STUDY PROTOCOL

Characterization of type 2 diabetes subgroups at diagnosis: a necessary step towards precision medicine in diabetes.

Protocol version and date	Version: 1.0 Date: 07/17/2021
Protocol Code	IIBSP-CDM- 2021-01
Sponsor	Fundació Institut de Recerca de l'Hospital de la Santa Creu i Sant Pau saint Quinti , 89 08041 Barcelona Phone: 93 5565661 Fax: 93 5565602
Coordinator Researcher	Dídac Mauricio, Department of Endocrinology and Nutrition Hospital of Santa Creu i Sant Pau saint Quinti , 89 08041 Barcelona Tel: 627953473 e-mail: dmauricio@santpau.cat
Principal researchers	Dídac Mauricio, Endocrinology and Nutrition Service Hospital of Santa Creu i Sant Pau saint Quinti , 89 08041 Barcelona Phone: 627953473 e-mail: dmauricio@santpau.cat
Number of protocol pages	80

Contents

GLOSSARY OF TERMS

T2DM: type 2 diabetes mellitus

CAP: primary care centers

HT: antihypertensive treatment

GAD: glutamic acid decarboxylase antibodies / anti -glutamic acid decarboxylase antibodies

BMI: body mass index

HbA1c: glycosylated hemoglobin

HOMA-B: homoeostasis model estimate of β- cell function / β -cell function homeostasis model estimates

HOMA-IR: homoeostasis model estimate of insulin resistance / insulin resistance

SAID: severe autoimmune diabetes

SIDD: severe insulin-deficient diabetes

SIRD: severe insulin-resistant diabetes

MOD: mild obesity-related diabetes

MARD: mild age-related diabetes

1. SUMMARY:

Promoter identification	Santa Creu and Sant Pau Hospital Research Institute - IIB Sant Pau c / Sant Antoni Maria Claret , 167 08025 Barcelona Tel: 93 553 78 69
Study Title	Characterization of subgroups of type 2 diabetes mellitus at diagnosis: a necessary step towards precision medicine in diabetes, COPERNICAN study
Protocol code	IIBSP-CDM- 2021-01
Protocol version and date	Version: 1.0 Date: 07/17/2021
Coordinating researcher	Dídac Mauricio, Department of Endocrinology and Nutrition Hospital of Santa Creu i Sant Pau saint Quinti , 89 08041 Barcelona Tel: 627953473 e-mail: dmauricio@santpau.cat
Principal investigator	Dídac Mauricio, Department of Endocrinology and Nutrition Hospital de la Santa Creu y Sant Pau c / Sant Antoni Maria Claret , 167 08025 Barcelona Tel: 627953473 e-mail: DMauricio@santpau.cat
Collaborating researchers	Lleida CAPs : • Neus Miró • Jesus Pujol • Monica Solanes • Antonieta Lafarga • Sandra Guerrero • Mireia Marin Barcelona CAPs : • Begoña Ribas • Natalia Riera • Ana M. Pedro Pijoan • Manel Mata • Helena Hernández Boluda • Xavier Cos

	USR Lleida • Marta Ortega • Eva Artigues • Inés García • Laura Azlor • Joaquim Sol Karen Rodriguez Ferran Primary Care Center (Lleida) Minerva Granado CIBER for Diabetes and Metabolic Diseases (CIBERDEM) Cintia González Blanco Hospital de la Santa Creu y Sant Pau Cristina Farras Martha Hernandez Merce Ibarz Maria Bernal Murillo Vilanova University Hospital USR of Barcelona Josep Franco University of Barcelona e-mail: josep.franch@gmail.com Jordi Real Gatius CIBER of Diabetes and Metabolic Diseases (CIBERDEM) Emerald Castelblanco CIBER of Diabetes and Metabolic Diseases (CIBERDEM) Bogdan Vlacho (project manager) Epidemiological Research Group on Diabetes from Care Primary (DAP-CAT)
Participating centers	Hospital de la Santa Creu i Sant Pau (IR-HSCSP, Barcelona) Primary care centers in Lleida and Barcelona Lleida Research Support Unit, Barcelona Research Support Unit
CEIC of reference	CEIm St. Paul For ICS centers: CEIm IDIAP Jordi Gol i Gurina for CAP's Consortium Comprehensive Sanitary: Consortium CEIC Comprehensive Sanitary
Main goal	This project aims to characterize a large group of people with type 2 diabetes mellitus (T2DM) recently diagnosed in the Spanish population.

Design	Observational prospective cohort study
Disease under study	Diabetes mellitus type 2
Methodology	We will conduct a multicenter prospective observational study. The scope of the study is made up of Primary Care Centers (CAP) in 2 care areas (Barcelona and Lleida). Study procedures will be performed by the local investigator, who will not be the subject's usual primary care physician.
Study population and total number of subjects	People with T2DM
Calendar. Expected duration of the study.	From January 2022 to December 2024 (3 years).
Ethical considerations	All newly diagnosed subjects with a clinical diagnosis of T2DM will be invited to participate and only those subjects who sign the informed consent for the study will enter the study.
Funding source	Instituto de Salud Carlos III Proyecto FIS PI21/01163
Keywords	Prospective cohort studies, characterization, T2DM

2. ABSTRACT

Objectives: Our primary objective is to identify and characterize the relevant subgroups (clusters*)* in type 2 diabetes (T2DM) at the time of diagnosis in our region. This objective will be addressed using the 6 main variables (see Methodology) that have been used so far in previous studies in other populations to develop the *clustering* of diabetes. Therefore, this is a study in the field of diagnostic precision medicine in diabetes. In addition, as secondary objectives we will evaluate other phenotypic characteristics of these subgroups (clinical, metabolic, and associated comorbidities).

Methodology: this project will establish a prospective observational cohort study of 1200 subjects newly diagnosed with T2DM in 11 primary care centers (PAC) of the healthcare areas of Barcelona city (5 CAP) and the territory of Lleida (6 PCC). We will identify and evaluate all newly diagnosed cases of T2DM. Participants will undergo a comprehensive phenotypic evaluation, including the 6 variables that will allow the characterization of T2DM subgroups: age, antibodies against glutamic acid decarboxylase (GAD), body mass index, glycated hemoglobin, and sensitivity indices (HOMA-IR) and insulin secretion (HOMA beta), based on the determination of C-peptide. We will carry out the latest generation cluster analysis (k- means and hierarchical clustering), following the method described in previous studies using the 6 variables mentioned above. Onset diabetes subgroups and their association with secondary outcome variables will be assessed. Other procedures of this project include: clinical (including complications) and biochemical evaluation, advanced lipoprotein profile, and validated questionnaires for the evaluation of diet and physical activity. We will also evaluate the initial prescription of antidiabetic medication.

Keywords: characterization, conglomerates, primary care centers, type 2 diabetes

3. THEORETICAL FRAMEWORK, PROBLEM STATEMENT:

3.1 Background

Diabetes mellitus is one of the most prevalent chronic diseases, ranking eighth in terms of the global burden of disease measure developed by the world health organization-WHO [1]. Specifically, type 2 diabetes (T2DM) is reaching epidemic proportions. This is also the case in Spain, where the T2DM prevalence in people over 18 years of age is around 14% [2]. Furthermore, diabetes reduces quality of life and increases years of life lost, especially due to chronic complications [3].

3.1.1 Heterogeneity of type 2 diabetes mellitus

Traditionally, diabetes is classified into different types: type 1, type 2, gestational diabetes, and other specific types of diabetes. However, one of the main problems is that T2DM is a diagnosis based on the exclusion of other types of diabetes. This has led to the inclusion in the diagnostic category of T2DM, subgroups of subjects clearly differentiated from each other in terms of their clinical characteristics, especially in terms of response to treatment and development of complications. Evidence suggests that T2DM is a highly heterogeneous complex metabolic disease encompassing different pathophysiological and genetic characteristics [4, 5]. Furthermore, the presentation and progression of the disease may vary between subjects, leading to poor metabolic and glycemic control [5]. Therefore, it is well known that heterogeneity is one of the main problems affecting T2DM. Actually, this is an important aspect in terms of disease management and prognosis decisions. For example, the response to different treatments varies widely among subjects with T2DM [6]; this is clearly leading to a significant proportion (44%) of people not achieving good glycemic control [7]. Furthermore, in Europe alone, during the period between 2005 and 2017, the European Medicines Agency (EMA) has approved 40 new drugs to treat diabetes, mainly for T2DM [8]. Thus, we have more complex combinations of treatments without having achieved a similar advance in the development of more precise tools to characterize the disease itself. Thus, although novel treatments are available, their use and the

individualization of treatment continue to be based on "classic" bases, from which considerable variability arises in clinical practice.

Our group has been a pioneer in research on the characterization of the heterogeneity of T2DM in Spain. In fact, we have carried out different studies on the so-called subtype of latent autoimmune diabetes in adults (LADA). So far we have been able to contribute to the characterization of LADA in Spain by carrying out immunological, metabolic, genetic and clinical characterization studies of this subtype of diabetes. The result of our research in this field consists of multiple articles based on our own cohort [9-12] or with internationally aggregated data [13-16]. Although this subtype of diabetes has not yet been accepted by the ADA [17], it should be underlined that in 2019 the WHO already included LADA as a different type of diabetes under the subset of hybrid forms of diabetes in its updated classification [18]. This is a paradigmatic example of what happens in terms of diagnosis, prognosis, therapeutic decisions, and follow-up in a different subgroup of subjects initially categorized as T2DM.

3.1.2 Current aspects of precision medicine in diabetes

Traditional methods using the clinical phenotype and basic physiology have been used by clinicians for decades. However, in view of the aforementioned complexity of new therapeutic approaches, it is clear that we need to introduce new tools in clinical decision making. This should lead to research in deep phenotyping and the introduction of advanced molecular tools, including different omics [5,20]. This implies a new stratified approach with the use of molecular and deep phenotyping, which is, adapting diagnosis or therapeutics (from prevention to treatment) to subgroups of populations that share similar characteristics, thus minimizing error and risk and maximizing the efficacy (i.e. precision medicine). Interestingly, a recent collaboration between the ADA and the EASD issued a consensus on the development of precision medicine in diabetes [21]. The full landscape of precision medicine includes precision diagnostics, precision prevention, precision treatment, precision prognosis, and precision monitoring. For this study we will focus solely on precision diagnoses, although the future

knowledge generated by this project and the follow-up of the subjects of this study will undoubtedly contribute to the rest of the fields of precision medicine in diabetes in our country.

An accurate diagnosis is a decision based on probability, which is usually made at a specific point in the natural history of the disease and is not an absolute or permanent state. Taking this into account, diabetes mellitus can be considered as a condition in which precision diagnostics should ideally be implemented. T2DM is already a paradigm of the need for this novel approach [22-24], but it has yet to come into force in routine clinical practice.

We can consider T2DM as a condition that was defined based on simple a clinical phenotypic trait that, however, contains different groups (clusters) of subjects. Recently published studies used cluster analyzes to refine the classification of T2DM beyond the simple use of blood glucose levels [25-29]. These studies identified subgroups of subjects with diabetes mellitus who have specific patterns of development of complications. In 2018, the seminal study by the group led by Leif Groop in Sweden [25], identified five new groups of subjects using 6 different variables, i.e., glutamic acid decarboxylase (GAD) antibodies, age at diagnosis, BMI, HbA1c, homeostasis model, cell function estimates β (HOMA-B) and insulin resistance (HOMA-IR). According to the grouping of subjects based on these phenotypic variables, 5 subgroups were defined: SAID severe autoimmune diabetes; SIDD, severe insulin deficiency diabetes; SIRD, severe insulin resistant diabetes; MOD, mild obesity-related diabetes; MARD, mild age-related diabetes. This classification based on combined information of these 6 variables proved to be a more precise and clinically useful stratification than the diagnosis based on a single metabolite such as glucose or HbA1c, together with all the traditional clinical factors. Another study from Germany studied this comprehensive phenotyping in its local diabetic population [26], in which the distribution of clusters differed slightly from that reported in the Swedish population-based cohort. They observed a higher prevalence of the SAID subtype of diabetes, and a lower prevalence of the SIDD subtype in their population and concluded that these specific groups showed different metabolic disturbances and patterns of risk of

developing diabetes-related complications [26]. In another similar study from the UK, researchers replicated the initial pooling of the German study, in a cohort of 6,810 people with prediabetes. Six distinct groups were identified: 1: low risk, 2: very low risk, 3: β-cell failure, 4: low-risk obesity, 5: high-risk insulin-resistant fatty liver, and 6: visceral fatty kidney disease high risk. Although groups 3, 5, and 6 were associated with higher blood glucose at baseline than the other groups, only groups 3 and 5 were associated with a higher risk of T2DM in longitudinal analyses. Furthermore, groups 3, 5, and 6 were associated with an increased risk of kidney disease [27]. Other investigators have built their research on the initial findings of Ahlqvist et al [28, 29], while others applied omics tools to clustering based on 6 phenotypic variables [30]. The most representative line of research in this line is that developed by the group of J. Florez [24], in which they have been able to combine the phenotypic and genetic approach with a combination of 94 loci that represent the genetic variant of T2DM to develop partitioned scores that identified 5 different subgroups of subjects harboring different traits of insulin deficiency (beta cell and proinsulin) and insulin resistance (obesity, lipodystrophy, and liver lipids) [30]. Furthermore, they were able to relate these groups to relevant clinical outcomes (cardiovascular and metabolic complications). Interestingly, the existence of different subgroups is not a static image, but a moving one, where the trajectory of the subjects in the follow-up does not always necessarily fit into the same group; that is, an individual can move from one group to another [23,26, 28,29]. The fact that the initial groups experience different diabetes component pathways emphasizes the importance of monitoring the trajectory of the different groups and the individuals included [23]. It should be noted that some of the studies that replicate the initial study by Ahlqvist et al. [3] do not include the SAID cluster in their approach, since it is not, strictly speaking, a subgroup of subjects belonging to T2DM once identified.

Finally, it should be noted that phenotypic characteristics have been shown to be associated with response to different drugs [19,28]. Therefore, it is highly relevant to initially implement a deeper phenotyping of each population (precision diagnoses) in order to further build therapeutics and precision prognosis.

Therefore, future phenomic and genetic characterization of large cohorts in other populations is needed to move towards better prediction of T2DM, complications, and precision medicine.

3.2 Justification

In this context, we strongly believe that the proper characterization of subjects clinically categorized as having type 2 diabetes is an important issue. Advances in this field should allow us to implement the different strategies of precision medicine in our country. It is of the utmost importance that we establish strategic initiatives to delve into our own population so that we can implement all the new potential tools of precision medicine for the diagnosis, prognosis, prevention and treatment of diabetes. Therefore, we think that the current project is fundamental as a pioneering study in our country. Our objective is to evaluate and characterize the different subtypes of T2DM in our population based on previous work by other groups in this field.

