 260-270 (2012).

3. E. Le Chatelier *et al.*, Richness of human gut microbiome correlates with metabolic markers. *Nature* **500**, 541-546 (2013).

4. R. E. Ley, P. J. Turnbaugh, S. Klein, J. I. Gordon, Microbial ecology: human gut microbes associated with obesity. *Nature* **444**, 1022-1023 (2006).

5. G. D. Wu *et al.*, Linking long-term dietary patterns with gut microbial enterotypes. *Science* **334**, 105-108 (2011).
6. M. Levy, E. Blacher, E. Elinav, Microbiome, metabolites and host immunity. *Curr Opin Microbiol* **35**, 8-15 (2017).
7. E. Rinninella *et al.*, What is the Healthy Gut Microbiota Composition? A Changing Ecosystem across Age, Environment, Diet, and Diseases. *Microorganisms* **7**, (2019).
8. P. Roggero *et al.*, Analysis of immune, microbiota and metabolome maturation in infants in a clinical trial of Lactobacillus paracasei CBA L74-fermented formula. *Nat Commun* **11**, 2703 (2020).
9. D. Zeevi *et al.*, Personalized Nutrition by Prediction of Glycemic Responses. *Cell* **163**, 1079-1094 (2015).
10. K. Tsilingiri *et al.*, Probiotic and postbiotic activity in health and disease: comparison on a novel polarised ex-vivo organ culture model. *Gut* **61**, 1007-1015 (2012).
11. L. Neri *et al.*, Constipation severity is associated with productivity losses and healthcare utilization in patients with chronic constipation. *United European Gastroenterol J* **2**, 138-147 (2014).
12. C. H. Chen, C. L. Lin, C. H. Kao, Irritable Bowel Syndrome Is Associated with an Increased Risk of Dementia: A Nationwide Population-Based Study. *PLoS One* **11**, e0144589 (2016).
13. J. P. Furet *et al.*, Comparative assessment of human and farm animal faecal microbiota using real-time quantitative PCR. *FEMS Microbiol Ecol* **68**, 351-362 (2009).
14. C. Milani *et al.*, Assessing the fecal microbiota: an optimized ion torrent 16S rRNA gene-based analysis protocol. *PLoS One* **8**, e68739 (2013).
15. R. C. Edgar, Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26**, 2460-2461 (2010).
16. J. G. Caporaso *et al.*, QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* **7**, 335-336 (2010).
17. S. Wold, M. Sjöström, L. Eriksson, PLS-regression: a basic tool of chemometrics. *Chemometrics and Intelligent Laboratory Systems* **58**, 109-130 (2001).
18. M. Sysi-Aho, M. Katajamaa, L. Yetukuri, M. Oresic, Normalization method for metabolomics data using optimal selection of multiple internal standards. *BMC Bioinformatics* **8**, 93 (2007).
19. B.-H. Mevik, R. Wehrens, The pls Package: Principal Component and Partial Least Squares Regression in R. *Journal of Statistical Software; Vol 1, Issue 2 (2007)*, (2007).
20. D. Müllner, Modern hierarchical, agglomerative clustering algorithms.(2011).
21. Z. Bar-Joseph, D. K. Gifford, T. S. Jaakkola, Fast optimal leaf ordering for hierarchical clustering. *Bioinformatics* **17 Suppl 1**, S22-29 (2001).

Data, 30/07/2021

Il Legale Rappresentante

Si autorizza al trattamento dei dati ai sensi del d.lgs. 196/2003

Il Legale Rappresentante

Ministero della Salute – Direzione Generale della Ricerca e dell'Innovazione in Sanità

Rendiconto 5 per mille ANNO 2018
Contributo percepito € 1.578.095,22 In data 03/07/2020

Ente della Ricerca Sanitaria

Denominazione Ente: IRCCS Istituto Clinico Humanitas

Codice fiscale: 10125410158
Sede legale: via Manzoni 56 – 20089 – Rozzano (MI)
Indirizzo di posta elettronica dell'ente: amministratore.delegato@humanitas.it
HUMANITASMIRASOLE@LEGALMAIL.IT
Dati del rappresentante legale: Luciano Ravera

Titolo del progetto: Identification of Artificial Intelligence based biomarkers to predict HCC response to medical treatment

