Department of Services Area SC
Transfusion Medicine Tel. 02
5503.4687
E-mail:daniele.prati@policlinico.mi.it

- Director: Dr. Daniele Prati

Gene-environment interaction study
in alcohol-related hepatocellular carcinoma

(European Project Title HORIZON-MISS-2021-CANCER- 02-03: "Understanding Gene ENvironment Interaction in ALcohol-related hepatocellular carcinoma")

Acronym: GENIAL

Version 2 of May 27, 2024

PROMOTER: IRCCS Ca' Granda Foundation, Maggiore Hospital, Policlinico, via Sforza 29, 20122 Milan, Italy,

COORDINATING CENTER: SC Transfusion Medicine
IRCCS Ca' Granda Foundation Maggiore Hospital Polyclinic, via Sforza 29, 20122 Milan Italy

PRINCIPAL INVESTIGATOR: Dr Serena Pelusi
Signature

DOCUMENT TYPE: non-pharmacological interventional study protocol

PRIVACY STATEMENT
All information presented in this document will be considered confidential and will remain the exclusive property of Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico. The use of such confidential information must be limited to the recipient for the agreed purpose and must not be disclosed, published or otherwise communicated to unauthorized persons, for any reason, in any form without the prior written consent of FondazioneIRCCSCa'GrandaOspedaleMaggiorePoliclinico.

Department of Services Area SC
Transfusion Medicine Tel. 02 5503.4687 - Director: Dr. Daniele Prati
E-mail:daniele.prati@policlinico.mi.it

List of abbreviations:

CI: informed consent
EC: Ethics Committee
CRF: case report form
NAFLD: non-alcoholic fatty liver disease.
ALD: alcoholic liver disease
HCC: hepatocellular carcinoma
EU: European Union
BMI: body mass index CA:
abdominal circumference AI:
artificial intelligence
CAP: controlled attenuation parameter
LSM: liver stiffness measurement IL-32:
interleukin 32
Pro-C3: pro-collagen 3 WES:
whole exome sequencing PRS:
polygenic-risk score
SNP: single nucleotide polymorphism

Responsibilities (role of the promoter and collaborators)

The Sponsor of the study will be the IRCCS Foundation Ca' Granda Ospedale Maggiore Policlinico of Milan.
The role of coordinating center of the study will be covered by the SC Transfusion Medicine of the IRCCS Foundation Ca' Granda Ospedale Maggiore Policlinico of Milan.
The Principal Investigator of the study will be Dr. Serena Pelusi who will be responsible for patient enrollment, collection of clinical and genetic data and follow-up visits.

Internal Collaborations

Operational Unit	Participant Name	Role and functions in the study
Complex Structure Transfusion Medicine IRCCS Foundation Ca' Granda Hospital Maggiore Policlinico, Milan	Dr. Serena Pelusi (principal investigator) Dr Daniele Prati Dr. Alessandro Cherubini Dr. Luisa Ronzoni Dr. Vittoria Moretti	Study coordination Genetic characterization of patients Clinical characterization of patients

External collaborations

Operational Unit	Participant Name	Role and functions in the study
Free University of Brussels_ULB	Prof. Eric Trepo	European Coordinator
University of Milan_UMIL	Professor Luca Valenti	Genetic characterization of patients
Technical University of Dresden_TUD	Dr. Jacob Kather	Integrating genetic and clinical data into artificial intelligence algorithms
Institut National de la Santé et de la	Prof. Jessica Zucman-Rossi	Identification of new variants

Department of Services Area SC
Transfusion Medicine Tel. 02
5503.4687
E-mail:daniele.prati@policlinico.mi.it
- Director: Dr. Daniele Prati

Recherche Medicale, Paris_INSERM		genetic predisposition to hepatocellular carcinoma
Hospitals of Paris_APHP	Professor Pierre Nahon	Characterization of the role of gene-environment interaction in the risk of hepatocellular carcinoma

Department of Services Area SC
Transfusion Medicine Tel. 02 5503.4687 - Director: Dr. Daniele Prati
E-mail:daniele.prati@policlinico.mi.it

INDEX

Department of Services Area SC
Transfusion Medicine Tel. 02 5503.4687 - Director: Dr. Daniele Prati
E-mail:daniele.prati@policlinico.mi.it

16 Appendix

Department of Services Area SC
Transfusion Medicine Tel. 02 5503.4687 - Director: Dr. Daniele Prati
E-mail:daniele.prati@policlinico.mi.it

1. INTRODUCTION

1.1 Background and rationale

Europe has the highest levels of alcohol consumption and burden of alcohol-attributable disease compared to other regions according to the World Health Organization (WHO). In 2019, 6.2% of all deaths in Europe were caused by alcohol. Alcohol-attributable deaths were mainly due to cancer (29%) and liver cirrhosis (20%).