4. HYPOTHESIS, PRIMARY AND SECONDARY OBJECTIVES

4.1 Assumptions

The previously described subtypes of subjects included in the diagnostic category of T2DM at the onset of the disease have been evaluated and confirmed in other populations, mainly Caucasians. However, we are not aware of any prospective study that is responsible for characterizing the subgroups of diabetes with onset in adulthood, especially T2DM, in a sufficiently potentiated sample in our country. Therefore, it has not yet been established whether the different conglomerates ("clusters") of adult-onset diabetes subtypes exist in our population. Our main hypothesis is that there is also heterogeneity of T2DM in our population, characterized by a relevant grouping of subtypes that are potentially similar to those described in other European populations.

We also hypothesize that the different subgroups identified would also show a different profile for other phenotypic characteristics that are represented by other clinical variables, such as comorbidities (fatty liver, clinical and subclinical atherosclerosis), metabolic/genetic characteristics (advanced metabolomics and lipoprotein profile, genetic alterations) and life habits (diet and physical activity).

4.2 Primary objective:

To assess and characterize the relevant subgroups (clusters) of type 2 diabetes at diagnosis in our population.

This objective will be addressed using the 6 main variables that have been used so far in previous studies in other populations to develop disease groupings, that is, age, GAD antibodies, body mass index, glycosylated hemoglobin (HbA1c), and estimates of the evaluation of the homeostatic model 2 of β-cell function and insulin resistance.

4.3 Secondary objectives:

The study design will allow us to study other characteristics of the study subjects and of the disease itself at the time of diagnosis:

Relevant phenotypic characteristics not included as variables in the main objective:

- Other clinical characteristics (for example, gender)
- Frequency of various associated comorbidities, including micro and macrovascular complications. As additional comorbidities we also include: non-alcoholic fatty liver and liver fibrosis scores; Subclinical carotid and femoral atherosclerosis.
- Advanced lipoprotein profile.
- Lifestyle variables: eating habits and physical activity.
- Use in real clinical practice of antidiabetics and other treatments at the beginning of the disease.
- Appearance of different micro- and macro-vascular complications of diabetes
- Metabolomics and genetic profile related to diabetes

5. METHODS:

5.1 Study design

We will conduct a multicenter prospective observational study. The study design contemplates a single visit that will coincide with the inclusion of the patient in the study. Prior to the inclusion visit, a pre-selection period is planned in which the possible candidate will be invited, all the study procedures will be explained and clarified, the selection criteria will be checked and the inclusion visit will be scheduled.

The scope of the study is made up of Primary Care Centers (CAP) in 2 large health areas (Barcelona and Lleida). In the **Annex 8 section,** the characteristics and expected recruitment capacity of each CAP are detailed. We have taken a conservative approach using the lower range of incidence calculated based on our own data from 2019 to 2020. The table provided used an expected acceptance rate for participation of 80%. Additionally, CAPs were selected based on commitment and performance in our previous studies. It is planned to include 1200 participants newly diagnosed with T2DM. Depending on compliance with the recruitment rate, the inclusion of other CAPs is possible to achieve the expected number of participants in the study.

5.2 Study period

The study period is three years, from January 1, 2022 to December 31, 2024.

5.2.1 Recruitment period

The recruitment period for participants is 2 years, from January 1, 2022 to December 31, 2023

5.3 Follow-up of participants and linkage with other databases

For this study, a pragmatic follow-up of the participants will be chosen, which is adapted to the usual clinical practice for the follow-up of diabetes. For this reason, the monitoring will be done from:

- SIDIAP/PADRIS databases, with the link to the electronic medical record of the participants who have accepted this option in the informed consent.
- Electronic medical record, in the event that the link with the SIDIAP/PADRIS databases is not possible for technical reasons (non-ICS centers), the collection of monitoring variables will be done directly from the clinical course (review of clinical history) of participants without scheduling additional study follow-up visits.
- For specific causes of mortality, data linkage with the National Institute of Statistics (INE) will be used.
- If the participants have already participated or are participating in other clinical research studies and want to transfer the data they have provided for other projects to this study, this linking (portability) of the data will be allowed.

All participants will be followed until the end of the study, death or withdrawal of informed consent.

5.4 Study population

People with type 2 diabetes mellitus treated in primary care centers in Barcelona and Lleida.

5.5 Selection criteria

5.5.1 Inclusion criteria:

- Subjects of both sexes, of legal age (≥18 years)
- Newly diagnosed T2DM according to the criteria of the *American Diabetes Association*
- With a maximum duration of three months of the disease at the time of inclusion.
- Informed consent signature

5.5.2 Exclusion criteria:

Subjects with a new diagnosis of other types of diabetes (type 1, MODY, gestational or other causes) will also be evaluated to ensure that the diagnostic criteria for selection are correctly implemented in all potential study participants.

*For final analysis, subjects with any diagnosis of diabetes other than T2DM will be excluded from the study.

5.5.3 Withdrawal criteria:

All included patients have the right to withdraw from the study at any time, withdrawing their consent, without having to justify this decision, and without this being detrimental to their clinical follow-up. Likewise, the participant may revoke the use of their data in the analysis, without justifying their decision, and without this resulting in any liability or damage.

5.6 Study procedures

This study will be launched in the participating primary care centers, once a favorable opinion has been obtained by the CEI, and following the legislative requirements for this type of observational study.

5.6.1 Preselection

Each study subject will be identified by their primary healthcare professional and they will notify the main investigator or collaborators of the center through the electronic medical record system about the possible candidate. All newly diagnosed subjects with a clinical diagnosis of T2DM will be invited to participate. The study procedures will be carried out by local researchers in each participating center, who in some cases will be the same doctors and nurses as the participants. A specific visit will be scheduled with the local investigator at each CAP to perform all study procedures in a single visit.

5.6.2 Inclusion

Each patient invited to participate will be informed about the study both verbally and in writing form, providing a document called "Patient Information Sheet" (**Annex 5**). No participant will be included until they have been duly informed by the researcher and have freely given their informed consent to participate in the study (**Annex 6**). Once compliance with all the selection criteria has been confirmed and informed consent has been granted, data collection will begin by completing an electronic case report form (eCRF), intended to record the information available in the patient's clinical history, patient or obtained during the inclusion visit (**Annex 1**).

The inclusion visit will take place within the first 3-month period after the diagnosis of T2DM. This will avoid interfering with the regular clinical practice of the GP, and will also ensure that we do not interfere with the doctor's prescribing habits. Since current clinical practice includes assessment of chronic diabetes complications (foot exam, eye disease screening program, microalbuminuria), these data will be extracted by linking with electronic record databases (SIDIAP/PADRIS) or through the electronic medical record in the event that the link with the SIDIAP/PADRIS database is not possible. The pathological antecedents of interest (comorbidities), clinical and analytical variables related to diabetes before inclusion, will be collected from the clinical history and from the SIDIAP/PADRIS databases. All participants will be followed until the end of the study, transfer from the SIDIAP/PADRIS databases, specific cause of mortality (INE) or withdrawal of informed consent. In each CAP, all the family doctors of the CAP will be informed so that they transfer the management of the routine initial evaluation of their newly diagnosed cases of T2DM to the investigator of the local center; this will prevent further unnecessary routine clinical procedures from being performed at the study visit. This measure will surely improve the participation of all local general practitioners in the study. Figure **1** describes the procedures and data flow in the study.

Initial clinical evaluation of all enrolled subjects will be performed during the study enrollment visit. Questionnaires related to lifestyle will be self-administered

during the visit. An appointment will be made for routine blood and urine tests as recommended by current guidelines. The laboratory procedures will be carried out in the central laboratories of each of the areas (each care area, that is, Lleida and Barcelona, has a local central laboratory). All routine clinical biochemical parameters will be determined according to standard clinical laboratory procedures. Additional laboratory procedures for the determination of C-peptide and antibodies that are considered non-routine parameters will also be carried out in local laboratories, and the cost is included in the project budget, with the same urine and blood sample as the initial analysis. This logistical organization provides a unique framework in which all data is included in the usual flow of routine clinical information. Thus, the usual clinical team and the participants will have access to all the data generated during the study. In addition, after the study visit, the local investigator will ensure that the participant gets an appointment for routine evaluation of foot and eye complications at the CAP. The data of these routine procedures are readily available and will be obtained from the electronic medical record.

5.6.3 Samples

Blood and urine samples will be obtained from all fasting study participants. Each biobank belongs to the same institution as the local laboratories, and the transfer of biological samples from the local laboratories to the biobanks will be done automatically. The samples will be sent and stored (finally deposited) to the central Biobank of the promoter at the end of the study. The visual diagram of the circuit for sending samples to the biobank is presented in **Figure 1.**

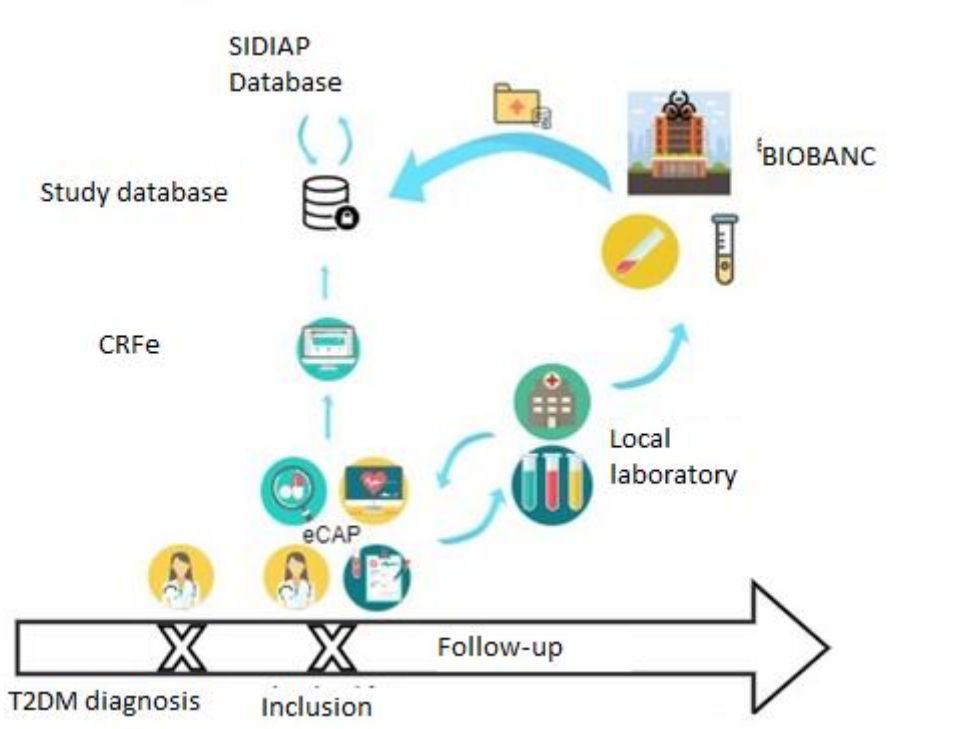

Figure 1. Diagram of study procedures

5.7 Variables

This study aims to include as many baseline phenotypic features as realistically possible. Next, we describe the variables that are considered essential, along with other additional markers that have been selected and that can help to carry out a more in-depth phenotyping of the study participants. This includes lifestyle questionnaires, liver fat and fibrosis grade, advanced lipoprotein profile, metabolites, and genetic markers; all of them have been widely used by our group in other studies.

5.7.1 Main variables

The six main variables to assess the outcome of the study are: age, anti-GAD antibodies, HbA1c, BMI, beta cell function (HOMA2-ß) and homeostatic model of insulin resistance assessment-2 (HOMA2-IR). These variables should allow the identification of the 5 relevant subgroups (clusters) of subjects initially described in the original study by Ahlqvist et al [25], ie SAID, severe autoimmune diabetes;

SIDD, severe insulin deficiency diabetes; SIRD, severe insulin-resistant diabetes; MOD, mild obesity-related diabetes; MARD, mild age-related diabetes.

1. biomarkers: Antibodies against islet cell antigens (GADA65 and IA-2) will be measured with two commercially available ELISA kits: GAD65 autoantibody ELISA and IA-2 autoantibody ELISA. The normal ranges of each local laboratory will be accepted.
2. HOMA2-IR and HOMA2-ß will be calculated with a calculator from Diabetes Trials Unit, University of Oxford: HOMA Calculator . (Diabetes Care. 1998; 21: 2191-2; this calculator is also available at: https://www.dtu.ox.ac.uk/homacalculator/).
3. C-peptide is the parameter used in previous studies on clusters in T2DM. C-Peptide will be determined using the CLIA Immunoassay Kit Procedure Analyzer High sensitivity direct chemiluminescent .

5.7.2 Other relevant variables

For a detailed description of sociodemographic variables and other relevant clinical variables (including clinical laboratory analysis), see Table 1.

Liver damage biomarkers: liver damage biomarkers will complement the characterization of metabolic complications associated with T2DM. See the detailed description in Table 1.

5.8 Specific sub-studies

The current design will be leveraged to address additional questions to further expand the phenotypic, metabolic, and genetic characterization at the time of T2DM diagnosis. Taking advantage of our previous experience and the available resources, we include a proposal of sub- studies that are clearly feasible. These sub-studies will be carried out in all the participants of the main study, except for the sub-study on subclinical atherosclerosis, which will be carried out only in centers with availability of carotid and femoral ultrasound equipment (Research Unit of Lleida).

5.8.1 Sub-study on diet and physical activity

We will assess physical activity and dietary pattern in all study subjects through validated questionnaires (see detailed description in Table 1) self-administered or by telephone.

Physical activity: it will be evaluated using the International Physical Activity Questionnaire validated in Spain (IPAQ) short version [31]. The IPAQ short form contains 7 questions that assess four domains: leisure-time physical activity, home and garden activities, work-related physical activity, and transportation-related physical activity. The IPAQ short form asks about three specific types of activities performed in the four domains listed above; These are light-intensity activities like walking, moderate-intensity, and vigorous-intensity.

The activity is classified as low, medium and high, according to the estimated energy expenditure with each activity. Activities with a metabolic equivalent task (MET) of 8 METs are classified as vigorous; moderate are those activities with an energy expenditure of 4 METs, and light are those activities with an energy expenditure of 3.3 METs, such as walking. Additionally, physical activity is classified as sedentary or active if study participants engage in some form of moderate physical activity for at least 30 minutes/day.

Dietary pattern: adherence to the Mediterranean Diet will be evaluated with the 14-item questionnaire (MEDAS) [32]. This questionnaire is an adaptation of the previously validated 9-item questionnaire [33], which is used in the control group of the PREDIMED-Plus study. It contains 12 questions about the frequency of food consumption, of which 2 questions are related to the habitual consumption of typical foods of the Spanish Mediterranean Diet. Each question is scored from 0 to 1. Low adherence is defined as a MEDAS of less than 9 points, and high adherence with a MEDAS of 9 points or more.