Data di inizio progetto: 01/07/2021	**Data di fine progetto**: 30/06/2024
Fondi 5 per mille assegnati al progetto: € 375.628,00	**Di cui:** **Quota sostenuta entro l'anno di rendicontazione: € 0,00** **Quota accantonata, da sostenere, per progetti pluriennali (durata massima tre anni): € 375.628,00**

VOCI DI SPESA	Quota sostenuta entro l'anno di rendicontazione	Quota accantonata, da sostenere, per progetti pluriennali (durata massima tre anni)
Personale di ricerca (borsista, a contratto e di ruolo in quota parte)		241.480,00

Apparecchiature (ammortamento, canone di locazione/leasing)		
Materiale d'uso destinato alla ricerca (per laboratori di ricerca, acquisto farmaci ecc.)		
Spese di organizzazione (manifestazioni e convegni, viaggi e missioni ecc.)		
Elaborazione dati		
Spese amministrative		34,148.00
Altro (storage e servizio di mappatura dati)		100.000,00
TOTALE	**0,00**	**375.628,00**

Data, 30/07/2021

Il Responsabile del Progetto
Prof. Luigi Terracciano

Il Legale Rappresentante

Si autorizza al trattamento dei dati ai sensi del d.lgs. 196/2003

Il Legale Rappresentante

Ministero della Salute – Direzione Generale della Ricerca e dell'Innovazione in Sanità

Rendiconto 5 per mille ANNO 2018
Contributo percepito € 1.578.095,22 In data 03/07/2020

Ente della Ricerca Sanitaria
Denominazione Ente: IRCCS Istituto Clinico Humanitas
Codice fiscale: 10125410158
Sede legale: via Manzoni 56 – 20089 – Rozzano (MI)
Indirizzo di posta elettronica dell'ente: amministratore.delegato@humanitas.it
HUMANITASMIRASOLE@LEGALMAIL.IT
Dati del rappresentante legale: Luciano Ravera

Titolo del progetto: Identification of Artificial Intelligence based biomarkers to predict HCC response to medical treatment

Relazione illustrativa

Hepatocellular carcinoma (HCC) is the most frequent primary liver cancer and the second cause of cancer death worldwide. More than 70% of patients need medical treatments but their benefit is minimal and achieved in few patients, due to the absence of markers of response.
Artificial intelligence (AI) algorithm can generate and learn the optimal feature to best separate the categories of interest without receiving any preexisting assumptions. Recently, AI has been used to discover information from digital images of pathological process and use them to make specific prediction, including treatment response.

Aim of this study is twofold:
1. to detect histological features identified by AI predicting the response to HCC medical treatment from H&E stained slides of advanced HCC
2. to compare them toward clinical features and to select that one performing best to be used as AI-based biomarker.

Working in collaboration with the Unit of Oncology we will:
- collect needle liver biopsies of 250 patients treated with Sorafenib (SOR) or Immune Checkpoint Inhibitors (ICI);
-separate patients into 4 categories (SOR/ICI and responders vs. nonresponders);
-populate each category with tiles obtained from digital slide of every patients.

Working in collaboration with the Artificial intelligence Center we will:
-set up an AI solution which fits with our aim;

-train the AI solution with 70% of tiles;
- test the AI-based features in the remaining 30% cases and compare their performance against clinical parameters.

BACKGROUND

Hepatocellular carcinoma (HCC) is the most frequent primary liver cancer and the second cause of cancer death worldwide (1). HCC incidence is increasing due to the rising of obesity/metabolic syndrome (2). More than 70% of patients need medical treatments (3). These latter were represented, until recently, by three anti-angiogenic systemic agents (Sorafenib, Lenvantinib, Regorafenib). In addition to the above-mentioned drugs, the combined use of atezolizumab (an immune check point inhibitor) and bevacizumab (an anti-angiogenic) was recently approved by FDA showing longer overall survival and progression-free survival in advanced HCC (4). This is in keeping with studies pointing out a significant interaction between endothelial cells and immune infiltrate (5).