It has been estimated that alcohol causes approximately 40% of premature liver-related deaths in Europe each year, although this number is likely underestimated. Alcohol-related liver disease (ALD) is the most common cause of liver cirrhosis and liver-related death in Europe with the peak age of death occurring in individuals in their 40s and 50s.

Thus, it remains a poorly understood disorder with research efforts and funding that contrasts sharply with its health relevance in Europe. Liver cancer is the second most common cause of cancer-related death (15-20% 5-year survival) and the second most common cause of alcohol-attributable cancer cases worldwide.

Hepatocellular carcinoma (HCC) accounts for approximately 90% of liver tumors and occurs most commonly against the background of ALD in Europe with the majority of cases (~90%) developing in patients with liver cirrhosis.

Some clinical features, including older age, male sex, obesity and type 2 diabetes, have been linked to alcohol-related HCC (ALDHCC) and, together with some common genetic determinants, bring alcoholic liver disease and NAFLD closer together and they seem to share similar pathogenetic patterns.

However, the effect of environmental variables including diet (e.g. type of alcoholic beverage), lifestyle and sociodemographic characteristics remain largely unknown. Both case-control studies and cancer databases have shown familial clustering suggesting genetic susceptibility. Liver carcinogenesis is the result of a complex multistep process generated by pro-oncogenic genetic alterations. These defects may pre-exist in normal cells (i.e. germline mutations) or result from DNA replication errors, spontaneous enzymatic conversions or DNA damage (i.e. somatic mutations). Recent studies have identified genetic variants in the WNT3A-WNT9A region as a novel susceptibility locus and confirmed the association of SNPs in PNPLA3 TM6SF2, HSD17B13. Note that SNP arrays are unable to detect

Department of Services Area SC
Transfusion Medicine Tel. 02 5503.4687 - Director: Dr. Daniele Prati
E-mail:daniele.prati@policlinico.mi.it

rare variants (allele frequency <1%) that can instead be assessed by whole exome sequencing (WES) and whole genome sequencing (WGS).

Like other complex diseases, ALD-HCC results from the interaction between environmental determinants and genetic variations but knowledge on gene-environment interactions is currently lacking in this area.
Although common genetic variants typically individually have a small effect size on disease risk, aggregating multiple common variants into polygenic risk scores (PRS) has been shown to successfully identify individuals at risk and outperform existing clinical models, especially when combined with multiple non-genetic information (e.g., lifestyle, anthropometric factors), thus enabling personalized cancer screening recommendations.

Advances in artificial intelligence (AI) algorithms have made it possible to extract clinically relevant information from complex and diverse clinical datasets of liver diseases and integrate genetic and non-genetic information into predictive information. Therefore, characterizing gene-environment interactions at the population level is of utmost importance to improve risk stratification for ALD-HCC and early diagnosis of the tumor.

This study is part of a larger European programme entitled “Understanding Gene ENvironment Interaction in ALcohol-related hepatocellular carcinoma”- HORIZON-MISS-2021-CANCER- 02-03- PROJECT CODE: GENIAL, whose general objective isassess the above needs through a comprehensive assessment of gene-environment interactions regarding ALD-HCC.
The project will assist in the identification of individuals at risk of developing alcohol-related hepatocellular carcinoma and in the development of strategies to prevent the development of alcoholic liver disease.

The GENIAL research project is coordinated by the UNIVERSITE LIBRE DE BRUXELLE. Among the various partners involved in the project, it includes the University of Milan (as Project Partner) and the Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico di Milano (as Affiliated Entity).
The GENIAL project has the following objectives:

Department of Services Area SC
Transfusion Medicine Tel. 02 5503.4687 - Director: Dr. Daniele Prati
E-mail:daniele.prati@policlinico.mi.it

- identification of novel genetic variants that modulate the risk of ALD-HCC (WP1)
- characterize how these variants are influenced by environmental factors in order to clarify the processes of carcinogenesis and identify potential targets for preventive therapies (WP2)
- provide a characterization of environmental factors (lifestyle, diet, socio-demographic factors) related to the development of ALD-HCC and integrate them with radiological, histological and genetic information using AI models to predict the risk of ALD-HCC (WP3)
- identify the machine learning approach with the greatest potential for clinical application in terms of improving diagnosis so as to help healthcare professionals provide better and faster recommendations and refine it for possible commercial application (WP3)

These evaluations will be carried out in 5 project work packages (WP) through a unified coordination, monitoring and evaluation component (WP4- WP5).
This study protocol in particular refers to what is foreseen in WP1 and WP3 which foresees the involvement of the State University of Milan and the IRCCS Foundation Ca' Granda Ospedale Maggiore Policlinico di Milano.