5.8.2 Sub-study on advanced lipoprotein profiling

A serum sample will be used for analysis of the advanced lipoprotein profile (Liposcale ®) by two-dimensional diffusion-ordered 1H-NMR spectroscopy

(DOSY) (J Lipid Res.2015 ;56: 737–46). This procedure evaluates lipid concentrations (i.e., triglycerides and cholesterol), the size and number of particles of 3 different classes of lipoproteins (VLDL, LDL, and HDL), as well as the number of particles of 9 subclasses (VLDL large, medium and small, LDL and HDL). A more detailed description of this methodology is available in our recent publication (Cardiovasc Diabetol 2020;19:126).

5.8.3 Sub-study on subclinical carotid and femoral atherosclerosis

The Lleida Research Unit will carry out a sub-study on the prevalence of subclinical carotid and femoral atherosclerosis (the expected recruitment in this health area is 509 participants). This sub-study is completely feasible since this research unit is already leading a large cohort study (ILERVAS) in the region. Researchers in this area have all the experience and resources necessary to carry out this sub-study. We estimate that a total of 15% of newly diagnosed subjects will already harbor established clinical cardiovascular disease. All other subjects will be invited to participate in this optional procedure. This will be done in a specific visit in the same CAP. Carotid and femoral ultrasound will be performed using B-mode ultrasound and color Doppler. The presence of atherosclerotic plaque will be assessed in ten vascular territories: internal artery, bulb and common carotid, common and superficial femoral arteries. Atheromatous plaque is defined as an intima media thickness (cIMT) >1.5 mm protruding into the lumen of the vessel, according to the ASE Consensus Statement (J Am Soc Echocardiogr. 2008; 21: 93-111) and methodology in previous publications of the group (Palanca et al. Diabetol ; 2019 ;18:93).

5.8.4 Sub-study on initial antidiabetic treatment and other diabetes-related therapies

Given that our organization has access to all the data generated during routine practice, we will also evaluate the use of antidiabetic therapies in the treatment of diabetes at the onset, and we will be able to analyze the degree of adherence to current guidelines (Red GDPS T2DM Guide; 2018; Guidelines for harmonization

pharmacotherapy , CatSalut ; 2019). eCAP data will be extracted for each participant for the first 6 months after diagnosis. Data on the type of medication, prescription frequency, dispensing data will be recorded. In addition, we will also collect data on lipids and lipid-lowering treatment.

5.8.5 Metabolomic and genetic sub-studies related to diabetes

The metabolomic characteristics will provide us with a more complete physiological and genetic view of the disease through the profile of metabolites and genetic markers in blood and urine samples. Metabolomic markers will be analyzed (free methionine and methionine sulfoxide, MetOx-148/Met-148 ratio in apo-A1, NMR of circulating lipoproteins and glycoproteins, specifically parameters of LDL and VLDL particles, and circulating GlycA glycoproteins, and GlycB) of T2DM and associated complications. Genetic characteristics that cause defects in insulin secretion, insulin action, or autoimmune destruction of ß cells will be analyzed. In case of clinical suspicion, possible dominant (eg HNF1A, HNF4A, GCK) or recessive (eg WFS1, INS, EIF2AK3) monogenic mutations will be studied, depending on the suspected clinical phenotype.

Table 1. Study variables.

Socio-demographic variables		
Variable	Source	Definition, (Value)
Age	eCAP *	Main, (≥18)
Sex	eCAP *	Others, (male/female)
MEDEA Index	SIDIAP/PARENTS *	Others, (R, U, U1-U5)
ABS	eCAP *	others,
Education level	eCAP *	Others, (Illiteracy, No primary studies, Primary studies, Secondary studies, University studies)
Toxic habits		

tobacco use	eCAP / SIDIAP / PADRIS **	Others, (Smoker, non-smoker, ex-smoker)
alcohol consumption	eCAP / SIDIAP / PADRIS **	Others, (Risk consumer, non-consumer)
drug use	eCAP / SIDIAP / PADRIS **	Others, (Consumer, non-consumer)
Comorbidities		
diabetes type	eCAP / SIDIAP / PADRIS *	Main, (T2DM, DM1, Gestational, secondary, others, ICD10: E10.xx-E13.xx, O24.xx)
Time of evolution of diabetes	eCAP *	Main, (months)
dyslipidemia	eCAP / SIDIAP / PADRIS **	Others, (ICD10: E78.xx)
Hypertension	eCAP / SIDIAP / PADRIS **	Others, (ICD10: I10.xx, I12.xx, I13.xx, I15.xx)
Peripheral arterial disease	eCAP / SIDIAP / PADRIS **	Others, (ICD10: I70.xx, I73.xx)
Heart failure	eCAP / SIDIAP / PADRIS **	Others, (ICD10: I50.xx)
ischemic heart disease	eCAP / SIDIAP / PADRIS **	Others, (ICD10: I20.xx, I25.xx)
stroke	eCAP / SIDIAP / PADRIS **	Others, (ICD10: G45.xx, G46.xx, I63.xx, I69.xx)
Liver disease (excluding steatosis)	eCAP / SIDIAP / PADRIS **	Others, (ICD10: K70.xx, K71.xx, K72.xx, K73.xx,K74.xx,K75.xx, K77.XX)
Steatosis hepatic	eCAP / SIDIAP / PADRIS **	Others, (ICD10: K75.81, K76.XX)
Renal insufficiency	eCAP / SIDIAP / PADRIS **	Others, (ICD10: N17.xx, N18.XX, N19)
acute pancreatitis	eCAP / SIDIAP / PADRIS **	Others, (ICD10: K85.xx)
diabetic nephropathy	eCAP / SIDIAP / PADRIS **	Others, (ICD10: E11.2, N08.3, R80)
Diabetic neuropathy	eCAP / SIDIAP / PADRIS **	Others, (CIE10: E11.4, G59, G56, G62, G63, G90, G99)
Diabetic retinopathy	eCAP / SIDIAP / PADRIS **	Others, (CIE10: E11.4, G59, G56, G62, G63, G90, G99)
Causes of Mortality	eCAP / SIDIAP / PADRIS / INE**	Others, (ICD10: R99), specific causes
Exploration variables		
heart rate	eCAP / SIDIAP / PADRIS **	Others, (EK401, [batecs /min])

BMI (Body Mass Index)	eCAP / SIDIAP / PADRIS **	Others, (TT103 [kgs /m2])
Abdominal perimeter	eCAP / SIDIAP / PADRIS **	Others, (EL401-[cm])
Diastolic Blood Pressure	eCAP / SIDIAP / PADRIS **	Others, (EK202 - [mmHg])
Systolic Blood Pressure	eCAP / SIDIAP / PADRIS **	Others, (EK201 - [mmHg])
/ Weight	eCAP / SIDIAP / PADRIS **	Others, (TT102 - [kgs])
Height / Size / Height	eCAP / SIDIAP / PADRIS **	Others, (TT101 -[cm])
Clinical variables		
GADA65 AI-2	eCAP / SIDIAP / PADRIS *	Principal
C -peptide	eCAP / SIDIAP / PADRIS *	Principal
Creatinine: Serum Creatinine / Creatinine / Creatinine Ion / Creatinemia	eCAP / SIDIAP / PADRIS **	Others, [c.subst . (serum)]
Filtrat (Estimated CKD-EPI): Estimated Glomerular Filtration (CKD-EPI)	eCAP / SIDIAP / PADRIS **	Others
HDL Cholesterol: Serum HDL Cholesterol / HDL / HDL-C	eCAP / SIDIAP / PADRIS **	Others, [c.subst . (serum)]
LDL Cholesterol: Serum LDL Cholesterol / LDL / / LDL-C	eCAP / SIDIAP / PADRIS **	Others, [c.subst . (serum)]
Cholesterol: Serum Cholesterol	eCAP / SIDIAP / PADRIS **	Others ,[c.subst . (serum)]
Triglyceride -: Triglycerides Serum / Triacylglycerols	eCAP / SIDIAP / PADRIS **	Others ,[c. Subst . (serum)]
Glycohemoglobin (a1c) - : Glycosylated Hemoglobin / hba1c	eCAP / SIDIAP / PADRIS **	Principal , [fr.subst .]
TSH	eCAP / SIDIAP / PADRIS **	
Glucose -: Serum Glucose / Glycaemia/Glycaemia	eCAP / SIDIAP / PADRIS **	Others ,[c.subst . (serum)]
leukocytes	eCAP / SIDIAP / PADRIS **	Others, [c.subst . (serum)]
red blood cells	eCAP / SIDIAP / PADRIS **	Others ,[c.subst . (serum)]
Ers (San)-Hemoglobin;c.massa	eCAP / SIDIAP / PADRIS **	Others, CI013

(CCMH)		
hematocrit	eCAP / SIDIAP / PADRIS **	others,
Medium corpuscular volume	eCAP / SIDIAP / PADRIS **	others,
Mean Corpuscular Hemoglobin	eCAP / SIDIAP / PADRIS **	others,
Red blood cell distribution width	eCAP / SIDIAP / PADRIS **	others,
platelets	eCAP / SIDIAP / PADRIS **	others,
mean platelet volume	eCAP / SIDIAP / PADRIS **	others,
neutrophils	eCAP / SIDIAP / PADRIS **	others,
lymphocytes	eCAP / SIDIAP / PADRIS **	others,
monocytes	eCAP / SIDIAP / PADRIS **	others,
eosinophils	eCAP / SIDIAP / PADRIS **	others,
Basophils	eCAP / SIDIAP / PADRIS **	others,
eosinophils	eCAP / SIDIAP / PADRIS **	others,
Basophils	eCAP / SIDIAP / PADRIS **	others,
Urat (serum); c. subst .	eCAP / SIDIAP / PADRIS **	Others , Q54485
To the girl aminotransferase (serum); c.cat.	eCAP / SIDIAP / PADRIS **	Others, Q02185
Aspartate aminotransferase (serum); c.cat	eCAP / SIDIAP / PADRIS **	Others, Q07585
To the girl aminotransferase (plasma); c.cat.	eCAP / SIDIAP / PADRIS **	Others
Gamma-glutamyltransferase	eCAP / SIDIAP / PADRIS **	Others
urine creatinine	eCAP / SIDIAP / PADRIS **	Others
Microalbuminuria	eCAP / SIDIAP / PADRIS **	Others
Albumin / Creatinine	eCAP / SIDIAP / PADRIS **	Others , [quotient dough /subst. (urine)]
urine pH	eCAP / SIDIAP / PADRIS **	Others
urine density	eCAP / SIDIAP / PADRIS **	Others

proteinuria	eCAP / SIDIAP / PADRIS **	Others
Glucosuria	eCAP / SIDIAP / PADRIS **	Others
Ketonuria	eCAP / SIDIAP / PADRIS **	Others
Bilirubin in urine	eCAP / SIDIAP / PADRIS **	Others
Hematuria	eCAP / SIDIAP / PADRIS **	Others
nitrites in urine	eCAP / SIDIAP / PADRIS **	Others
Urobilinogen urine	eCAP / SIDIAP / PADRIS **	Others
urine leukocytes	eCAP / SIDIAP / PADRIS **	(Negative positive)
Leukocyte concentration in urine	eCAP / SIDIAP / PADRIS **	Others
erythrocytes in urine	eCAP / SIDIAP / PADRIS **	Others
Liver fat index (FLI)	eCAP / SIDIAP / PADRIS **	Others
Liver Fibrosis Index (NFS)	eCAP / SIDIAP / PADRIS **	Others
FIB4 Index	eCAP / SIDIAP / PADRIS **	Others
HOMA2-IR HOMA2-ß	eCAP / SIDIAP / PADRIS *	Principal
Variables Sub-study on diet and physical activity		
IPAQ	Questionnaire	Seven questions related to physical activity
Activities with a metabolic equivalent task (MET)	Questionnaire	Physical activity mild, moderate and high
Adherence to the Mediterranean Diet	Questionnaire	Contains 12 questions about the frequency of food consumption and 2 questions about the consumption of typical foods of the Mediterranean Diet
Variables Sub-study on advanced lipoprotein profile (Liposcale ®)		
VLDL LDL HDL	1H-NMR spectroscopy / (DOSY) biological samples	the size and number
VLDL large medium and small, LDL and HDL	1H-NMR spectroscopy / (DOSY) biological samples	number of particles of 9 subclasses

GlycA , and GlycB	1H-NMR spectroscopy / (DOSY) biological samples	number of particles
Variables Sub-study on subclinical carotid and femoral atherosclerosis B-mode and color Doppler ultrasound		
Carotid common drag < 50%	eCAP /SIDIAP	Altered / Normal
Carotid common right > 50%	eCAP / SIDIAP	Altered / Normal
Carotid common left <50%	eCAP / SIDIAP	Altered / Normal
Carotid common left > 50%	eCAP / SIDIAP	Altered / Normal
Bulb right <50%	eCAP / SIDIAP	Altered / Normal
Bulb right > 50%	eCAP / SIDIAP	Altered / Normal
Bulb left <50%	eCAP / SIDIAP	Altered / Normal
Bulb left > 50%	eCAP / SIDIAP	Altered / Normal
Right internal carotid <50%	eCAP / SIDIAP	Altered / Normal
Right internal carotid artery > 50%	eCAP / SIDIAP	Altered / Normal
Left internal carotid artery <50%	eCAP / SIDIAP	Altered / Normal
Left internal carotid artery> 50%	eCAP / SIDIAP	Altered / Normal
Right external carotid <50%	eCAP / SIDIAP	Altered / Normal
Right external carotid artery > 50%	eCAP / SIDIAP	Altered / Normal
carotid Left <50%	eCAP / SIDIAP	Altered / Normal
carotid left> 50%	eCAP / SIDIAP	Altered / Normal
Common femoral right	eCAP / SIDIAP	Altered / Normal
Common femoral left handed	eCAP / SIDIAP	Altered / Normal
Right superficial femoral	eCAP / SIDIAP	Altered / Normal
Left superficial femoral	eCAP / SIDIAP	Altered / Normal
Carotid right not valued	eCAP / SIDIAP	If not
carotid not valued	eCAP / SIDIAP	If not
Right femoral not valued	eCAP / SIDIAP	Otherwise

Femoral Left unfavorable	eCAP /SIDIAP	Otherwise
Right intima-media thickness (IMT)	eCAP /SIDIAP	Value
Left intima-media thickness (IMT)	eCAP /SIDIAP	Value
Medium intima-media thickness (IMT)	eCAP /SIDIAP	Value
carotid plates	eCAP /SIDIAP	Number plates
plate type	eCAP /SIDIAP	(Hypoechoic, Hyperechoic, Calcified)
femoral plates	eCAP /SIDIAP	Number plates
EA0	eCAP /SIDIAP	1
EA2	eCAP /SIDIAP	3
EA3	eCAP /SIDIAP	4
Sub-study on initial antidiabetic treatment and other diabetes-related therapies		
Drugs antidiabetics	eCAP / SIDIAP / PADRIS **	Others, (ATC/DDD: A10, dosage, frequency, dispensing)
Lipid-lowering	eCAP / SIDIAP / PADRIS **	Others, (ATC/DDD: C10, posology, frequency, dispensation)
Antiaggregants	eCAP / SIDIAP / PADRIS **	Others, (ATC/DDD: B01AC, posology, frequency, dispensation)
Anticoagulants	eCAP / SIDIAP / PADRIS **	Others, (ATC/DDD: B01AA, posology, frequency, dispensation)
hypotensive	eCAP / SIDIAP / PADRIS **	Others, (ATC/DDD: C07, posology, frequency, dispensation)
Metabolomic and genetic sub-studies related to diabetes		
Metabolomics	biological samples	free methionine, methionine sulfoxide, MetOx-148/Met-148 ratio, apo-A1
Genetics	biological samples	HNF1A, HNF4A, GCK, WFS1, INS, EIF2AK3

***Baseline, **Baseline and follow-up, eCAP : Clinical history**

6 . DATA COLLECTION AND ANALYSIS

6.1 Data collection schedule

The study period is estimated to be 36 months (from January 1, 2022 to December 31, 2024), although recruitment is expected in the first 24 months, and can be extended until the expected recruitment is achieved. However, these times may be modified by the deadlines of the administrative process of starting up the study or by the situation generated by the current COVID-19 pandemic.

