

Executive summary

Cardiometabolic diseases (CMDs) and their comorbidities currently represent a considerable burden for health care systems in the EU and worldwide. In Europe, the economic burden of cardiovascular diseases (CVD) alone is currently estimated to be €210 billion per year (€370/person) despite these diseases being seen as largely preventable through lifestyle and dietary measures. However, once established, CMDs and particularly CVD, obesity, and diabetes, often remain resistant to treatment and dietary or lifestyle changes and are characterized by a chronic evolution in terms of symptoms. Continually mounting research suggests that alterations in the gut microbiota richness and composition may contribute to alter subject metabolic phenotypes. As such, the gut microbiota may impact host cellular and tissue biology that result in the development and ongoing complications of CMDs and chronicization.

In an effort to elucidate the interaction between the gut microbiota and host health and diseases, MetaCardis set out to apply a systems medicine and multi-level approach. This effort combined metagenomics to understand the gut microbiome and advanced phenotyping approaches in humans (metabolomics and transcriptomics) in existing cohorts from EU and national networks. The strength of MetaCardis consortium was to constitute a newly recruited cohort of over 2000 patients spanning the CMD spectrum.

The specific groups recruited through MetaCardis included patients with 1) Metabolic Syndrome, 2a) Obesity, 2b) Obesity with bariatric surgery, 3) Type 2 Diabetes Mellitus, 4) Acute Coronary Artery Disease without Congestive Heart Failure, 5) Chronic Coronary Artery Disease without Congestive Heart Failure, 6) Chronic Coronary Artery Disease with Congestive Heart Failure, 7) Chronic Congestive Heart Failure without Coronary Artery Disease, and 8) Healthy volunteers. The contribution of patient associations and medical doctors was critical in this context. Over 2700 patients from existing cohorts and the MetaCardis cohort were sequenced for gut metagenomics. Over 2100 serum and urine samples were analyzed for metabolomics and a sub-set of 250 patients were targeted for gene expression through transcriptomics. Data produced from patient phenotyping and omics technologies were centralized into a project data hub in Germany which is mirrored remotely in France. The project data hub is home to numerous methodological and analytical advances integrating high-dimensional data, which were developed from and applied to both existing and the Metacardis cohort.

Through existing cohort analyses, research groups have elucidated how dietary modifications and medical interventions (medication, surgery) influences the gut microbiota and how this impacts the patient phenotype. Efforts have provided networks demonstrating the interactions between specific dietary constituents and bacterial populations. Using results from human studies, the project advanced the mechanistic understanding of the microbiota's effects on host phenotypes using experimental models.

Although MetaCardis has provided many insights into the microbiota's implication in cardiometabolic diseases, it is also in the critical stage of providing the first results on the novel MetaCardis cohort with an aim at examining the microbiota and cardiometabolic diseases, and specifically: 1) the transition from the healthy state to obesity and diabetes and how the microbiota and metabolites associate with disease progression, 2) the escalation from metabolic to cardiovascular diseases, 3) the impact of polypharmacy on gut microbiota profiles in addition to their biological effects, and 4) the characterization of "a healthy microbiota" and its link with circulating metabolites and factors influencing gut microbiota composition.

Project context and the main objectives

The main objective of MetaCardis was to examine the impact of qualitative and quantitative changes in gut microbiota on the pathogenesis of cardiometabolic diseases and their numerous co-morbidities. The ultimate goal of this research objective was to eventually translate clinical and basic scientific research discoveries into new diagnostics and preventive clinical approaches. MetaCardis aimed to achieve these objectives through experimental studies, systems biology approaches, and the recruitment of a deeply phenotyped clinical cohort (the MetaCardis cohort).