In this study we intend to contribute to investigate the genetic predisposition and the interaction between genetic and environmental factors in the development and progression of liver disease towards HCC. To do so, we will share the clinical and genetic data of two cohorts of patients belonging to two previously approved studies (EPIDEMIC-NAFLD cohort, SERENA study cohort) for the performance of WP1 and WP3 of the GENIAL project.

2.OBJECTIVE/S/I HYPOTHESIS OF THE EXPERIMENTATION

2.1 Primary objective

The primary objectives of the study are:

Department of Services Area SC
Transfusion Medicine Tel. 02 5503.4687 - Director: Dr. Daniele Prati
E-mail:daniele.prati@policlinico.mi.it

1) To evaluate whether the rare and common genetic variants identified by the Consortium in ALD-HCC patients (ULB and INSERM partners) predispose to the risk of HCC both in patients affected by NAFLD and ALD or whether they are ALD specific

2) To assess whether there is an association between the identified genetic variants and ALD-HCC within the general population (UK Biobank)

2.2 Secondary objective Secondary objectives of the study are:

1) Identify a polygenic risk score (PRS) that predicts HCC risk, combined with non-genetic variables.

2) Provide a comprehensive characterization of environmental factors (e.g. lifestyle, diet, sociodemographic factors) linked to HCC onset; and how these can be integrated with histological, radiological and genetic information and biomarkers of liver damage, making the best use of AI models to predict individual HCC risk.

Given that alcoholic liver disease and NAFLD share similar pathogenetic patterns, we will contribute to achieving these goals by sharing within the Consortium the clinical and genetic data and the characterization of environmental risk factors of two cohorts of patients already characterized from two previously approved studies (EPIDEMIC-NAFLD cohort, SERENA study cohort) involving patients affected by metabolic liver disease.

The polygenic risk score (PRS) for HCC risk prediction will be calculated by summing the number of steatosis-predisposing alleles in the genesPNPLA3, TM6SF2, GCKR, MBOAT7AndHSD17B13,weighted by the size of the effect of the genetic factor on liver fat as already described, Dongiovanni 2018). This score will be combined with non-genetic data (clinical risk factors for HCC such as age, sex, presence of diabetes, presence of severe liver fibrosis) to create a combined clinical-genetic score.

Department of Services Area SC
Transfusion Medicine Tel. 02 5503.4687 - Director: Dr. Daniele Prati
E-mail:daniele.prati@policlinico.mi.it

The combined risk score was developed considering acquired and genetic risk factors to predict HCC using the following formula as described elsewhere (Donati 2017): 1/ (1 + e− ((−12.588 + (0.162 * age) + (0.404 * Sex: 1 if male, −1 if female) + (0.259 * Obesity: 1 present, −1 absent)+(0.587 * T2DM: 1 present, −1 absent)+(1.299 * Severe fibrosis: 1 yes, −1 no)+(0.442 * number of risk alleles among common genetic factors associated with liver fat:inPNPLA3, TM6SF2, GCKR, MBOAT7And HSD17B13).

3. STUDY DESIGN

3.1 Study design

Interventional, with collection of biological, non-pharmacological, monocentric, retrospective-prospective material.

3.2 Inclusion criteria

Patients enrolled in the EPIDEMIC (approval no. 1822 of 27 August 2013) and SERENA (approval of latest amendment no. 1151_2021 of 9 November 2021) studies, already approved by the CE Milano Area 2, will be included.

That is, subjects with the following characteristics:

1. Diagnosis of NAFLD or cryptogenic liver disease, allowing a more liberal alcohol intake limit (<60/40 g/day in M/F), so as to also include subjects with a moderate alcohol component of the liver disease, an important factor given the high epidemiological burden of this group
2. Age between 45 and 75 years
3. Any of the following criteria:
   - 3.a. F3-F4 fibrosis, determined histologically, or by non-invasive techniques (stiffness > 7.9 kPa on Fibroscan and positive NAFLD fibrosis score or APRI or FIB4), or evidence of cirrhosis resulting from biochemical tests or imaging techniques;
   - 3.b. Family history of primary liver cancer in first degree of kinship, or carrier status of rare mutations associated with the development of HCC (such as mutations in APOB and TERT)
   - 3.c. Male patient with type 2 diabetes or obesity carrying at least three genetic variants in PNPLA3, TM6SF2, MBOAT7.