6.2 Source of information

The source of information will in all cases be the patient's clinical history and the electronic Case Report Forms (eCRF) for the physical activity and diet questionnaires.

6.2.1 Information collected

This study proposes a prospective collection of information, that is, baseline data will be obtained in a single visit. The monitoring variables will be collected by linking with the SIDIAP/PADRIS databases. In the event that the link with the SIDIAP/PADRIS databases is not possible for technical reasons (non-ICS centers), the available monitoring variables will be collected from the electronic medical record.

Next, all the variables that will be collected during the study visit and during follow-up are described, as long as they are available in the patient's clinical history or can be obtained during the interview between the doctor and the patient during said visit. without applying any type of diagnostic or therapeutic intervention outside of routine clinical practice.

The data will be collected by completing an ECRF (Annex **1)**, in order to record the information available in the patient's clinical history. The linking of patients with large SIDIAP/PADRIS databases will be carried out annually. The link will be made through the encrypted CIP (encrypted) of the patients included in the study that have allowed the data to be linked with the signed informed consent. The linking of the

encrypted CIPs with INE (specific causes of mortality) is managed through SIDIAP. The identification of the complications of T2DM, comorbidities and treatments will be carried out using the SIDIAP/PADRIS population databases, mainly where possible, which corresponds to the data of the registered clinical history of Primary Health Care in Catalonia. In the case of non-ICS centers, the collection of monitoring variables will be done directly from the clinical course of the participants.

Table 2. *Information collection scheme during the study.*

Information related to the patient	Preselection	Visit inclusion	Tracing SIDIAP/PADRIS/INE/electronic clinical record
Review of selection criteria	x	x	
Informed consent	x	x	
Data recording at study visit		x	
Socio-demographic data [a]			
Clinical [datab]		x	x
History of T2DM		x	
Comorbidities of interest		x	x
Specific mortality			x
Laboratory data [c]		x	x
Smoking habit/alcohol/drug use		x	
Pharmacological treatment		x	x
Diet and physical activity variables		x	
Variable subclinical carotid and femoral atherosclerosis [and]		x	
Biological samples		x	
Variables advanced lipoprotein profile (Liposcale) [f]			
Metabolomic and genetic variables [g]			

[to] Age, sex, ABS, the MEDEA will be collected from SIDIAP/PADRIS

[b] Weight, height, (auto calculated BMI), and abdominal circumference, SBP/DBP and heart rate

[c] The laboratory parameters, the main variables will be collected during the inclusion visit, the rest of the laboratory parameters will be collected from SIDIAP/PADRIS

[d] Treatment for inclusion will be collected at the visit and from the SIDIAP/PADRIS database Includes prescription, posology, dose

[e] They will be collected at the inclusion visit only in centers that have equipment to assess atherosclerosis.

[f] They will be collected at the end of the study, from the blood samples obtained at the inclusion visit

[g] They will be collected at the end of the study, from the blood/urine samples obtained at the inclusion visit.

6.2.2 Recruitment procedure for possible candidates

The health professional who sees for the first time a patient who meets the inclusion criteria and none of the exclusion criteria, will inform the principal investigator of the center (IPC) or the collaborating investigators (CI) with authorization to carry out this task. After that, the IPC/IC will make an appointment with the patient in a time not to exceed one week in order to obtain informed consent and include the patient in the study and, subsequently, proceed to physical examination, collection of history data, schedule the analytical study and collection of biological samples. All study procedures will be carried out only by the center's research team.

6.3 Data analysis

We will perform the cluster analysis (k- means and hierarchical clustering) following the method described by Ahlqvist et al [25]. New adult-onset diabetes subgroups and their association with outcomes: a cluster analysis based on six-variable data. Men and women will be grouped separately to avoid stratification due to sex-dependent differences in group variables. Continuous variables will be centered on the mean and standardized, and continuous variables greater than five standard deviations from the mean will be excluded; categorical variables will be included as binary variables. Two-step clustering will be used, in which the first step estimates the optimal number of clusters based on the width of the silhouette and the second step performs hierarchical clustering using the log probability as the distance measure and the Bayesian criterion of Schwarz for grouping. The stability of the groups will be evaluated by resampling the data set 2000 times, and calculating the Jaccard similarities with the original group. Generally, stable clusters should produce a Jaccard similarity greater than 0.75. Once the clusters are defined, we will assign the same names as in the original study based on the distribution of cluster features. We will assign each subject to their cluster according to their Euclidean distance from each cluster center. Differences between subtypes will be assessed by descriptive analysis and tested by Fisher's exact test, ANOVA test, and post hoc comparisons. A p value of less than 0.05 will be considered statistically significant in the phenotypic analyzes

and corrected for multiple comparisons. In addition, the Benjamini & Hochberg correction will be used to determine the significance of multiple tests (by variables). Statistical analysis will be performed using the R statistical software package (version > 3.6.3), and cluster analysis will be performed in the fpc (Flexible Procedures for Clustering) package, version > 2.2.

6.4 Sample size

This is a descriptive study without confirmatory analysis of a specific hypothesis. Therefore, the sample size has not been calculated according to an expected hypothesis or a specific measure of desirable association. Additionally, we have evaluated the number of subjects in previous studies similar to the current one in other countries (references 26 and 29, in Background); all used a similar approach with a sample size similar to or smaller than that proposed for this study. As stated above, we plan to recruit a total of 1,200 participants, spread across 11 centers over a 24-month recruitment period. This estimate has been based on an expected incidence of 4 T2D per 1000 person-years according to our own study in our population, as part of a recent international collaboration [34].

6.5 Limitations of the design, the source of information and the methods of analysis

One of the limitations may be the relative representativeness of the participants included in this study. We opt for a realistic approach that allows us to develop an initial project that strikes the most appropriate balance between viability and sufficient representativeness. We believe that this proposal addresses this issue and may allow us to cooperate in this field with other national and international researchers. In fact, due to lack of time and limited funding for a call like the current one, we have not been able to undertake a large collaborative study in the country; however, we plan to approach other Spanish researchers in the field to complement our research with other regions of the country.

Another limitation lies in the cluster analysis, which is an exploratory unsupervised analysis. It does not assume any underlying knowledge of the pathophysiology of the

condition and needs clinical interpretation. Some forms of cluster analysis may produce slightly different results each time the statistical analysis is run.

7. DATA HANDLING

7.1 Data collection and confidentiality

The data will be entered by the researchers themselves and/or authorized personnel directly in ECRF (Annex **1)**. In order to guarantee the confidentiality of the data obtained from the study, only the local researcher and his team of collaborators (after signing **Annex 3)**, the sponsor or the person designated by him, the CEI will have access to them. , the relevant health authorities and those responsible for data analysis.

Regarding the ECRF, each researcher will be given a username and a password made up of between 4 and 6 digits by means of a closed document. These codes are considered confidential and non-transferable, and are subject to the same confidentiality rules as the rest of the documents, including the protocol. It is the researchers' responsibility to keep their passwords secret and not to disclose them to third parties. To obtain access to the ECRF, the researchers must sign the study researcher commitment and confidentiality commitment (**Annex 3**).

The content of the ECRF, as well as the documents generated during the study and the database, will be protected from unauthorized use by people outside the research and, therefore, will be considered strictly confidential and will not be disclosed to third parties. In addition, the ECRF will be automatically disconnected in the case of lack of activity for more than 10 min, in this way the ECRF will be protected from the entry of unauthorized persons, in the event that for some reason the ECRF has been left open by the / the study researcher.

The patients included in each center will have a correlative number (code), automatically generated by the ECRF once the informed consent has been signed and the inclusion criteria in the ECRF have been completed. This patient code will be documented in the participant's clinical record and in the confidential patient inclusion record in the investigators' folder of each participating center. The linking of patients with the SIDIAP/PADRIS databases will be carried out annually. The link will be

made through the encrypted CIP (encrypted) of the patients included in the study who have allowed the data to be linked with the signed informed consent.

7.2 Mandatory requirements on the Data Protection Law regulation (EU) No. 2016/679 and Organic Law 3/2018 Protection of Personal Data and digital rights guarantee (RGPD)

7.2.1 Identification of the data and the subjects that treat them

The variables necessary to carry out the study are those specified in the variables section (**Table 1**), and will come from the participants/electronic medical record (eCAP) after obtaining informed consent and from the SIDIAP/PADRIS population databases such as it is specified in sections (5.3; 5.6; 5.7; 6.2; 7.1) The variables of the participants/electronic clinical record (eCAP) will be crossed with information from the SIDIAP/PADRIS population databases for the centers where possible (ICS) or it will be collected directly from the electronic medical record (Non-ICS centers). The information will be encrypted. Coding will be done through the ECRF as specified in section 7.1 Data collection and confidentiality.

7.2.2 Identification of the treatments and legitimizing basis of the treatments

The processing of personal data required in this study is governed by Organic Law 3/2018, of December 5, on the protection of personal data. The Institute Català de la Salut acts as data controller within the framework of this study. The data finally registered in the central database will be the property of the promoter. The identity of the participants will not be revealed to any other person except the health authorities, when required or in cases of medical emergency. The Research Ethics Committees, the representatives of the Health Authority in matters of inspection and the personnel authorized by the Promoter, may only access to check the personal data, the clinical study procedures and compliance with the standards of good clinical practice. (always maintaining the confidentiality of the information).

The Researcher and the Promoter are obliged to keep the data collected for the study for at least 5 years after its completion. Subsequently, the personal information will only be kept by the patient's health center, and by the promoter for other scientific research purposes if the patient has given their consent to do so, and if permitted by applicable law and ethical requirements.

The Information System for the Development of Research in Primary Care, (SIDIAP/PADRIS), managed by the IDIAP Jordi Gol i Gurina foundation, when crossing patient data, must have their data (CIPS) for the provision of the service, so it will act as the person in charge of the treatment. The project database will be hosted on the SIDIAP/PADRIS servers, for which it will act as data processor.

The analysis of the coded data will be carried out in the research support units of Lleida and Barcelona. The final encrypted database without the possibility of identifying the subjects will be hosted on the servers of these units that comply with the Catalan Institute of Health data protection regulations.

Legal basis of the treatment: RGPD: 6.1 a) the interested party gave their consent for the treatment of their personal data for one or more specific purposes.

7.2.3 Tools used to process the data

To carry out the project, the REDCAP platform will be used, hosted on the promoter's servers. Research Institute of the Hospital de la Santa Creu i Sant Pau, and which has the security measures determined by the institution. The data is stored on the local web server where the organization has installed the software and is therefore only accessible on computers that have a trusted connection via VPN and secure credentials (certificates, RSA keys, or complex passwords). A system has been incorporated so that only the application service can send the data to the back office, through a firewall that only allows requests from the application's IP addresses. The web server has the X- Frame - Options HTTP header configuration enabled with the value " same-origin " to prevent clickjacking attacks.

7.2.4 Procedure for linking data with SIDIAP/PADRIS

The information registered in the SIDIAP/PADRIS databases is pseudonymized. For this study, a standardized procedure already established in SIDIAP/PADRIS will be used, which allows data to be crossed in a controlled and confidential manner. This procedure is also a technical and functional separation between the research team and those who participate in carrying out pseudonymization. Pseudonym data They will only be accessible to the research team according to the Guidelines established in Organic Law 3/2018, of December 5, on the Protection of Personal Data and the guarantee of digital rights.

The procedure for comparing data from external sources with SIDIAP/PADRIS data is explained below:

1. Research team encrypts the identifiers of the patients (CIP) who have given informed consent for linking the data through a password.
2. Research team sends the data (encrypted identifier + variables) to the ICS (via SIDIAP/PADRIS)
3. Research team sends the key to the ICS (via SIDIAP/PADRIS)
4. The ICS decrypts the ECAP patient identifier with the investigator team's key
5. ICS links information
6. The ICS exchanges the encrypted identifier of the external source for the SIDIAP/PADRIS identifier
7. The ICS sends the data to SIDIAP/PADRIS
8. SIDIAP/PADRIS sends the data to the research team

*The data crossing for specific mortality with INE is managed in the same way through SIDIAP.

7.2.5 International data transfers

Similar studies have been carried out in other countries, such as Sweden, Germany, UK [25,26]. It is possible that our group conducts collaborative studies with other national and international research groups related to diabetes. In no case, raw and

confidential data of the participants will be sent. International transfers will be made as regulated in chapter V of the RGPD (article 44, 45, 46 and 49 a: the interested party has explicitly given their consent to the proposed transfer, after having been informed of the possible risks for them of said transfers). transfers due to the absence of an adequacy decision and adequate guarantees). The transfer of data from our study would only be for patients who have accepted this option in the informed consent. If we transfer the encrypted data within the EU to the entities of our group, to service providers or to scientific researchers who collaborate with us, the participant's data will be protected with safeguards such as contracts or other mechanisms by the data protection authorities. data.

7.2.6 Identification of treatments that may pose a high risk to the rights and freedoms of the participants in the research project.

In accordance with the provisions of article 35 of the RGPD, the project does not meet the necessary characteristics that require the corresponding impact assessment to be carried out. As secondary objectives, it is planned to study the genetic factors related to diabetes mellitus, in order to better characterize the disease, the data will be coded. However, the impact assessment related to data protection for research projects and clinical innovation in health is attached in Annex 9, it has been evaluated as low risk.