However, medical treatment benefit is minimal and achieved in few patients. These frustrating result reveals, despite the undeniable advances in HCC biology (6-10), that we are far to identify markers of response. Taking this background into consideration our group has recently shifted the focus of its research from HCC cells to the interaction between HCC cells and non-tumor cell component (the so called tumoral micro environment, TME). We are currently investigating the role of immune (LMT) and vascular (LDT) TME, and both these researches had been supported by IG-AIRC grants.

This study aims to experiment an alternative approach using Artificial Intelligence algorithm to identify novel biomarkers embedded within tissue but not detected by routine histological analysis.

The term Artificial intelligence (AI) refers to the branch of computer science in which the machines attempt to emulate the human intelligence (11,12). Machine Learning (ML) is a branch of AI where machines are fed with data and learn by them. Deep learning (DL) is a type of ML based on the use of artificial representation of human neural architecture (13,14). DL algorithm can generate and learn the optimal feature to best separate the categories of interest without receiving any pre-existing assumptions. In particular, DL algorithms have been recently used to discover information from whole slide images (WSI) of pathological process and use them to make specific prediction, including treatment response (15,16), mutational profile (17,18), or patient survival (19).

AIM

To explore the existence of any histological specific feature detected by AI (AI-based biomarker) predicting the response to HCC medical treatment.

MATERIAL and METHODS

WP 1: Case selection and Slide digitalization The research will be conducted in collaboration with the group of Oncologists (prof. L. Rimassa and A. Santoro) active within the Liver Clinical Network of Humanitas. This network has a strong experience in Translational research and extensive access to human liver samples. We will focus our study on two series of patients with advanced HCC diagnosed on a liver biopsy, treated at Humanitas Clinical and Research centre, with complete follow-up data:

- 150 patients treated with Sorafenib (SOR);
- 100 patients treated with immune checkpoint inhibitors (ICI).

For each series, we will

a) Separate patients into “good” and “bad” responders according to clinical criteria (mRECIST; TTR)
b) Select one H/E for each patient;
c) Digitalize H/E and create a WSI (whole slide image);
d) Divide each WSI into several tiles;
e) Assign tiles to four specific categories established according to treatment and response: SOR good, SOR bad, ICI good and ICI bad.

Expected result WP1:
1) to generate at least 2000 different tiles;

2) to assign to every category at least 250 tiles.
Timing: 3 months.

WP 2: Neural network selection and training
a) Perform an Exploratory Data Analysis (EDA) to assess the profile and relevant features of the obtained data.
b) Through a Principal Component Analysis (PCA) evaluate the presence of redundant features and finalize a feature reduction process.
c) Build and train (training set, about 70% of total data) several common classifiers based on Neural Networks. Then assess their performances four our specific task and select the best Network in terms of accuracy, sensitivity and F1-score.

Expected results WP2:
1) to develop a neural network able to ingest and analyse clinical data and biopsy slices tiles an to discover correlation patterns among patients clusters.
2) to identify a list of AI features associated with “good” and “bad” response in SOR and ICI cohort.
Timing: 12 months.

WP 3: AI biomarkers validation
a) We will use the remaining 30% of tiles generated in WP1 to test the performance of AI features to predict response to SOR and ICI
b) We will compare the performance of best AI features to classical clinical parameters.

Expected results of WP3:
1) To identify a set of AI features usable to predict response to medical therapy in combination with classical clinical parameters to improve clinical response.
Timing: 3 months.

IMPACT ON SSN
A widespread use of Artificial Intelligence is a milestone of Digital Health. The introduction of AI within the workflow of a Pathology Department is likely to reduce the workload and to help standardize the otherwise subjective diagnosis that can lead to suboptimal treatment of patients.
Moreover promising high-level scientific researches recently used AI algorithms to discover information ignored by human review of image and use them to make specific prediction. These AI-based biomarkers will help to discover new perspectives in human biology, and progress on personalized diagnostics and patient care.