Department of Services Area SC
Transfusion Medicine Tel. 02 5503.4687 - Director: Dr. Daniele Prati
E-mail:daniele.prati@policlinico.mi.it

4.Willingness to sign informed consent.

3.3 Exclusion criteria

The exclusion criteria are those already foreseen for the two studies mentioned above, that is, they will beexcluded from the study patients with:

1.Alcohol intake >60/40 g/day in M/F
2.Chronic viral or autoimmune hepatitis
3.Any previously diagnosed genetic liver disease associated with increased risk of HCC (such as hereditary haemochromatosis, Wilson's disease, Alpha-1 Antitrypsin deficiency)
4.Use of drugs known to induce steatosis and liver disease
5.HCC diagnosed prior to study start date.
6.Other medical conditions with a prognosis of less than two years.

4. PROCEDURES RELATING TO THE STUDY

4.1 Intervention

In the context of this study, patients already included in the above-mentioned studies will be enrolled at the Foundation and the enrollment of 200 patients within the SERENA study will be continued.

Patients who agree to be included in the GENIAL study, as already foreseen in the context of the SERENA and EPIDEMIC studies, will be subjected (if prospectively enrolled) to the collection of the following data useful for the characterization of clinical risk factors: biohumoral parameters, histological parameters relating to liver biopsy, lifestyle and metabolic factors (body mass index BMI, abdominal circumference (AC), smoking, alcohol, diet, ongoing medical therapy), liver fat content and liver damage (by abdominal ultrasound and measurement of liver stiffness measurement (LSM) and controlled attenuation parameter (CAP) by Fibroscan), as indicated by clinical practice in these subjects.
Upon enrollment, patients will be subjected (if prospectively enrolled) to blood sampling with collection of a 6 ml EDTA tube and a 6 ml serum tube, respectively for DNA extraction and the dosage of some biomarkers associated with liver damage (IL-32, Pro-C3) (this last sampling is

Department of Services Area SC
Transfusion Medicine Tel. 02 5503.4687 - Director: Dr. Daniele Prati
E-mail:daniele.prati@policlinico.mi.it

repeated annually ad hoc for the study). Finally, as already foreseen by clinical practice, patients will be followed every six months for hepatological follow-up.

Any residual biological material not used by the end of the study will be biobanked, subject to the patient's consent to biobanking, for future research purposes in the same pathology at the Foundation's Biobank.

At the end of theprimary objectivesof the study, DNA samples from prospectively enrolled patients will be subjected to targeted sequencing to identify rare genetic variants and genotyping through SNP arrays to identify common variants.
DNA samples already collected in the context of the SERENA study that have not yet been sequenced will instead be analyzed, as per protocol, by whole-exome sequencing (WES).

In general, DNA will be extracted from peripheral blood using the automatic platform of the Fondazione IRCCS Ca' Granda. Quality control will be performed by evaluating the absorbance ratio 260/280 nM and by gel electrophoresis. The enrichment of the DNA library for exome sequencing will be performed using SureSelect Human All Exon v6 (Agilent). Sequencing will then be performed on the HiSeq 4000 platform (Illumina). This procedure will ensure an average coverage of 70-80X of the target region.

The obtained raw data will be analyzed using standard bioinformatics pipelines, including alignment, variant annotation, and function annotation. Functional analysis of mutations will focus on integrating the data by performing different analyses to comprehensively characterize variants in known genes associated with liver disease, as well as to search for novel genes and genetic risk factors potentially involved in HCC progression.
Multiple available resources such as genetic variant frequency data (ExAC, ESP, 1000Genomes), pathogenic variants (ClinVar) and insilico predictors of damage (SIFT, PolyPhen, CADD) will be used to prioritize candidate variants and genes involved in liver disease predisposition.