7.2.7 Rights of interested persons

In accordance with regulation (EU) No. 2016/679 and Organic Law 3/2018 Protection of Personal Data and digital rights guarantee, the patient can limit the processing of data that is incorrect, request a copy or transfer it to a third party (portability) the data you have provided for the study. To exercise these rights, you will contact the main researcher, the researchers at the study site or the Data Protection Delegate of the center as long as this figure exists in the center or promoter Institut de Recerca de l'Hospital de la Santa Creu i Sant Pau – IIB Sant Pau Sant Quintí , 89, 08041 Barcelona, Tel: 93 553 78 69, email dpo_ir@santpau.cat , or with the Spanish Data Protection Agency at Jorge Juan, 6, 28001 Madrid (https: // www .aepd.es / es) or the Catalan Dades Protection Authority (https://apdcat.gencat.cat).

If the participant wants to know more about it, they can contact the Data Protection Delegate of the institution or the promoter Institut de Recerca de l'Hospital de la Santa Creu i Sant Pau – IIB Sant Pau Sant Quintí , 89, 08041 Barcelona, Tel: 93 5537869, email dpd@santpau.cat, or with the Spanish Data Protection Agency at Jorge Juan, 6, 28001 Madrid (https: // www .aepd.es/es) or the Catalan Data Protection Authority (https://apdcat.gencat.cat).

7.2.8 Data validation and quality control

All data received through the use of the ECRF will be stored, and will be subjected to the appropriate work procedures to comply with the FDA 21 ECRF standard. Part 11 guaranteeing, therefore, their confidentiality, security and authenticity.
The adaptation of the 21 CFRe standard Part 11 ensures that data received via electronic transmission is as valid as originals received on paper. Said standard sets the rules for the use of electronic data and defines the requirements of all systems for their collection, storage, maintenance and security.

All the data registered by the researcher in the system will be submitted to an automatic validation program that will review them and, accordingly, discrepancy reports will be generated to facilitate the correct and complete registration of the data (automatic queries). In addition, the data entered will be reviewed and verified by the study monitor. The corrections of the ECRF will be made by the researchers themselves. Raw data should be available and ready for review during scheduled monitoring visits.

Likewise, all the documents of this study must be available to the duly qualified personnel responsible for the development of the audits or inspections that, if they occur, are carried out. The verification of the ECRF data must be carried out by direct inspection of the original documents.

Biological samples

Collection and use

The samples obtained will be labeled with a code to maintain confidentiality and will initially be processed by the Biobank IRBLleida and the Vall University Hospital Biobank d'Hebron (HUVH Biobank). At the end of the study they will be deposited in the Biobank – IIB Sant Pau.

Responsible for data processing

Biobank Sant Pau Biomedical Research Institute
Santa Cruz and San Pablo Hospital Research Institute Foundation (hereinafter, FIRHSCSP or Foundation). Registered office: Calle San Antonio María Claret, 167, 08025 Barcelona. Registered in the Registry of Foundations of the Government of Catalonia, with number 708. CIF: G-60136934. Telephone: 93.553.76.13. Contact email: biobanc@santpau.cat .
Promoter: The principal investigator (Dr. Dídac Mauricio DNI: 46525502K) of the study and the Promoter - Research Institute of the Hospital de la Santa Creu i Sant Pau - IIB Sant Pau, c / Sant Antoni Maria Claret, 167 08025 Barcelona, Tel: 93 553 78 69. Contact email: dmauricio@santpau.cat

Purposes

Your data will be processed solely to carry out research to advance knowledge of diabetes (advanced lipo -protein profile, metabolomics and genetics sub studies).

Legal basis of the treatment

RGPD: 6.1 a) the interested party gave their consent to the processing of their personal data for one or more specific purposes.

Rights of interested persons

The same rights specified in point: 7.2.7

8 . WORK PLAN (tasks, milestones and study timeline):

Month -2

Participating centers: IR-HSCSP, CAPs of Lleida and Barcelona, USR Barcelona and Lleida,

Meeting of the researchers for the coordination and resolution of doubts. Preparation of the logistical aspects for the start-up of the study (month -2).

IR-HSCSP

Presentation to CEIC (months -2, -1).

1 months

Participating centers: IR-HSCSP, CAPs of Lleida and Barcelona, USR Barcelona and Lleida,

Inception visits with participating centers

Months 1-24

Participating centers: CAPs of Lleida and Barcelona, USR Barcelona and Lleida,

Recruitment of participants

Months 4, 8, 12, 16, 20, 24

Participating centers: IR-HSCSP, USR Barcelona and Lleida

Data monitoring

Months 1-24

Participating Centers: IR-HSCSP

Collection of biological samples in Biobank

Months 14,15,16,27,28,29

Participating centers: USR of Barcelona

Data management and linkage with SIDIAP/PADRIS databases

Months 17-20, 30-33

Analysis of data

Months 22-24, 34-36

Participating centers: IR-HSCSP, USR Barcelona and Lleida,

Elaboration of the report for the data collected during 2 years

Interpretation and discussion of the results

Preparation of communications and manuscripts

Researchers meeting to discuss the results and close the study.

9 . ETHICAL ASPECTS:

The present study complies with all the ethical aspects and protection of the participating subjects by complying with the ethical precepts formulated in Royal Decree 957/2020 and the Biomedical Research Law 14/2007, of July 3, and in the Declaration of Helsinki and all your reviews.

Since this study is observational in nature, patient participation is considered to carry minimal risk; that is, the risk is similar to that which the patient would have in clinical practice without mediating their participation in the study. It is an unconditional prerequisite for the participation of a patient to obtain his consent after having been informed by the researcher about it both verbally and in writing in a vocabulary that allows its content to be completely readable and understandable for the patient (Annex **5)**.

The protocol will be submitted to the evaluation of a CEI prior to the start of the inclusion of the patients. Any data required by the protocol may be subject to audits by the promoter, independent organizations and/or competent authorities, but the confidentiality of the data in accordance with the aforementioned law will always be an essential condition.

The participating subject may at any time revoke their consent for the use of their data in the analysis, without giving cause and without this resulting in any liability or damage whatsoever.

Before accepting and signing the investigator's commitment, the participating clinical investigators must ensure that their participation in the study does not interfere with their prescription habits or their care duties.

9 .1 Benefit-risk assessment of the research

This project affects the priority of developing personalized medicine. This project adopts a research approach that is clearly aimed at improving clinical decision-making that affects the people we serve. The potential impact may be equivalent to that mentioned for diabetes itself.

The patient will not suffer any risk in this study since it is completely observational and, therefore, will not be subjected to any intervention, neither dietary nor pharmacological treatment. The present study is limited to an anonymized data record in a database that does not allow access to the patient's personal data.

9.2 Ethical considerations, regarding information to subjects and informed consent

The study will be carried out strictly following the international ethical recommendations for medical research in humans. The investigator will be responsible for ensuring that the study is carried out in accordance with the standards set out in the Declaration of Helsinki.

Before starting the study, the Ethics Committee of the Hospital de la Santa Creu i Sant Pau must approve the study protocol, the information that will be given to the subject and the model of informed consent that will be used.

For linking with the SIDIAP/PADRIS databases and carrying out the study in the primary care centers of the Institut Català de la Salut - ICS, the study will be evaluated by the Jordi Gol y Gurina IDIAP Scientific and Ethics Committee.

For centers that have other reference RECs different from those previously mentioned, before starting recruitment in these centers, if necessary, the protocol will be passed for evaluation and approval by the corresponding reference RECs for these centers.

The REC will be informed of any subsequent amendments to the protocol and its opinion should be sought if a new evaluation of the ethical aspects of the study is necessary.

It is the investigator's responsibility to obtain informed consent from the patient. The patient will not be able to participate in any specific procedure of the study before obtaining their consent, or that of their legal guardian/relative when the patient is not capable of giving their consent due to their clinical situation.

Before including any subject in the study and before obtaining informed consent, the researcher or the person designated by the researcher will explain to the potential participating subject or their legal guardian/relative, the objectives, methods and potential risks of the study and any inconvenience it may cause. The explanation about the nature, scope and possible consequences of the study will be made in an understandable language. The rights related to the data protection law will be explained, as well as the possibility of being contacted in the case of future studies related to diabetes.

The potential participating subject or their legal guardian/relative must have time to consider their decision to participate in the study, and have the opportunity to ask questions. After this explanation, and before entering the study, the consent must be properly recorded by the signature of the subject or their legal guardian/relative. The informed consent will be signed in three copies, one for the participant, another for the researcher and the third copy for the biobank.

9.3 Considerations on the treatment of biological samples

Regarding the obtaining, handling, identification and storage of biological samples, the provisions of the Biomedical Research Law 14/2007, of July 3, specifically in chapters III and IV of title V, as well as provided in Royal Decree 1716/2011, of November 18, specifically in chapter I of title II, which establishes the treatment of biological samples of human origin for biomedical research purposes.

In accordance with said regulations, the obtaining of biological samples for biomedical research purposes may only be carried out when the prior written consent of the source subject has been obtained and prior information on the consequences and risks that such obtaining may entail for their health. Said consent will be revocable. The participants will be asked in the same informed consent for the study samples, also the use of the samples related to other future diabetes studies and other future non-diabetes related studies.

The biological samples of the study will be processed by the Biobank IRBLleida and the Vall University Hospital Biobank d'Hebron (HUVH Biobank) at the end of the study will be stored by the Biobank – IIB Sant Pau.

9.3 Confidentiality of data

Regarding the confidentiality of the study data, the provisions of Organic Law 3/2018, of December 5, on the Protection of Personal Data and Guarantee of Digital Rights and the General Data Protection Regulation (EU) will be followed. 2016/679.

9.4 Interference with the physician's prescribing habits

The study will not interfere with the clinical management habits of patients

10. PLANS FOR DISSEMINATION OF RESULTS:

It is expected that, from the second year of the project, the results generated will be presented at national and international congresses that are benchmarks in diabetes research, such as the congresses of the Spanish Diabetes Society and the European Diabetes Society, or other societies (from the second year).

In addition, it is expected that a scientific article will be published for each objective of the study in high-impact journals in the field.

11. RESOURCES FOR CONDUCTING THE STUDY

The research group has members with extensive experience in clinical research. They have led several projects, similar to the current one in terms of organizational and logistical capabilities, which have given rise to multiple publications. We will only mention a few of the studies that require similar knowledge, experience, and skills to that required for this proposal: D. Mauricio led a prediabetes cohort study (the Mollerussa cohort recruited more than 600 individuals and has already produced multiple publications; see publications of D. Mauricio); J. Franch led the PREDAPS study (Medicine; 2015;94:44); C. Farràs , leads the ILERVAS cohort study, in which J. Franch and D. Mauricio are also researchers (more than 8,000 people recruited PCC in the province of Lleida) (Nefrologia . 2016;36:389-396); D. Mauricio led a pragmatic trial (INTEGRA) in different CAPs in the territories of Barcelona, Lleida and Girona (more than 600 subjects included; Molló et al. BMC Family Practice 2019 ;20:25). M. Hernández and D. Mauricio recruited the Action LADA international collaborative Spanish cohort (this has

resulted in previous publications; see background section). Therefore, the previous experience and knowledge of the research team in the implementation of this type of study ensure that the project will be conducted successfully.

Ultrasound: USR Lleida researchers (C. Farràs), as well as the PI, have solid experience in the characterization and interpretation of subclinical atherosclerosis by vascular ultrasound. The Principal Investigator has contributed to several articles on clinical research on atherosclerosis in diabetes.

USR (Primary Care Research Support Unit) Barcelona: J. Franch-Nadal (JF) is PI of the leading Epidemiological Research Group in Diabetes from Primary Care (DAPCAT). He has a long experience in conducting clinical research in the study of diabetes in primary health care through epidemiological studies based on the exhaustive analysis of routinely collected data. B. Vlacho (BV) is a project manager expert in managing clinical research projects. Other members of the team are those of each CAP: B. Ribas (BR), a family doctor with experience in clinical research, has been participating as an investigator in different trials and a project financed by the ISCIII (ITADI study). Ana M. Pedro Pijoan (AM.PP) family doctor and clinical researcher with experience Dos de Maig /ABS Gaudí. N. Riera (NR), experienced family physician and clinical research investigator, involved in intervention programs to improve medical care. C.González -Blanco (CG-B) medical endocrinologist Hospital Dos de Maig . H. Hernández Boluda (HH), experienced family physician and site investigator. He participated as a site investigator in the INTEGRA study.

USR Lleida: M. Ortega (MO), family doctor specializing in primary care therapeutics, and coordinator of the Primary Care Therapies Research Group (GRETAP). His team includes E. Artigues (EA), J. Sol (JS), L. Azlor (LA), and JI Garcia (JIG). E. Artigues is a nurse doctor specializing in health education, and an expert in community care and health promotion. J.Sol is a predoctoral researcher in biomarkers metabolomics and lipidomics . L. Azlor has a master's degree and is an expert in Biostatistics. JI. García is a pharmacist with a specialization in Hospital Pharmacy and a Master's degree in Clinical Research, Psychopharmacology and Aging. Other members of the team are those of each PCC: J. Pujol (JP), family doctor with experience in clinical trials (as PI in 3 trials) with a special interest in diabetes research. M. Solanes (MS), family doctor and field researcher in the two intervention trials (INTEGRA and E-COVID), and as

principal investigator in two projects (ECOGRAFIES Lleida and Ultrasound in primary care) A. Lafarga (AL) : family doctor and researcher trained at the clinical center, participated in the INTEGRA study. S. Guerrero (SG), family doctor and highly motivated researcher in diabetes research (different local studies). M. Marin (MM), family doctor; Over the years, he participated in a local intervention study to improve diabetes care.

Additionally, the sub -study on preclinical atherosclerosis that will take place in the health area of Lleida (6 CAP; half of the study participants) will be carried out without problems since the USR Lleida already has the stage to carry out the study without additional resources, since which are already running the second wave of the ILERVAS study.

The study is funded by the Carlos III Health Institute (FIS 2020 PI21/01163).

12. CHANGES TO THE PROTOCOL:

Any modification of the study protocol will always take the form of a written amendment or addendum. For its formalization, the approval of all the people responsible for the study will be required. In the case of relevant modifications, the express approval of the Clinical Research Ethics Committee will be requested.

1 3. PRACTICAL CONSIDERATIONS:

Start, follow-up and end reports

The start of the study will be notified to the ethics committee. Subsequently, annual monitoring reports will be sent. After obtaining the conclusions of the study, a final report will be prepared and presented to the ethics committee.

14. BIBLIOGRAPHY:

[1] Busse et al.Tackling chronic disease in Europe: strategies, interventions and challenges, WHO Regional Office Europe, 2010.

[2] F. Soriguer, A. et al.Revalence of diabetes mellitus and impaired glucose regulation in Spain: the Di@ bet. es Study, Diabetologia. 55 (2012) 88 -93.