Riferimenti bibliografici
1.Ferlay J, et al. Int J Cancer. 2015; 136:E359-86.
2.El-Serag HB. N Engl J Med. 2011; 365:1118-27.
3.Bruix J, et al. Gastroenterology. 2016; 150:835-53.
4.Finn RS et al. N Engl J Med. 2020.
5.Pinyol R. Clin Cancer Res.
6.Boyault S, et al. Hepatology. 2007; 45:42-52.
7.Hoshida Y et al. Cancer Res. 2009; 69:7385-92.
8.Zucman-Rossi J, et al. Gastroenterology. 2015; 149:1226-1239.
9.Cancer Genome Atlas Research Network.. Cell. 2017;
10.Harding JJ, et al. Clin Cancer Res. 2019; 25:2116-2126.
11.McCarthy J, et al. 1955
12.McCarthy JJ, et al. 1959
13.Yao X. Proc. IEEE 87 1423–1447 (1999).
14.Haykin S Neural Networks (Prentice Hall, 1994).
15.Wang X et al J. Clin. Oncol. 36 12061–12061 (2018).
16. Barrera C et al. J. Clin. Oncol. 36, 12115–12115 (2018).
17.Coudray N et al., Nat. Med. 24 (2018) 1559–1567.
18.Kather JN, et al. Nat. Med. 2019, 25, 1054–1056.
19.Skrede OJ, et al. Lancet. 395 (2020) 350–360

Innovazione e trasferibilità
Artificial Intelligence (AI) covers a broad choice of solutions to collect, organize, interpret and use clinical data. Innovative technologies, such as AI, are adopted to the extent that they a) solve a problem users readily recognize; b) fit within existing physical environment and processes; c) influence analogs to telegraph what it does and how I work it.
Innovative technologies with penetration >15% have significant potential to become mainstream. In a recent survey for adopting digital clinical tools conducted by the American Medical Association, 31% of physicians stated to expect a definite advantage in the ability to care for patients. One of the most relevant aspect prompting physicians do adopt digital healthcare tools was improving diagnostic ability. Our proposal aims to develop such a highly innovative technique with a potential revolutionary impact on the management of HCC patients and, accordingly, a high potential of transferability.

Data, 30/07/2021

Il Legale Rappresentante

Si autorizza al trattamento dei dati ai sensi del d.lgs. 196/2003

Il Legale Rappresentante

Ministero della Salute – Direzione Generale della Ricerca e dell'Innovazione in Sanità

Rendiconto 5 per mille ANNO 2018
Contributo percepito € 1.578.095,22 In data 03/07/2020

Ente della Ricerca Sanitaria

Denominazione Ente: IRCCS Istituto Clinico Humanitas

Codice fiscale: 10125410158
Sede legale: via Manzoni 56 – 20089 – Rozzano (MI)
Indirizzo di posta elettronica dell'ente: amministratore.delegato@humanitas.it
HUMANITASMIRASOLE@LEGALMAIL.IT
Dati del rappresentante legale: Luciano Ravera

Titolo del progetto: Machine-learning ready Endoscopy-Data Bank for personalized and precision medicine

Data di inizio progetto: 01/07/2021	**Data di fine progetto**: 30/06/2024
Fondi 5 per mille assegnati al progetto: € 377.467,22	**Di cui:** **Quota sostenuta entro l'anno di rendicontazione: € 0,00** **Quota accantonata, da sostenere, per progetti pluriennali (durata massima tre anni): € 377.467,22**

VOCI DI SPESA	Quota sostenuta entro l'anno di rendicontazione	Quota accantonata, da sostenere, per progetti pluriennali (durata massima tre anni)
Personale di ricerca (borsista, a contratto e di ruolo in quota parte)		220,820.00

Apparecchiature (ammortamento, canone di locazione/leasing)		
Materiale d'uso destinato alla ricerca (per laboratori di ricerca, acquisto farmaci ecc.)		
Spese di organizzazione (manifestazioni e convegni, viaggi e missioni ecc.)		
Elaborazione dati		
Spese amministrative		34,315.20
Altro (storage e servizi di mappatura dati)		122,332.02
TOTALE	**0,00**	**377.467,22**

Data, 30/07/2021

Il Responsabile del Progetto
Prof. Alessandro Repici

Il Legale Rappresentante

Si autorizza al trattamento dei dati ai sensi del d.lgs. 196/2003

Il Legale Rappresentante

Ministero della Salute – Direzione Generale della Ricerca e dell'Innovazione in Sanità

Rendiconto 5 per mille ANNO 2018
Contributo percepito € 1.578.095,22 In data 03/07/2020