Department of Services Area SC
Transfusion Medicine Tel. 02 5503.4687 - Director: Dr. Daniele Prati
E-mail:daniele.prati@policlinico.mi.it

To complement these analyses, a genome-wide association study (GWAS) will also be conducted to evaluate genetic variants associated with severe liver disease and HCC. The GWAS study aimed at identifying genetic variants associated with severe liver disease and HCC will specifically include approximately n~1000 cases with severe liver disease related to NAFLD from the EPIDEMIC and SERENA studies vs n~10,000 controls with uncomplicated fatty liver whose genetic data are already available (public data - UK Bionbank)).

As explained in thesecondary objectives,a polygenic risk score (PRS) will also be created that predicts the risk of HCC. This score will then be combined with non-genetic data (clinical risk factors for HCC such as age, sex, presence of diabetes, presence of severe liver fibrosis) to create a combined clinical-genetic score.

In summary, the clinical and genetic data described above, whose collection and analysis is already foreseen for the EPIDEMIC and SERENA studies, will be shared in pseudonymised form within the GENIAL consortium in order to:
1) identify new genetic variants associated with HCC (this will be carried out in collaboration with the Institut National de la Santé et de la Recherche Medicale, Paris and the Universite Libre De Bruxelles under the responsibility of Prof Jessica Zucman-Rossi and Prof. Eric Trepo respectively);
2) evaluate the interaction between genetic, clinical and biohumoral factors in modulating the risk of HCC (at Hopitaux de Paris under the responsibility of Prof Pierre Nahon);
3) train an AI model to predict the risk of HCC (this algorithm will be developed at the Technical University of Dresden under the responsibility of Dr Jacob Kather) and identify hypotheses of clinical recommendations that healthcare professionals can use in the future in the diagnosis and follow up of ALD and NAFLD-related liver disease.

Finally, within the scope of the study proposal and in reference to the objectives of WP3, the genetic data resulting from the study will be sent to allow Start

Department of Services Area SC
Transfusion Medicine Tel. 02 5503.4687 - Director: Dr. Daniele Prati
E-mail:daniele.prati@policlinico.mi.it

up Stratipath Research and development with registered office at Nanna Svartz väg 4° Solna, Sweden to create a predictive model to reach the diagnosis of HCC.

5. ENDPOINT

5.1 Primary Endpoints

1) Identification of novel genetic variants associated with HCC (ALD and NAFLD related)
2) Association between identified genetic variants and ALD-HCC within the general population

5.2 Secondary Endpoints

1) Creation of polygenic-clinical risk scores for the development of HCC
2) Integration of genetic and non-genetic information into AI algorithms with the aim of deriving predictive information.

6. DURATION / TIMELINE OF THE STUDY

Gantt Chart

	0-6 month	7-12 month	13-18 month	19-24 month	25-36 month	37-42 month	43-48 month	49-54 month	55-60 month
Screening criteria inclusion/exclusion	X	X	X	X	X				
Explanation of the study and signing of the informed consent	X	X	X	X	X				
Anamnesis collection and parameters (clinical risk factors, lifestyle and metabolic factors (body mass index BMI, abdominal circumference (AC), smoking, alcohol, diet, ongoing medical therapy, presence of diabetes)	X	X	X	X	X				
Blood tests (indicators of necrosis and liver damage: AST, ALT, GGT, ferritin,	X	X	X	X	X				

Department of Services Area SC
Transfusion Medicine Tel. 02 5503.4687 - Director: Dr. Daniele Prati
E-mail:daniele.prati@policlinico.mi.it

blood count, blood sugar, glycated hemoglobin, cholesterol, triglycerides)									
Blood sampling for genetics*	X	X	X	X	X				
Non-invasive assessment of liver damage (LSM) and liver fat (CAP) using Fibroscan and abdominal ultrasound		X	X	X	X	X	X		
Annual blood sampling for serum biomarker assay (interleukin 32, IL-32; procollagen 3; pro-C3)*	X	X	X	X	X				
Hepatological follow-up			X	X	X	X	X	X	X
Data Analysis					X	X	X	X	X

* Ad hoc for the study

Study start: December 2023
Enrollment end: December 2026
Study end: December 2028 The study will last 60 months.

7. STATISTICAL ANALYSIS

7.1 Sample size

The EPIDEMIC-NAFLD cohort includes approximately 2800 subjects with NAFLD (n=347, 11.6% with advanced fibrosis, n=251, 8.9% with HCC). The SERENA cohort includes approximately 400 patients with advanced NAFLD.