[3] Mauricio et al. Chronic Diabetes Complications: The Need to Move beyond Classical Concepts, Trends Endocrinol. Metab. 31 (2020) 287 -295. https://doi.org/10.1016/j.tem.2020.01.007.

[4] Hurtado et al. What is type 2 diabetes?, Medicine (Baltimore). 47 (2019) 10 -15.

[5] Pearson et al. Lecture 2013 Stratified approaches to the management of diabetes, Diabet. Med. 31 (2014) 393 -398.

[6] Mata-Cases et al. Clinical inertia in the treatment of hyperglycemia in type 2 diabetes patients in primary care, Curr. Med. Res. Opin. 29 (2013) 1495 -1502. https://doi.org/10.1185/03007995.2013.833089.

[7] Mata-Cases et al. Glycaemic control and antidiabetic treatment trends in primary care centres in patients with type 2 diabetes mellitus during 2007 -2013 in Catalonia: a population-based study, BMJ Open. 6 (2016) e012463. https://doi.org/10.1136/bmjopen-2016-012463.

[8] Blind et al. The European Medicines Agency's approval of new medicines for type 2 diabetes, Diabetes, Obes. Metab. 20 (2018) 2059 -2063. https://doi.org/10.1111/dom.13349.

[9] Hernández et al. Latent Autoimmune Diabetes in Adults: A Review of Clinically Relevant Issues, in: Diabetes from Res. to Clin. Pract. Vol. 4, Springer, 2020: pp. 29 -41. https://doi.org/10.1007/5584_2020_533.

[10] Castelblanco et al. Low-grade Inflammatory Marker Profile May Help to Differentiate Patients With LADA, Classic Adult-Onset Type 1 Diabetes, and Type 2 Diabetes, Diabetes Care. 41 (2018) 862 -868. https://doi.org/10.2337/dc17-1662.

[11] Hernández et al. Preclinical carotid atherosclerosis in patients with latent autoimmune diabetes in adults (LADA), type 2 diabetes and classical type 1 diabetes, Cardiovasc. Diabetol. 16 (2017) 94. https://doi.org/10.1186/s12933-017-0576-9.

[12] Mollo et al. Latent autoimmune diabetes in adults is perched between type 1 and type 2: evidence from adults in one region of Spain, Diabetes. Metab. Res. Rev. 29 (2013) 446 -451. https://doi.org/10.1002/dmrr.2411.

[13] Hawa et al. Metabolic Syndrome and Autoimmune Diabetes: Action LADA 3, Diabetes Care. 32 (2009) 160 -164. https://doi.org/10.2337/dc08-1419.

[14] Hawa et al. Adult-Onset Autoimmune Diabetes in Europe Is Prevalent With a Broad Clinical Phenotype: Action LADA 7, Diabetes Care. 36 (2013) 908 -913. https://doi.org/10.2337/dc12-0931.

[15] Mishra et al. Relative contribution of type 1 and type 2 diabetes loci to the genetic etiology of adult-onset, non-insulin-requiring autoimmune diabetes, BMC Med. 15 (2017) 88. https://doi.org/10.1186/s12916-017-0846-0.

[16] Mishra et al. Genetic Discrimination Between LADA and Childhood-Onset Type 1 Diabetes Within the MHC, Diabetes Care. 43 (2020) 418 -425. https://doi.org/10.2337/dc19-0986.

[17] ADA, Introduction: Standards of Medical Care in Diabetes --2021, Diabetes Care. 44 (2021) S1 -S2.

[18] WHO, Classification of diabetes mellitus, WHO, 2019. https://www.who.int/health-topics/diabetes#tab=tab_1.

[19] Buzzetti et al. Management of Latent Autoimmune Diabetes in Adults: A Consensus Statement From an International Expert Panel, Diabetes. 69 (2020) 2037 -2047. https://doi.org/10.2337/dbi20-0017.

[20] Fitipaldi et al. Global Overview of Precision Medicine in Type 2 Diabetes, Diabetes. 67 (2018) 1911 -1922. https://doi.org/10.2337/dbi17-0045.

[21] Chung et al. Precision Medicine in Diabetes: A Consensus Report From the American Diabetes Association (ADA) and the European Association for the Study of

Diabetes (EASD), Diabetes Care. 43 (2020) 1617 -1635. https://doi.org/10.2337/dci20-0022.

[22] Prasad et al. Precision medicine in type 2 diabetes, J. Intern. Med. 285 (2019) 40 -48. https://doi.org/10.1111/joim.12859.

[23] McCarthy, Painting a new picture of personalised medicine for diabetes, Diabetologia. 60 (2017) 793 -799. https://doi.org/10.1007/s00125-017-4210-x.

[24]Udler et al. Genetic Risk Scores for Diabetes Diagnosis and Precision Medicine, Endocr. Rev. 40 (2019) 1500 -1520. https://doi.org/10.1210/er.2019-00088.

[25] Ahlqvist et al. Novel subgroups of adult-onset diabetes and their association with outcomes: a data-driven cluster analysis of six variables, Lancet Diabetes Endocrinol. 6 (2018) 361 -369. https://doi.org/10.1016/S2213-8587(18)30051-2.

[26] Zaharia et al. Risk of diabetes-associated diseases in subgroups of patients with recent-onset diabetes: a 5-year follow-up study, Lancet Diabetes Endocrinol. 7 (2019) 684 -694. https://doi.org/10.1016/S2213-8587(19)30187-1.

[27] Wagner et al. Pathophysiology-based subphenotyping of individuals at elevated risk for type 2 diabetes, Nat. Med. 27 (2021) 49 -57. https://doi.org/10.1038/s41591-020-1116-9.

[28] Dennis et al. Disease progression and treatment response in data-driven subgroups of type 2 diabetes compared with models based on simple clinical features: an analysis using clinical trial data, Lancet Diabetes Endocrinol. 7 (2019) 442 -451. https://doi.org/10.1016/S2213-8587(19)30087-7.

[29] Bello-Chavolla et al. Clinical characterization of data-driven diabetes subgroups in Mexicans using a reproducible machine learning approach, BMJ Open Diabetes Res.

[30] Udler et al. Type 2 diabetes genetic loci informed by multi-trait associations point to disease mechanisms and subtypes. A soft clustering analysis, PLOS Med. 15 (2018) e1002654. https://doi.org/10.1371/journal.pmed.1002654.

[31]. Román Viñas B, Ribas Barba L, Ngo J, Serra Majem L. Validation in the Catalan population of the international physical activity questionnaire [Validity of the international fitness activity questionnaire in the Catalan population (Spain)]. gac Sanit . 2013 May-June ;27 (3):254-7. Spanish. doi : 10.1016/j.gaceta.2012.05.013. Epub 2012 Oct 24. PMID: 23103093.

[32] Schröder H, Fitó M, Estruch R, Martínez-González MA, Corella D, Salas-Salvadó J, Lamuela -Raventós R, Ros E, Salaverría I, Fiol M, Lapetra J, Vinyoles E, Gómez-Gracia E, Lahoz C, Serra- Majem L, Pintó X, Ruiz- Gutierrez V, Covas MI. A short screener is valid for assessing Mediterranean diet adherence among older Spanish men and women. J Nutr . 2011 Jun ; 141 (6): 1140-5. doi : 10.3945 / jn.110.135566. Epub 2011 Apr 20. PMID: 21508208.

[33] Martínez-González MA, Fernández-Jarne E, Serrano-Martínez M, Wright M, Gomez-Gracia E. Development of a short dietary intake questionnaire for the quantitative estimation of adherence to a cardioprotective Mediterranean diet. Eur J Clin Nutr. 2004;58:1550–2.

[34] Magliano DJ, Chen L, Islam RM, Carstensen B, Gregg EW, Pavkov ME, Andes LJ, Balicer R, Baviera M, Boersma-van Dam E, et al. Trends in the incidence of diagnosed diabetes: a multicountry analysis of aggregate data from 22 million diagnoses in high-income and middle-income settings. Lancet Diabetes Endocrinol (2021) 9:203–211. doi:10.1016/S2213-8587(20)30402-2

15. EXPERIENCE OF THE RESEARCH TEAM

The research team is made up of professionals from different disciplines, a good part of whom have been collaborating for a number of years in this area of research.

D. Mauricio leads the Diabetes Research Group of this institution (HSCSP-IR), which belongs to the CIBER of Diabetes and Associated Metabolic Diseases (CIBERDEM), Carlos III Health Institute (ISCIII). He has published more than 240 peer-reviewed articles (H-index: 42), mainly in the field of diabetes. He has extensive experience in clinical studies in diabetes, having participated in multiple national and international research initiatives such as PI. He has been funded by the Carlos III Health Institute, and international research organizations, leading 8 projects as PI in the last 10 years. The main research activities focus on the study and treatment of diabetes and its associated complications, and also on the evidence in real practice (see below).

J. Franch is a family doctor and principal investigator of the DAP-CAT group, heading the Epidemiological Research Group on Diabetes from Primary Care (IDIAP Jordi Gol Research Institute). He is a national primary care expert and leader in the field of type 2 diabetes. He has long experience conducting clinical research on diabetes in primary health care through epidemiological studies. In 2009, D. Mauricio and J. Franch pioneered clinical evidence research in real practice using large database epidemiology studies in the field of diabetes. Together, they have also conducted pragmatic primary care trials. J. Franch will be instrumental in leading the project at the primary care level.

The research team also includes experience in the field of endocrinology in both participating territories. M. Hernández (Lleida) and C. González (Barcelona) will contribute their experience in this field. M. Hernández will contribute with his experience in the field of LADA, and C. González will be essential for his experience in technology in diabetes and communication with patients.

Fortunately, we have been able to gather the necessary experience and commitment from the 2 IDIAP Research Support units involved in Lleida and Barcelona, where the relevant leadership is provided by local leaders M. Ortega and J. Franch, respectively. M. Ortega coordinates research initiatives in primary care in the health area of Lleida, with extensive experience in interventional studies in CPCs in this region. Relevant

experience and key members of the local teams are outlined in the section on Available Resources.

The proven national and international leadership of different members in clinical diabetes has resulted in participation in guidelines from different scientific organizations: Primay Care Diabetes Europe (J. Franch ; Prim Care Diabetes. 2021 ;15: 31-51); redGDPS in Spain (J. Franch; Diabetes Practice. 2020;11:117-162 and Diabetes Practice. 2020;11:47-54); International group of experts in LADA (D. Mauricio; Diabetes. 2020;69:2037-2047);

The team includes 1 senior statistician and collaborators from the statistical team of the USR units. J. Real is a senior data scientist, with more than 13 years of experience in data management and statistical analysis of large population databases. In addition, the bioinformatics group of Dr. Perera (UPC, Univ. Politècnica de Catalunya) will advise and support these analyses.

Two members of the team (M. Granado, RD and PhD, and K. Rodríguez, RD, MD) will lead the lifestyle studies. Together with the main researcher, M. Granado, M. Mata and J. Franch have carried out several studies in this field.

Finally, we want to highlight the fact that the key researchers of this project are members of the CIBERDEM research group led by D. Mauricio: J. Franch, J. Real and M. Granado. This group has achieved excellence among the 31 CIBERDEM groups, having reached the first and second position in the annual evaluations of 2018 and 2019, respectively. Our research group has stable and ongoing collaborations with other international researchers: Kamlesh Khunti (University of Leicester, UK) and Dianna Magliano (Monash University , Melbourne, AUT) in the field of real world data and diabetes epidemiology; Angelo Avogaro (University of Padova , IT), Peter Rossing (Steno Diabetes Center, DK) and Per- Henrik Group (University of Helsinki, FI) in the field of diabetes complications.

16. APPLICABILITY OF THE PROJECT

This project entails the application of research in different priority areas defined in the current AES call.

Capacity of the project to address the objectives and priorities of the Social Challenge of Health, Demographic Change and Well-being of the Spanish Strategy for Science, Technology and Innovation.

This project addresses important questions about the current objectives and priorities of the Spanish investigative authority. First of all, the study focuses on a relevant health problem, that is, diabetes mellitus. Actually, type 2 diabetes mellitus is one of the main causes of morbidity and mortality in our country and the prevalence of the disease is growing more and more to an epidemic proportion as in other developed countries. Second, the project is an effort on precision medicine in diabetes. This is also considered a priority both locally and internationally. The project focuses on precision diagnosis, as a first step towards the implementation of other aspects of precision medicine (prevention, prognosis, therapeutics, follow-up). Individualization and shared decision making are now at the center of diabetes care; however, the optimal distribution of these important aspects of clinical care must be based on the integrity of the relevant information (in this case, deep phenotyping). Therefore, precision medicine in diabetes is a requirement for optimal individualization and decision-making.

translational research proposal

The findings of the current project should result in a simple application to develop new strategies for an improvement in the diagnostic approach of T2DM (accuracy diagnosis). In addition, the proposed cohort is the basis for future projects in the field of precision medicine in diabetes. In fact, the follow-up of the prospective cohort will provide very relevant information on the prognosis of the disease (precision prognosis), the prevention of complications (precision prevention) and, more importantly, information on the response to antidiabetic drugs and other therapies (precision therapies). Future application of the findings of this approach should translate into benefits for the person with diabetes and for healthcare professional decision-making.

The implementation of precision medicine in diabetes will also result in an improvement for our health system, since the proper use of resources must be based on relevant information (cost-effectiveness).

We firmly believe that the new information resulting from studies such as ours, which combine innovative methodologies, will ultimately help physicians, health authorities and patients themselves, and help develop strategies to implement better management of the disease, using a more individualized targeted approach.

17. FUNDING SOURCE

Carlos III Health Institute Health Research Projects PI21/01163

18. CONFLICT OF INTEREST

The researchers have no conflict of interest

LIST OF INDEPENDENT DOCUMENTS

ANNEX 1. NOTEBOOK FOR COLLECTING DATA
ANNEX 2. COORDINATING RESEARCHER COMMITMENT
ANNEX 3. RESEARCHER COMMITMENT CENTER
ANNEX 4. CEI CONFORMITY
ANNEX 5. PATIENT INFORMATION SHEET
ANNEX 6. INFORMED CONSENT FORM
ANNEX 7. FINANCIAL REPORT
ANNEX 8. CHARACTERISTICS OF CENTERS AND EXPECTED RESULTS
ANNEX 9. PROMOTIONAL STUDY MATERIAL TO RECRUIT PATIENTS
ANNEX 10. LIST OF CENTERS, PRINCIPAL INVESTIGATORS CENTERS

ANNEX 1. DATA COLLECTION NOTEBOOK

Attached in a separate document.