Ente della Ricerca Sanitaria
Denominazione Ente: IRCCS Istituto Clinico Humanitas
Codice fiscale: 10125410158
Sede legale: via Manzoni 56 – 20089 – Rozzano (MI)
Indirizzo di posta elettronica dell'ente: amministratore.delegato@humanitas.it
HUMANITASMIRASOLE@LEGALMAIL.IT
Dati del rappresentante legale: Luciano Ravera

Titolo del progetto: Machine-learning ready Endoscopy-Data Bank for personalized and precision medicine

Relazione illustrativa

Thanks to new digitalization technologies in healthcare, it is now possible to correlate data in ways that hardly could be achieved a few decades ago. In particular, thanks to the ability to digitise tissue samples with high fidelity, algorithms and models now have the opportunity to infer meaningful information from those. The ramifications of the technologies are many, to the support in support clinicians in the analysis, provide early diagnoses and provide new research opportunities. Examples are multi-scale multi-modal data correlation to predict outcomes from data originating at different levels of the patient physiology or AI based detection and classification of pathologies. In order to achieve this task, it is important to provide a solid base where clinicians and researchers can test their hypothesis, but many challenges arise: Videos deriving from endoscopic examinations may have different durations, formats, artifacts and may result complex to handle, Data compression may impact the information and storage requirements amongst others, Training AI algorithms may be difficult because of large amounts of image annotations required. Even if some datasets with annotations are already available, there is still a significant need of detecting pathology predictors in digital recording and documentation and precisely treating the area. Several researches highlight the impact of these technologies . A solid database would translate in a strong.

An endoscopy is an examination consisting of the insertion of a long, flexible fiber-optic tube, into one of the cavities of the body. The endoscope is a medical device made of fiber-optic, with a light and camera at the end, that let the MD see inside the patient's body. This kind of examination is useful and necessary for diagnosing or preventing several pathologies, from polyps to cancer related to the human digestive system. In case of doubt or anything unusual, during an examination, the doctor may use a surgical tool in order to remove and collect a sample of tissue to

study. The endoscope is connected to a display and the video is streamed live during the examination. The availability of the results and the video even after the exam could be helpful for research purposes and other studies in gastroenterology. Specifically, data could be used both for new findings or the development of algorithms and approaches that could be helpful for new findings, personalised medicine or predictive models.

The project will be structured in different phases:
Phase 1 - Numerosity and cohort assessment:
Initially, the metadata of all samples collected in ICH from 2015 afterwards will be used to provide an outcome baseline to estimate the processing capacity of the EHDB (Endoscopy Health Data Bank). In particolar trends, diagnoses and demographics of patients will be clustered to identify trends in the population with clinical guidance. The analysis will provide an indication of the data population of the final data base.
Phase 2 - Rate of digital conversion
Secondly, once several clusters of patients are chosen, an assessment on the monthly-rate at which samples can be collected will be used to estimate timings to populate the EHDB. The collection-rate will turn important in the selection of the architecture of the database and its evolution.
Phase 3 - Tags identification
Identification of the relevant information to correlate with the videos, such as, but not only to, patient information, diagnoses and information about the anatomical region. Moreover the format itself will be investigated, perhaps the DICOM communication format may yield to good results, being heavily utilized.
Phase 4 - Pilot Phase
An anonymized pilot will be created on a cloud infrastructure. Cloud tools provide elasticity and scalability needed to fine tune the shape of the database. Machine learning and statistical elements will be used to identify the correct parameters of the storage capacity, rate of in-bound data, and processing power to handle the samples. This phase will last until the cumulative samples produced from phase 2 will reach numerosities in the order of the hundreds. Around these numbers it will be possible to provide solid estimates on the evolution of the dataset as well as its performance on a large scale. In this phase different communication protocols and data models will be tested, to find the best trade-off between user accessibility and data processing potential.
Phase 5 - On-premise solution
Eventually, an on-premise solution may be considered depending on the stability of the database in phase 4.

The database will enable the possibility to correlate on a large scale information captured at different levels of the patient anatomy and physiology, so to capture uncovered correlations that ultimately benefit patients treatment.