Based on the hypothesis based on previous literature that we will observe a 1.5% HCC incidence rate in the entire cohort over a mean follow-up of 5 years, and that based on data generated in a cross-sectional cohort (Donati, Scientific Reports 2017 and Cancer Medicine 2017) we will be able to identify at least 30% of individuals with at least a threefold increased risk of HCC, for a type I error rate of 5%, we planned to enroll 375 patients to have a statistical power of 80% to detect a significant association of genetic score with HCC risk. However, we estimated the need to increase the number of enrollees to account for patients

Department of Services Area SC
Transfusion Medicine Tel. 02 5503.4687 - Director: Dr. Daniele Prati
E-mail:daniele.prati@policlinico.mi.it

potentially censured due to the occurrence of other major clinical events prior to the onset of HCC (death, cancer, major cardiovascular events, and liver failure) and a small group lost to follow-up. Therefore, we intend to enroll at least 600 patients to complete the study. We then expect to prospectively enroll another 200 patients over the next 3 years.
The GWAS study aimed at identifying genetic variants associated with severe liver disease will include approximately n~1000 cases with severe NAFLD-related liver disease from the EPIDEMIC and SERENA studies vs n~10,000 controls with uncomplicated fatty liver disease whose genetic data are already available (public data - UK Biobank).

7.2 Data analysis

Time to progression to HCC will be estimated by Kaplan-Meier estimates of cumulative incidence rates. Survival distributions among genotypes will be compared using Cox proportional hazard models. Hazard ratios and P values will be adjusted for confounding factors. In the analysis, risk genotypes will be coded using additive genetic models.
We will evaluate the impact of genetic risk variants on HCC risk, using Cox regression models adjusted for classical HCC risk factors (age, sex, BMI, type 2 diabetes, advanced fibrosis, smoking, alcohol intake).
As a secondary outcome, we will also evaluate the interaction between genetic and acquired risk factors in the pathogenesis of HCC. As a secondary objective of the study, we will develop a combined risk score to predict the incidence of HCC, which will be cross-validated within the cohort. Statistical analyses will be performed using JMP 16.0 Pro (SAS Institute) and version 3.3.2 of the statistical analysis software R (http://www.Rproject . org/). P-values <0.05 will be considered statistically significant.

P-values <0.05 (two-tailed test) will be considered statistically significant.

8. ADVERSE EVENTS

Department of Services Area SC
Transfusion Medicine Tel. 02 5503.4687 - Director: Dr. Daniele Prati
E-mail:daniele.prati@policlinico.mi.it

The project does not involve the administration of drugs or other substances or invasive clinical practices other than those normally foreseen in clinical practice. Therefore, no adverse events are expected.

9. RISK/BENEFIT ASSESSMENT

We expect to define whether a genetic HCC risk score can accurately predict the onset of NAFLD-HCC in a high-risk cohort. This "Precision Medicine" approach would have an important clinical implication, allowing to identify subjects for whom ultrasound screening for HCC could be cost-effective with potential clinical benefits for patients and savings for the health system.

Validation of genetic variants associated with ALD-related HCC within our cohort will also help to elucidate the pathogenic mechanisms of alcoholic and non-alcoholic liver disease, possibly demonstrating how these conditions share similar processes.

10. STUDIO MANAGEMENT

10.1 Data collection and management

Company software designed to manage outpatient visits and laboratory tests will be used to collect data.

Each participant, at the time of enrollment, will be assigned a unique code. The de-identification of the data will be done in such a way that the people who access the database will not be able to trace the identity of the subjects in any way. Only local experimenters will be able to trace the identity of the enrolled subjects.

The data needed for the study will be recorded in a specific eCRF in a Data Management System validated according to national regulations, provided by the Scientific Direction of the Foundation. The platform used will be RedCap (Research Electronic Data Capture).

The REDCap Consortium is composed of >1000 institutional partners worldwide (research institutes, universities, ministries etc). The consortium supports a secure web application (REDCap) designed exclusively to support data acquisition for research studies. The REDCap application allows users to create and manage online databases quickly and securely, and is currently in use for more than

Department of Services Area SC
Transfusion Medicine Tel. 02 5503.4687 - Director: Dr. Daniele Prati
E-mail:daniele.prati@policlinico.mi.it

110,000 projects with approximately 150,000 users covering numerous areas of research interest across the consortium.

Through REDCap, for this study the following will be implemented: a) user-level identification, with specific restrictions based on the role in the study b) real-time data integrity validation and control c) patient de-identification before data export d) centralized data storage with daily backup, a secure server within the Foundation's IT structure.