ANNEX 2. COMMITMENT OF THE PRINCIPAL INVESTIGATOR AND COORDINATOR

The data of the researcher that appear in this document will be entered and processed in a file owned by the Research Institute Foundation of the Hospital de la Santa Creu i Sant Pau – IIB Sant Pau c/ Sant Antoni Maria Claret, 167, 08025 Barcelona, Tel: 93 553 78 69, declared as a Data Management Center, and will be treated in accordance with the provisions of Organic Law 3/2018 Protection of Personal Data and digital rights guarantee. The researcher may exercise their rights of information, opposition, access, rectification or cancellation of the data recognized in said Law.

Dr. Dídac Mauricio with DNI 46525502K from the Hospital de la Santa Creu i Sant Pau in Barcelona.

It states:

That has evaluated the study "Characterization of subgroups of type 2 diabetes at diagnosis: a necessary step towards precision medicine in diabetes, study: COPERNICAN" and whose promoter is the IIB Sant Pau Foundation.

And you agree to:

- Sign a commitment in which they recognize themselves as investigators of the study and affirm that they know the protocol and any modifications to it, and agree with it in all its terms.

- Inform the research subjects and obtain their consent.

- Collect, record and notify the data correctly, responding to its updating and quality before the appropriate audits.

- Notify the sponsor of adverse events as established in the protocol.

- Respect the confidentiality of the data of the participating subjects

- Facilitate the monitoring visits of the monitor, the audits of the promoter and the inspections of the health authorities.

- Know how to answer about the objectives, basic methodology and meaning of the results of the study before the scientific and professional community.

- Sign the protocol and any modifications to it together with the promoter.

- Contribute to disseminating the results of the study, in collaboration with the promoter.

Investigator Signature:

Date:

ANNEX 3. RESEARCHER COMMITMENT

The data of the researcher that appears in this document will be entered and processed in a file owned by the foundationFoundation Institut de Recerca de l'Hospital de la Santa Creu i Sant Pau – IIB Sant Pau c/ Sant Antoni Maria Claret, 167, 08025 Barcelona, Tel: 93 553 78 69, declared as Data Management Center, and will be treated in accordance with what is provided by Organic Law 3/2018 Protection of Personal Data and guarantee digital rights. The researcher may exercise their rights of information, opposition, access, rectification or cancellation of the data recognized in said Law.

Dr. with DNI of the service of the center It states:
That it has evaluated the study "Characterization of subgroups of type 2 diabetes at diagnosis: a necessary step towards precision medicine in diabetes, study: COPERNICAN" and whose promoter is the IIB Sant Pau Foundation.

That he agrees to participate as principal investigator, and undertakes to:

-Sign a commitment in which they recognize themselves as investigators of the study and affirm that they know the protocol and any modifications to it, and agree with it in all its terms.
-Inform the research subjects and obtain their consent.
-Collect, record and notify the data correctly, responding to its updating and quality before the appropriate audits.
-Notify the promoter of adverse events as established in the protocol.
-Respect the confidentiality of the data of the participating subjects
-Facilitate the monitoring visits of the monitor, the audits of the promoter and the inspections of the health authorities.
-Know how to answer about the objectives, basic methodology and meaning of the results of the study before the scientific and professional community.
-I freely agree to participate in the study and give my consent for the access and use of my data under the conditions detailed in the study protocol.

Firma del Investigador:		Fecha:	
Investigadores colaboradores:			
Nombre:	DNI:	Firma del Investigador:	Fecha:
Nombre:	DNI:	Firma del Investigador:	Fecha:

It will be included once the aforementioned agreement is obtained.

ANNEX 5. PATIENT INFORMATION SHEET

STUDY TITLE: Characterization of type 2 diabetes subgroups at diagnosis: a necessary step towards precision medicine in diabetes, study: COPERNICAN

PROMOTER: IIB Sant Pau Foundation

INTRODUCTION

We are writing to inform you about a research study in which we invite you to participate. The study has received the approval of the Ethics Committee of the IIB Sant Pau Foundation and IDIAP Jordi Gol i Gurina , in accordance with current legislation. This study is funded by the IIB Sant Pau Foundation through the award of a FIS 2021 PI21/01163 project from the Carlos III Health Institute, which allocates its resources to the study of diabetes mellitus type 2.
Our center participates in this study carried out in Catalonia.
Our intention is that you receive the correct and sufficient information so that you can evaluate and judge whether or not you want to participate in it. To do this, read this information sheet carefully and we will clarify any doubts that may arise after reading it. In addition, you can consult with the people you consider appropriate.

VOLUNTARY PARTICIPATION

Your participation in this study is voluntary, and you may decide not to participate or change your decision and withdraw your consent at any time, without changing your relationship with your doctor or affecting your health care.

THE PURPOSE OF THE STUDY

To evaluate and characterize the relevant subgroups of type 2 diabetes at the time of diagnosis in our country. The results of this study could help us better understand the variety of type 2 diabetes, and identify different important subgroups that allow improving the management and treatment of the disease.

STUDY OVERVIEW

The study will be carried out in the primary care centers of the various health areas of Catalonia. This characterization consists of collecting a very complete series of clinical characteristics in all people, including those specifically characteristic of diabetes and also those of the person himself. This should allow us to contribute new knowledge in what is known as precision medicine, in this case applied to the precision diagnosis of diabetes. One of the problems we have in the management of type 2 diabetes is the variety of possible subgroups of patients that comprise it. We know, from previous research, that people with this type of diabetes do not have all the common characteristics at the beginning of the disease, nor do they evolve in the same way, nor do they respond in the same way to treatments, and they do not have the same risk of developing diabetes. complications. In addition, no treatment measures

are currently recommended for each person individually, and sharing decisions between the patient and health professionals. This makes it increasingly necessary that we obtain much more precise information on the different subgroups of people diagnosed with type 2 diabetes and individually characterize those that correspond to each specific patient profile. This information, which is what this project intends to generate, when diagnosing diabetes, should lead us to have much more precise information for decision-making by professionals, and shared with the affected person. The application of this strategy with more precise information in each case (precision medicine) should allow in the future that each person with type 2 diabetes receive individualized management of their condition, considering the specific factors of their case (precision medicine). personalized).

Diagnosis of type 2 diabetes

Type 2 diabetes used to appear in older ages, but the growing epidemic of overweight and obesity in the population of all ages is resulting in an increase in the number of people with type 2 diabetes also in the younger population. Obesity, which is in many cases linked to a sedentary lifestyle and an unhealthy diet, is one of the main causes that ultimately leads to the development of type 2 diabetes.

The diagnosis of type 2 diabetes is usually late since the symptoms may go unnoticed or not be identified by the person as typical of diabetes, so it may even take several years without the person going to health centers to diagnose this condition.

Visits and follow-up program

Before participating in the study, you will be contacted by your family doctor or researcher at the center to explain the possibility of participating and explain all the study procedures.

If your decision is to participate in the study, before scheduling your first visit, called an inclusion visit, you will be given a telephone questionnaire in order to detect possible symptoms or contacts of COVID-19 infection. If there is any suspicion, you will proceed according to local clinical practice recommendations, and your participation in the study will be postponed.

If you agree to participate in the study, the local investigator at the center will review your medical records. Only one visit will be made for the study in which data related to alcohol or tobacco consumption, weight and height, waist circumference, blood pressure, information on other diseases suffered, medications taken, request for laboratory analysis will be collected. related to diabetes, possible complications you may have related to diabetes, and questionnaires about your diet and physical activity.

Some primary care centers participating in this study have equipment to assess atherosclerosis. Atherosclerosis is a condition in which plaque made of fat, cholesterol, calcium, and other substances found in the blood builds up inside the arteries (blood vessel). In these centers you will be offered the possibility of performing a painless, safe and fast non-invasive diagnostic test (ultrasound) with a

short duration to evaluate the involvement of atherosclerosis in the blood vessels of the neck and legs.

You will be scheduled for routine blood and urine tests, as recommended by current clinical practice guidelines. Laboratory procedures will be performed at local laboratories. Specific additional laboratory procedures for the determination of C-peptide and antibodies that are considered non-routine parameters will also be performed at local laboratories from the same blood draw. These tests will help us better characterize your diabetes.

Since current clinical practice includes assessment of chronic complications of diabetes (foot exam, eye disease screening program, problems with kidney function), other diseases and related clinical and laboratory variables of interest with diabetes, mortality, these data will be extracted by linking them with the database of electronic records of medical records (databases of Catalonia-SIDIAP/PADRIS, for specific causes of mortality through the INE) or through the electronic clinical record in the event that the link with the SIDIAP/PADRIS databases is not possible. Only if you have accepted monitoring of your disease with the Catalonia-SIDIAP/PADRIS databases, your encrypted CIP (Personal Identification Code) (transcribing your personal identification code with a key) will be sent to these databases. data to do data binding.

Portability of data from other studies

If you are participating or have participated in any other clinical research projects, you may transfer your data from these studies to this study. The data portability procedure from another study will be carried out by the research team with the permission granted by you in this informed consent.

Withdrawal from the study

Your participation in this study is completely voluntary, you can withdraw from the study at any time if you wish without having to explain to your doctor and without compromising in any way the treatment you have to receive.

BENEFITS AND RISKS OF YOUR PARTICIPATION IN THE STUDY

The study has no possibility of generating risk because it is an observational study without the specific use of medication, which is limited to an anonymous data record in a database that does not allow access to the patient's personal data.

CONFIDENTIALITY

The treatment, communication and transfer of personal data of all participating subjects will comply with the new legislation in the European Union (EU) on personal data, specifically Regulation (EU) 2016/679 of the European Parliament and of the

Council of April 27, 2016 on Data Protection (RGPD). Organic Law 3/2018, of December 5, on the Protection of Personal Data and guarantee of digital rights is also in force. In accordance with the provisions of the aforementioned legislation, you can exercise your rights of access, modification, opposition and cancellation of data, for which you must contact the study doctor.
The data collected for the study will be identified by a code and only your study doctor and collaborators will be able to associate said data with you and with your medical history. Therefore, your identity will not be revealed to any person except for exceptions, such as in case of medical emergency or legal requirement.
If the study is published or the results are disseminated in scientific publications or congresses, no data will be included that could provide information that helps to identify patients.
Access to your personal information will be restricted to the study doctor/collaborators, health authorities (Spanish Agency for Medicines and Health Products), the Ethics Committee for Research with Medicines (CEI) and personnel authorized by the sponsor, when required to do so. verify the data and procedures of the study, but always maintaining their confidentiality in accordance with current legislation.

It is also important that you know the following information related to your rights:

In addition to the rights that you already know (access, modification, opposition and cancellation of data), you can currently also limit the processing of data that is incorrect, request a copy or transfer to a third party (portability) the data that you have provided for the study. To exercise your rights, you can contact your treating doctor and you must provide a document that proves your identity as the owner of the data (DNI), and indicate the right you wish to exercise. The promoter's data protection delegate will manage the request and will issue a response according to the case. The following steps will be carried out: your doctor will contact those responsible for managing the study using the data indicated in the study protocol, without disclosing in any case the personal data of the patient, but only indicating the code to which the holder corresponds, code of the study and code of the center, as well as the right that you have been asked to exercise. The doctor must not send the promoter the document proving the identity, since it was verified by him at the time of the request.
We remind you that the data cannot be deleted, even if you stop participating in the study, to guarantee the validity of the research and to comply with legal duties and drug authorization requirements. You also have the right to contact the Spanish Data Protection Agency. C/ Jorge Juan, 6. 28001 Madrid (https://www.aepd.es/es) or the Catalan Dades Protection Authority (https://apdcat.gencat.cat), if you are not satisfied.

The center is responsible for the processing of your data and undertakes to comply with current data protection regulations. The data collected for the study will be identified by a code, so that information that can identify you is not included, and only your study doctor / collaborators will be able to relate said data to you and your medical history. Therefore, your identity will not be disclosed to any other person except the health authorities, when required or in cases of medical emergency. The Research Ethics Committees, the representatives of the Health Authority in matters of inspection and the personnel authorized by the Promoter, may only access to check the

personal data, the clinical study procedures and compliance with the standards of good clinical practice. (always maintaining the confidentiality of the information).

Responsible for data processing: Institute Catalan of Health

Responsible for data processing: SIDIAP/PADRIS

Legal basis of the treatment: RGPD: 6.1 a) the interested party gave their consent for the treatment of their personal data for one or several specific purposes.

The Researcher and the Promoter are obliged to keep the data collected for the study for at least 5 years after its completion. Subsequently, your personal information will only be kept by the health care center and by the sponsor for other scientific research purposes if you have given your consent to do so, and if permitted by applicable law and ethical requirements.

If we transfer your encrypted data within the EU to our group entities, service providers or scientific researchers who collaborate with us, the participant's data will be protected with safeguards such as contracts or other mechanisms by data protection authorities. data. If you want to know more about it, you can contact the Data Protection Officer of the institution or the promoter Institut de Recerca de l'Hospital de la Santa Creu i Sant Pau – IIB Sant Pau c/ Sant Antoni Maria Claret , 167, 08025 Barcelona, Tel: 93 553 78 69, email dpo_ir@santpau.cat, or with the Spanish Data Protection Agency at Jorge Juan, 6, 28001 Madrid (https: // www .aepd.es / es) or the Catalan Dades Protection Authority (https://apdcat.gencat.cat)

Biological samples

Collection and use

Your participation in this study entails obtaining and using biological samples for research purposes, for which Law 14/2007 on biomedical research and Royal Decree 1716/2011 will be observed, regulations that guarantee respect for your rights . . By signing this document, reviewed and favorably evaluated by the Research Ethics Committee that has approved this clinical study, you agree to the use of your samples for the purposes of this study.

What is a biobank?: Institution to promote research and health

Biobanks are authorized establishments that house an organized set of biological samples and associated information under the conditions and guarantees of quality and safety required by current legislation. Said samples and their associated information are available to those national or international research centers or institutions that officially request them from the Biobank. The samples and data may therefore be transferred to European countries or from other parts of the world, always for projects that have been approved by the External Committees and by the Scientific Director of the Biobank.

From the donated samples, in the cases in which the research requires it, genetic studies will be carried out that on occasions could refer to the complete sequencing of

your genome, and from them information about your health and that of your loved ones can be obtained. relatives. Action will always be taken to ensure the protection of this information (see section on data protection and confidentiality).

Biological samples will be kept indefinitely in the Biobank or until they are consumed in research projects. In the event of a possible closure of the Biobank, the samples and associated data would be transferred to another authorized Biobank. The information on the destination of the samples will be available to you in the National Registry of Biobanks for Biomedical Research of the Carlos III Health Institute. If an additional sample is necessary, we request your permission for the health institution to contact you to request your collaboration again.

The samples will always be transferred anonymously, without associated personal information.

Discomfort and possible risks

No major risk to participants
The only discomfort may derive from the blood draw. Sampling can cause a burning sensation at the point where the needle is inserted into the skin and cause a small bruise or a mild infection that disappears in a few days. More rarely, dizziness may appear at the time of blood collection.