Data, 30/07/2021

Il Legale Rappresentante

Si autorizza al trattamento dei dati ai sensi del d.lgs. 196/2003

Il Legale Rappresentante

Ministero della Salute – Direzione Generale della Ricerca e dell'Innovazione in Sanità

Rendiconto 5 per mille ANNO 2018
Contributo percepito € 1.578.095,22 In data 03/07/2020

Ente della Ricerca Sanitaria

Denominazione Ente: IRCCS Istituto Clinico Humanitas

Codice fiscale: 10125410158
Sede legale: via Manzoni 56 – 20089 – Rozzano (MI)
Indirizzo di posta elettronica dell'ente: amministratore.delegato@humanitas.it
HUMANITASMIRASOLE@LEGALMAIL.IT
Dati del rappresentante legale: Luciano Ravera

Titolo del progetto: Assessing early predictors of mortality and treatment strategies for life-threatening, time-dependent diseases in the Emergency Department using the Artificial Intelligence approach

Data di inizio progetto: 01/07/2021	**Data di fine progetto**: 30/06/2024
Fondi 5 per mille assegnati al progetto: € 55.000,00	**Di cui:** **Quota sostenuta entro l'anno di rendicontazione: € 0,00** **Quota accantonata, da sostenere, per progetti pluriennali (durata massima tre anni): € 55.000,00**

VOCI DI SPESA	Quota sostenuta entro l'anno di rendicontazione	Quota accantonata, da sostenere, per progetti pluriennali (durata massima tre anni)
Personale di ricerca (borsista, a contratto e di ruolo in quota parte)		30.000,00

Apparecchiature (ammortamento, canone di locazione/leasing)		
Materiale d'uso destinato alla ricerca (per laboratori di ricerca, acquisto farmaci ecc.)		20.000,00
Spese di organizzazione (manifestazioni e convegni, viaggi e missioni ecc.)		
Elaborazione dati		
Spese amministrative		5.000,00
Altro (indicare quali)		
TOTALE	**0,00**	**55.000,00**

Data, 30/07/2021

Il Responsabile del Progetto
Dott. Antonio Voza

Il Legale Rappresentante

Si autorizza al trattamento dei dati ai sensi del d.lgs. 196/2003

Il Legale Rappresentante

Ministero della Salute – Direzione Generale della Ricerca e dell'Innovazione in Sanità

Rendiconto 5 per mille ANNO 2018
Contributo percepito € 1.578.095,22 In data 03/07/2020

Ente della Ricerca Sanitaria

Denominazione Ente: IRCCS Istituto Clinico Humanitas

Codice fiscale: 10125410158
Sede legale: via Manzoni 56 – 20089 – Rozzano (MI)
Indirizzo di posta elettronica dell'ente: amministratore.delegato@humanitas.it
HUMANITASMIRASOLE@LEGALMAIL.IT
Dati del rappresentante legale: Luciano Ravera

Titolo del progetto: Assessing early predictors of mortality and treatment strategies for life-threatening, time-dependent diseases in the Emergency Department using the Artificial Intelligence approach

Relazione illustrativa

Precision Medicine is defined as: "The tailoring of medical treatment to the individual characteristics of each patient…to classify individuals into subpopulations that differ in their susceptibility to a particular disease or their response to a specific treatment. Preventative or therapeutic interventions can then be concentrated on those who will benefit, sparing expense and side effects for those who will not". As the definition suggests, the power of precision medicine lies in its ability to guide health care decisions toward the most effective treatment for a given patient, and thus, improve care quality, while reducing the need for unnecessary diagnostic testing and therapies. Health care has changed since the decline in mortality caused by infectious diseases, chronic and non-contagious diseases as well as time-dependent diseases (such as Stroke, Myocardial Infarction and Sepsis) with a direct impact on the cost of public health and individual health care. We must now transition from traditional reactive medicine based on symptoms, diagnosis and treatment to a system that targets the disease before it occurs and, if it cannot be avoided, treats the disease in a personalized manner. Precision Medicine is that new way of thinking about medicine. To emphasize the importance of disease prevention, a critical component of precision medicine can be referred to as precision health, which is defined as the use of all available information pertaining to specific subjects (including family history, individual genetic and other biometric information, and exposures to risk factors for developing or exacerbating disease), as well as features of their environments, to sustain and enhance health and prevent the development of disease.