10.2 Regulatory aspects and ethical considerations

10.2.1 Approval by the Competent Authority

In accordance with applicable regulations, the principal investigator must obtain approval from the appropriate Competent Authority prior to initiating the clinical study.
This study will be conducted in accordance with ICH/GCP (International Conference of Harmonization/Good Clinical Practice) rules and all applicable laws, including the Declaration of Helsinki.
of June 1964, as last amended by the World Medical Association General Assembly in Seoul, 2008.

10.2.2 Ethics Committee Approval

The investigator must ensure that the protocol has been reviewed and approved by the local independent Ethics Committee (EC) before starting the study.
The CE must also review and approve the informed consent (IC) form and all written information received from the patient prior to enrollment in the study. Should it be necessary to modify the protocol and/or the IC during the study, the investigator will be the guarantor and therefore the person responsible for ensuring the review and approval of such modified document as requested by the CE.
The content of these changes will be implemented only after the CE has approved them. Until then, it will be necessary to refer to the previous version of the document already approved.

10.2.3 Informed consent (IC)

The investigator or other personnel designated by him have the task of informing the subjects about all aspects and procedures of the study.

Department of Services Area SC
Transfusion Medicine Tel. 02 5503.4687 - Director: Dr. Daniele Prati
E-mail:daniele.prati@policlinico.mi.it

The informed consent of the participants will be collected by the physician in charge of the clinical and metabolic evaluation during the visit.
The process for obtaining informed consent must comply with applicable regulatory procedures. The investigator (or designated collaborator) and the subject must date and sign the informed consent form before the patient begins any study-related procedures.
The subject will receive a copy of the CI dated and signed by both parties; the original copy will be kept in the designated study archives. Neither the investigator nor designated personnel should in any way exert any coercion or influence on a subject to induce him or her to participate or continue to participate in the study. A subject's decision to participate in the study must be completely voluntary. The investigator and designated personnel should emphasize to the subject that he or she may withdraw consent at any time without penalty or loss of any benefits to which he or she may be entitled. Written or oral information about the study, including the written consent form, must not contain any language that forces the subject to waive (even apparently) his or her legal rights, or that would exonerate the investigator, agency, or sponsor from liability for negligence.

The time that the subject will have to dedicate to participating in the study does not exceed that normally foreseen by normal clinical practice for the treatment of the condition from which he/she suffers.

10.3 Duties of the experimenter

In accordance with applicable local regulations, the investigator must submit periodic reports regarding the progress of the study at his/her site to the CE and notify the CE of study closure. Periodic reports and closure notification are part of the investigator's responsibilities.

10.4 Study monitoring

In accordance with applicable regulations and good clinical practice (GCP), the monitor must visit or contact the center periodically. The duration, nature and frequency of such visits / contacts depend on the recruitment frequency, the quality of the documents held by the center and their adherence to the protocol. Through such contacts, the monitor must:
• monitor and evaluate the progress of the study

Department of Services Area SC
Transfusion Medicine Tel. 02 5503.4687 - Director: Dr. Daniele Prati
E-mail:daniele.prati@policlinico.mi.it

• examine the collected data
• conduct source document verification
• identify each problem and related solutions
The purposes of the monitoring activity are to verify that:
• the rights and well-being of the subject are respected
• the study data are accurate, complete and verifiable from the original documents
• the study is conducted in accordance with the protocol and any approved amendments, GCP and
applicable regulations
The experimenter must:
• give the monitor direct access to all relevant documentation
• dedicate part of his time and his staff to the monitor to discuss the monitoring results and any other possible aspects.
The monitor must also contact the center prior to the start of the study to discuss the protocol and data collection procedures with the staff.

10.5 Study quality assurance

As the sponsor, Fondazione IRCCS Ca' Granda, Ospedale Maggiore Policlinico may, at its discretion, carry out a quality control of the study. In this case, the investigator must allow the monitor direct access to all relevant documentation and dedicate part of his or her time and personnel to the reviewer to discuss the results of the monitoring and any other aspects of the study.
In addition, Regulatory Authorities may conduct inspections. In this case, the investigator must allow the inspector direct access to all relevant documentation, and dedicate part of his time and personnel to the inspector to discuss the monitoring results and any other aspects of the study.

10.6 Closing of the study

At the time of study closure, the monitor and the experimenter must activate a series of procedures:
• review all study documentation
• reconcile study data
• reconcile all clarifying reports.

Department of Services Area SC
Transfusion Medicine Tel. 02 5503.4687 - Director: Dr. Daniele Prati
E-mail:daniele.prati@policlinico.mi.it

10.7 Document storage

In accordance with current national regulations, the investigator must keep a copy of all documentation and store it in a dry and safe place after the study has been closed.