During the routine blood and urine clinical visit, part of the blood/urine samples will be stored for later analysis. These samples will be used to determine the advanced lipoprotein profile (analysis of fats in the blood, such as cholesterol and triglycerides), which will allow the detection of some specific parameters as part of the objectives of the study. Your samples will also be used to determine metabolites (substance produced during digestion or other body chemical processes) and genes (information that determines our traits) related to diabetes.

Apart from this purpose of biological samples, it is also possible that your samples will be used for other studies related to diabetes, non-diabetes related, or genetic studies only and when you accept one of these options in the informed consent form. for donation and preservation of biological samples. In the event that you accept these purposes of use of your samples, you will be informed by your doctor regarding the results obtained if you wish.
The samples obtained will be labeled with a code to maintain confidentiality, will be processed in local laboratories, and later stored by the Biobank of the Research Institute of the Hospital de la Santa Creu i Sant Pau.

Responsible for data processing
The personal data collected will be obtained, processed and stored in a file under the responsibility of the Hospital de la Santa Cruz y San Pablo Research Institute Foundation for research purposes. The file wants to guarantee the quality, security and traceability of the data and biological samples stored as well as of the associated procedures, in the terms established in Law 14/2007, of July 3, on Biomedical

Research; complying at all times with the duty of secrecy, in accordance with current legislation on the protection of personal data.

The identification of the biological samples will be subjected to a coding process. Each sample is assigned an identification code that is detached from personal data. Only authorized personnel may associate their identity with the aforementioned codes. Through this process, the researchers who request samples will not be able to know any information that reveals their identity. Likewise, although the results obtained from the research carried out with your samples are published in scientific journals, your identity, or that of your family in the event that they participate jointly, will not be disclosed.
In those studies in which no potentially useful results for your health are expected, and in accordance with the corresponding Ethics Committee, the samples and data may be anonymized, that is, there will be no possibility of re-associating the sample with your identity. .
You may exercise your rights of access, rectification, cancellation and objection, as well as obtain information on the use of your samples and associated data, by contacting the legal person responsible for the study

Biobank. Sant Pau Biomedical Research Institute
Santa Cruz and San Pablo Hospital Research Institute Foundation (hereinafter, FIRHSCSP or Foundation). Registered office: Calle San Antonio María Claret , 167, 08025 Barcelona. Registered in the Registry of Foundations of the Government of Catalonia, with number 708. CIF: G-60136934. Telephone: 93.553.76.13. Contact email: biobanc@santpau.cat .
Promoter: The principal investigator (Dr. Dídac Mauricio DNI: 46525502K) of the study and the Promoter- Research Institute of the Hospital de la Santa Creu i Sant Pau – IIB Sant Pau, c/ Sant Antoni Maria Claret , 167 08025 Barcelona, Tel: 93 553 78 69. Contact email: dmauricio@santpau.cat

Purposes
Your data will be processed solely to carry out research to advance knowledge of diabetes.

Legal basis of the treatment
RGPD: 6.1 a) the interested party gave their consent to the processing of their personal data for one or more specific purposes.

Rights of interested persons
Interested parties may exercise, if they wish, the rights of access, rectification and deletion of data, as well as request that the processing of their personal data be limited, oppose the processing, and request the portability of their data, through the following media:
- By sending a written request to the indicated postal address.
- By sending a request to the email: dpo_ir@santpau.cat

The interested parties may, in relation to those treatments that are based on obtaining their consent, withdraw their consent through the procedure detailed in the previous paragraph.

Also, if you believe that the protection of any of your rights has been violated, you can file a claim with the Spanish Agency for Data Protection (www.aepd.es).

GRATITUDE

Thank you very much for reading this study information document. If you decide to be part of the study, you will need to sign the informed consent form. A copy of this information document and the consent document will be given to you for your retention.

We think that it is a necessary study, which will help us to better understand this pathology and thus be able to improve patient care.

Your participation is important and we greatly appreciate your collaboration. We remain at your disposal for any clarification or question you may have.

CONTACT IN CASE OF DOUBTS

If you need any information or for any other reason, do not hesitate to contact the study doctor, Dr._____, at the primary care center_____, Address :_ _____, contact telephone number: _____

ANNEX 6. INFORMED CONSENT FORM

STUDY TITLE: Characterization of type 2 diabetes subgroups at diagnosis: a necessary step towards precision medicine in diabetes, studio: COPERNICAN
PROMOTER: IIB Sant Pau Foundation

I (name and surnames -written in handwriting by the participating subject-)

....
.....
- I have read the information sheet that has been given to me
- I was able to ask questions about the study.
- I have received enough information about the study.
- I have spoken with:

....
(Name and surname of the researcher)
I understand that my participation is voluntary.
I understand that I can withdraw from the study:
- Whenever you want
- Without having to explain.
- Without this affecting my medical care.

	YES	NO
I accept all the options mentioned below	☐	☐
I accept that my electronic medical record data is crossed with clinical databases to track my disease	☐	☐
I accept that the follow-up of my disease is done through the electronic medical record	☐	☐
I agree to be contacted for future studies related to diabetes	☐	☐
I accept the transfer of my data under the conditions detailed in the information sheet	☐	☐
I accept portability of my data from other research studies	☐	☐

I freely give my consent to participate in the study, and I give my consent for the access and use of my data under the conditions detailed in the information sheet.
Patient Signature: Investigator Signature:
Name: Name:

Date: ____/____/____ Date: ____/____/____

(Name and date, handwritten by the participating subject)
This document will be signed in duplicate, one copy being kept by the researcher and another by the patient.

INFORMED CONSENT FORM FOR WITNESS

STUDY TITLE: Characterization of type 2 diabetes subgroups at diagnosis: a necessary step towards precision medicine in diabetes, study: COPERNICAN

PROMOTER: IIB Sant Pau Foundation,

I……. (name and surname of the witness) declare under my responsibility that ……..……. (Name and surname of the participant) I have read (or it has been read to him, in the event that the patient cannot read), the information sheet that has been given to him.

- You have been able to ask questions about the study.
- You have received enough information about the study.
- Have you talked to:

....
(Name and surname of the researcher)

- You understand that your participation is voluntary.
- You understand that you can withdraw from the study:
  - Whenever you want
  - Without having to explain.
  - Without this affecting my medical care.

You will receive a signed and dated copy of this informed consent document
You freely agree to participate in the study and give your consent for the access and use of your data under the conditions detailed in the information sheet.

Witness Signature: Investigator Signature:

Name: Name:

Date: ____/____/____ Date: ____/____/____

(Name and date, handwritten by the witness)

This document will be signed in duplicate, keeping one copy for the investigator and another for the witness.

INFORMED CONSENT FORM FOR DONATION AND CONSERVATION OF BIOLOGICAL SAMPLES FOR THE BIOBANK

STUDY TITLE: CHARACTERIZATION OF SUBGROUPS OF TYPE 2 DIABETES AT DIAGNOSIS: A NECESSARY STEP TOWARDS PRECISION MEDICINE IN DIABETES, COPERNICAN STUDY

PROMOTER: IIB Sant Pau Foundation

- I understand why the promoter is preserving the biological samples for the biobank.
- I have had the opportunity to ask questions. All my questions have been answered.
- I understand that I will be given a copy of this consent form after I sign it.
- I understand and authorize the processing of my personal data as indicated in the Patient Information Sheet

1. I agree that my biobanked blood and urine samples **may be** used for:

	YES	NO
I accept all the options mentioned below	☐	☐
Research related to the study	☐	☐
Future research related to diabetes	☐	☐
Future research not related to diabetes	☐	☐
Future genetic studies	☐	☐
I want to know the results of the tests carried out with my biological samples	☐	☐

As for the full study, your participation in the biobanking portion of the study sample donation is voluntary and you are free to withdraw your consent at any time for any reason without influencing your participation in the main study.

…… …….

Patient name Patient signature Date

…… …….

Name of the investigator/ delegate Signature of the investigator/ delegate Date

REVOCATION OF CONSENT form for participation in the study Characterization of subgroups of type 2 diabetes at diagnosis: a necessary step towards precision medicine in diabetes, study: COPERNICAN

I wish to withdraw from the study Characterization of subgroups of type 2 diabetes at diagnosis: a necessary step towards precision medicine in diabetes

However, I allow:

- Follow-up of the study as described in the protocol

☐ If not ☐

……...…….

Name of patient

…….. …….

Place and Date Signature of the patient

……...…….

Name of the investigator/delegate

…….. …….

Place and Date Signature of the researcher/delegate

REVOCATION OF CONSENT Form FOR DONATION AND CONSERVATION OF BIOLOGICAL SAMPLES FOR BIOBANK in the study Characterization of type 2 diabetes subgroups at diagnosis: a necessary step towards precision medicine in diabetes, COPERNICAN study

PER DONOR:

/ Mrs. .... with DNI .... I cancel the consent given on the date ..... of .... of 20 ..... and I do not wish to continue voluntary donation in the study Characterization of subgroups of type 2 diabetes at diagnosis: a necessary step towards precision medicine in diabetes, which I am ending with today's date.

☐ I REQUEST THE DELETION OF THE SAMPLE ONLY

☐ I REQUEST THE DELETION ONLY OF MY PERSONAL DATA

The sample will be irreversibly anonymized and may be used in research projects.

☐ I REQUEST THE TOTAL DELETION OF MY DATA AND SAMPLES

signed :

……...…….

Name of patient

…….. …….

Place and Date Signature of the patient

……...…….

Name of the investigator/delegate

…….. …….

Place and Date Signature of the researcher/delegate

ANNEX 7. FINANCIAL REPORT

Study title: "Characterization of type 2 diabetes subgroups at diagnosis: a necessary step towards precision medicine in diabetes."

Protocol code: COPERNICAN

This study is independent of the Pharmaceutical Industry, which is why the exemption from fees is requested from the Clinical Research Ethics Committees.

The study is funded by a grant from the Health Research Fund (FIS). This aid will go to:

- Researcher meetings
- Preparation of the electronic CRD
- monitoring
- Data extraction
- Data management and analysis
- Biobanks
- GADA65, IA-2 and C-peptide determinations

No payment will be made to the investigators for participating in this trial or to participating patients.

To clarify that in order to carry out this study, tests other than those used in the treatment of this disease in primary care in routine clinical practice are not required. The IIB Sant Pau Foundation assumes the financing of the study in accordance with the guidelines of this protocol, which in any case will be independent of the results obtained.

Center Investigator Signature: Date:

ANNEX 8. Center characteristics and expected results

A. Health region, type 2 diabetes incidence and participants recruitment rates estimations

Primary care centers Lleida	Total attended population	Incident cases diabetes per year (with rate 4/1000)	With a rate of 4/1000 inhabitants and 20% losses	Incident cases of T2DM in 2 years	Minimum number of participants per center	Minimum monthly reclutment (24 months)
Tarrega - Agramunt	26172	105	84	168	126	5
Balaguer+ Artesa + Pons	30462	122	97	195	146	6
Onze septembre	14670	59	47	94	70	3
Bordeta	13354	53	43	85	64	3
Balafia - Pardinyes	21537	86	69	138	103	4
Alcarras	11367	45	36	73	55	2
			Estimated number of cases	752	564	24
Primary care centers Barcelona	**Total attended population**	**Incident cases diabetes per year (with rate 4/1000)**	**With a rate of 4/1000 inhabitants and 20% losses**	**Incident cases of T2DM in 2 years**	**Minimum number of participants per center**	**Minimum monthly reclutment (24 months)**
Mina	15851	63	51	101	76	3
Besòs	27921	112	89	179	134	6
Sant Martí de Provençals	42734	171	137	273	205	9
ABS Sagrada Familia	20727	83	66	133	99	4
ABS Gaudi	25670	103	82	164	123	5
			Estimated number of cases	851	638	27

B. Preliminary results on the cluster with bigdata (non-published results) in collaboration with Research Center for Biomedical Engineering (CREB) of the Universitat Politècnica de Catalunya(Polytechnic University of Catalonia, UPC)

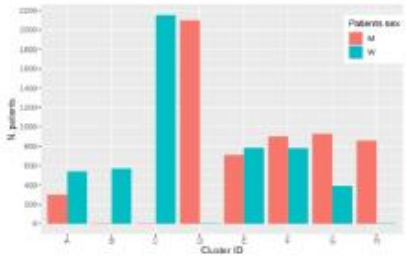

Figure B.1 Sex distribution among the clusters

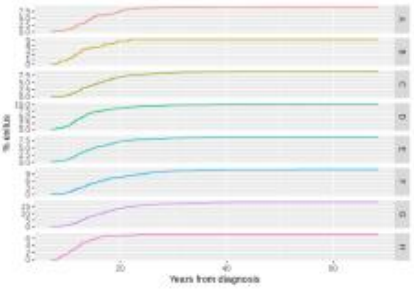

Figure B.2 Mortality distribution among the clusters

ANNEX 9. Study promotional material to recruit patients

# Do you have newly diagnosed type 2 diabetes? In our center, the clinical field study is being carried out: Characterization of subgroups of type 2 diabetes at diagnosis: a necessary step towards precision medicine in diabetes. If you have been diagnosed with type 2 diabetes in the last 3 months, you may be a candidate for this prospective study that is being carried out in Catalonia. With your participation, you can help us better characterize your disease, learn more about its evolution and complications, in order to apply more personalized treatments to each patient.	Procedures: A single study visits with your GP A blood and urine test	If you are interested in participating and would like to receive more information, contact us by email: XXX@XXXX.COM Contact information participating centers: XXX

ANNEX 10. List of centers, principal investigators centers

Center	Principal Investigator Center
Hospital de la Santa Creu i Sant Pau-Barcelona	Dr. Didac Mauricio
Consortium Centers Integral Sanitation	
CAP Sagrada Familia- Barcelona	Dra. Natalia Riera
Centers Institute Catalan Health	
Vilanova University Hospital - Lleida	Dra. Marta Hernandez
CAP Tàrrega	Dra. Neus Miró
CAP Balaguer	Dr. Jesús Pujol
CAP September 11th _	Dr. Monica Solanes
CAP Bordeta	Dr. Antonieta Lafarga
CAP Balafia-Pardinyes	Dra. Sandra Guerrero
CAP Alcarràs	Dra. Mireia Marin
CAP Agramunt	Dra. Meritxell Torres
CAP Eixample (Lleida)	Dra. Laia Llubes
CAP La Mina	Dr. Manel Mata
CAP Besòs	Dra. Elena Hernandez Boluda
CAP Sant Martí	Dr. Francesc Xavier Cos Claramunt
CAP Sarrià	Dr. Gabriel Cuatrecasas
CAP Casc Antic	Dr. Cristian Llacer
CAP El Clot	Dra. Anna Massana Raurich
CAP Poblenou	Maria Isabel Prieto Fernandez

CAP: Primary Care Center;