Background
Precision Medicine, Personalized Medicine, and Individualized Medicine (and even older terms, such as pharmacogenomics or pharmacogenetics) have often been used over the last years to mean the use of biological markers (obtained by ordering a few supplementary tests by physicians) that may indicate particular genetic characteristics of the different affected individuals and of the various diseases that affect them (such as genetic, oncologic, or infectious diseases). These biomarkers may be used to indicate the most adequate treatment and follow-up for these disorders, which can be different for each individual. The term 'Precision Medicine' is currently preferred, since besides Genetics, it also involves lifestyle and environmental characteristics. This topic has been increasingly discussed by the medical community over these last years. Many scientific events on the theme have been held all over the world to discuss the use of new biomarkers, and to improve diagnosis and treatment of diverse diseases.
The combination of technological advances in molecular chemistry, assay automation, data storage and transmission, and computational mathematics now allow quantification of hundreds (e.g., metabolomics) to tens of thousands (e.g., transcriptomics) of analytes from a single patient specimen. In the case of genomics, billions of data points from a single patient sample can be generated. Some common techniques for systematic molecular phenotyping include genomics, genome-wide association studies (GWAS), epigenetics, transcriptomics, proteomics, metabolomics, and microbiomics, among others.
The purpose of the current project is to introduce Precision Medicine in the Emergency Department in order to assess early clinical predictors of mortality and individual-based treatments for time-dependent, life-threatening conditions.

Methods
We aim to develop predictive algorithms for patients outcome (in terms of mortality) affected by time-dependent diseases (such as Stroke, Myocardial infarction and Sepsis) using clinical parameters, instrumental parameters and biomarkers collected at the time of ED admission.
Fundamental elements required in order to develop a Precision Medicine Project are:
1) clinical expertise;
2) clinical data and records;
3) a dedicated testing laboratory.
The testing laboratory will require two separated parts:
i) "Wet part" which involves a pre-processing time dedicated to the biobanking of the biological samples;
ii) "Dry part" which involves a post-processing time dedicated to the extraction of the material.

Data will be analyzed using deep machine learning methods in order to find early predictors of mortality and new therapeutic strategies for time-dependent diseases.

Impact on the NHS
Genomics and other-omics knowledge and technologies are transforming the way healthcare can be delivered through greater understanding of disease detection and therapeutics. Responsible decision-making in the climate of escalating healthcare costs is required to ensure that precision medicine can be properly tested on a scale to determine if this approach will lead to better patient outcomes. Additionally, traditional decision-making paradigms must be agile to the precision medicine approach to ensure knowledge and discovery can be translated effectively and efficiently for better patient care.

Riferimenti bibliografici
1. National Research Council, Committee on A Framework for Developing a New Taxonomy of Disease. Toward Precision Medicine: Building a Knowledge Network for Biomedical Research and a New Taxonomy of Disease. National Academies Press; Washington DC: 2011.
2. Gameiro GR, Sinkunas V, Liguori GR, Auler-Júnior JOC. Precision Medicine: Changing the way we think about healthcare. Clinics (Sao Paulo). 2018;73:e723. Published 2018 Dec 3. doi:10.6061/clinics/2017/e723
3. Galli SJ. Toward precision medicine and health: Opportunities and challenges in allergic diseases. J Allergy Clin Immunol. 2016 May;137(5):1289-300. doi: 10.1016/j.jaci.2016.03.006. PMID: 27155026; PMCID: PMC4872702.

4. Hodson R. Precision medicine. S49Nature. 2016;537(7619)
5. Pinho JRR. Precision Medicine. Einstein (Sao Paulo). 2017;15(1):VII-X. doi:10.1590/S1679-45082017ED4016
6. Limkakeng AT Jr, Monte AA, Kabrhel C, et al. Systematic Molecular Phenotyping: A Path Toward Precision Emergency Medicine?. Acad Emerg Med. 2016;23(10):1097-1106. doi:10.1111/acem.13027
7. Bilkey GA, Burns BL, Coles EP, Mahede T, Baynam G, Nowak KJ. Optimizing Precision Medicine for Public Health. Front Public Health. 2019;7:42. Published 2019 Mar 7. doi:10.3389/fpubh.2019.00042

Data, 30/07/2021

Il Legale Rappresentante

Si autorizza al trattamento dei dati ai sensi del d.lgs. 196/2003

Il Legale Rappresentante