10.8 Disclosure of information regarding scientific discovery

10.8.1 Confidentiality

The investigator and other personnel involved in the study must handle all information relating to the study (including the protocol, data obtained and all documentation generated during the study) and must not use such information, data or reports for purposes other than those described in the protocol. These restrictions do not apply to:

1) information that becomes publicly available, not due to negligence on the part of the investigator or his staff;

2) information that requires confidential disclosure to CE for the sole purpose of evaluating the study;

3) information that must be disclosed in order to obtain appropriate medical care for a study subject.

10.8.2 Publications

Each Party is the owner of the intellectual and industrial property rights relating to its Background. Furthermore, the Affiliated Entity will grant the Beneficiary access rights to the Background and the Results, if necessary for the implementation and exploitation of the results, in order to fulfill its obligations and perform its tasks as described in the Grant agreement and in the Affiliated Entity Agreement stipulated between the Parties. This provision applies vice versa to the Affiliated Entity if it requests Access for the performance of its tasks and/or for the valorization of its Results.

The results are owned by the Party that generates them. If the Parties have jointly produced results and it is not possible to establish their respective contributions or separate them for protection, the Parties automatically become co-owners. In case of co-ownership of the Results, the ownership share of the co-owners will be proportional to the intellectual contribution invested in generating that specific Result. The co-owners agree on the ownership shares, all protection measures and the distribution of related costs in a co-ownership agreement.

Department of Services Area SC
Transfusion Medicine Tel. 02 5503.4687 - Director: Dr. Daniele Prati
E-mail:daniele.prati@policlinico.mi.it

co-ownership to be negotiated in advance and in any case before any commercial exploitation of the joint results.
In any case, the Parties undertake to adequately protect their Results taking into account all relevant considerations, including the prospects of commercial exploitation.
For further information, please refer to the Affiliated Entity Agreement.

11. Indemnity and compensation in case of damages

In case of unwanted events or any damages that may arise from participation in research, our Institute's Insurance Policy also extends to cover subjects participating in research projects.

12. Amendments to the Protocol

Communications regarding significant changes to the protocol will be the responsibility of the researchers involved in the study and will be under the responsibility of the principal investigator.

13. Financial Agreements

The interventions planned within the study are partly foreseen within normal clinical practice.
The costs of study procedures exceeding normal clinical practice (genotyping, genetic analysis and biomarkers) will be covered by the following funding:
The European Union, HORIZON-MISS-2021-CANCER-02-03 program “Genial” under grant agreement “101096312”

14. Disclosure on conflicts of interest

The experimenters declare no conflict of interest.

15. References

- World Health Organization. Status report on alcohol consumption, harm and policy responses in 30 European countries 2019. (2019).
- World Health Organization. https://www.who.int/publications/i/item/9789241565639 (2018). GBD 2019 Diseases and Injuries Collaborators. The
- Lancet. 396, (2020).

Department of Services Area SC
Transfusion Medicine Tel. 02 5503.4687 - Director: Dr. Daniele Prati
E-mail:daniele.prati@policlinico.mi.it

- Sheron, NJ Hepatol. 64, (2016). Karlsen, T. H.
- The Lancet 399, (2022). Rumgay, H. Lancet
- Oncol. 22, (2021). Llovet, J.M. Nat. Rev. Dis.
- Primer 7, (2021). Hemminki, K. Gut 52,
- (2003).
- Vogelstein, B. Science 339, (2013). Guichard,
- C. Nat. Genet. 44, (2012). Schulze, K. Nat.
- Genet. 47, 505–511 (2015). Ally, A. Cell 169,
- (2017).
- Buniello, A. Nucleic Acids Res. 47, (2019).
- Trépo, E. Lancet Oncol. 23, (2022).
- ENCODE Project Consortium Nature 583, (2020).
- Gatchel, J.R. Nat. Rev. Genet. 6, (2005). Timpson,
- NJ Nat. Rev. Genet. 19, (2018). Lappalainen, T.
- Science 373, (2021).
- Clayton, DG PLoS Genet. 5, (2009).
- Ye,Y., Trends Genet. TIG 36, (2020).
- King, EA, PLoS Genet. 15, (2019).
- Maas, P. JAMA Oncol. 2, (2016).
- Donati, Scientific Reports (2017)
- Dongiovanni, J Intern Med (2018)

16. Appendix