To this end, the project was divided into 5 scientific workpackages and 3 workpackages for management, dissemination, and ethics with the overall objective to:

- I. Identify novel targets and shared pathways involved in the natural progression of CMD (from cardiometabolic risks towards more chronic stages such as atherothrombosis with or without heart failure) through a multidisciplinary strategy applying systems biology approaches combining (gut) metagenomics and host /transcriptomic/metabolic phenotyping approaches in patients and healthy subjects characterised for lifestyle and CMD advanced phenotypes.
- II. Validate the contribution of novel gut microbiome-derived targets and prognostic markers identified in original experimental systems including human cellular systems and rodent models as well as in newly recruited CMD patients at successive stages of disease progression.
- III. Deliver new insights into shared pathophysiological targets of CMD and associated comorbidities that can be used as targets for the development of novel therapeutics and diagnostic tools, possibly redefining stages of CMD progression.
- IV. Transfer and disseminate new knowledge to the scientific community, health professionals and scientists via training programs and workshops, and to the public and stakeholders through patient association contributions and MetaCardis project outreach actions (open houses, newsletters, press releases).

The specific workpackage context and objectives within the MetaCardis project are summarized below

WP1: CMD risk biomarkers: gut microbiome and metabolome-derived target identification

Using established national and European cohorts, such as the EU FP7 MetaHIT cohort, the French MicrObes cohort, the Gothenburg study, and the French Microbaria cohort, researchers in WP1 explored existing cohort datasets in-depth, developed analytical infrastructure and made new findings. These cohorts spanned from healthy subjects to diabetic or severely obese patients as well as patients with ischemic heart diseases. The datasets included clinical and biological data as well *de novo* metabolomics profiling of blood and urine to identify proteins, lipids, metabolites and gut metagenomics. Through WP1, the project identified some environmental and gut microbiome markers, which associated with various cardio-metabolic disease phenotypes. These new results were not only published but generated hypothesis that were examined through systems biology (WP5) in the newly phenotyped MetaCardis cohort (WP3).

WP2: Characterization of gut microbiome-derived targets in experimental systems

There is strong evidence of an association between changes in the architecture of gut microbiota and cardiometabolic diseases. This WP brought together experts in translational microbiome studies in humans and preclinical models, in vivo physiology, and molecular signaling to assess causal relationships between microbiota and products of gut microbial activity identified in cohort studies by consortium partners. Transfer of microbiota from patients affected with cardiometabolic diseases to germ free mice has demonstrated the impact of gut microbiota composition on phenotypes that are relevant to diseases (impaired glucose

regulation, obesity). Deeper analyses have underlined the role of specific bacteria in this phenomenon, but researchers in WP2 had also the goal to identify novel underlying host physiological, cellular and molecular mechanisms. In particular, an objective was to find novel gut microbiota-derived metabolites possibly impacting on metabolic health. Thus, these objectives were to provide new insights into the cross-talk between commensal gut bacteria and mammalian host biology through genomic and physiological characterization of products of gut microbial activity in mouse and rat models exhibiting spontaneously or experimentally the main pathophysiological features of cardiometabolic diseases.

WP3: European biobanks and phenotype databases for studies of common cardiometabolic disorders

The major objective of WP3 was to establish a new database with detailed molecular and clinical phenotyping with harmonized European biobank from subjects at different stage of their cardiometabolic disease progression. This was done in order to identify biomarkers and predictors of well-defined stages of CMD development as well as to refine CMD phenotypes in these heterogeneous disorders. Within the MetaCardis project, WP3 focused on cohort recruitment to provide the foundation in responding to novel clinically-driven hypotheses tested in WP5 and validating findings from WP1 legacy cohorts.

MetaCardis aimed to provide one of the first cohorts spanning cardiometabolic diseases and to be characterized in-depth through examination of clinical, biological, stool metagenomics, serum and urine metabolomics, and monocyte and adipose tissue transcriptomics. In order to establish the MetaCardis cohort, clinical centers set out to develop SOPs for 1) phenotyping and lifestyle questionnaires in 3 different languages and 2) biobanking. Patients were then recruited through routine care visits, announcements in newspapers, public advertisements, and through word of mouth involving patients' families. Patient data was collected through a harmonized eCRF in each local country language. Finally, in order to capitalize on the recruited cohort and collected data, clinical center physicians and medical personnel, along with data managers, worked on validating variables and datasets for WP5 examination. The contribution of patient associations was key for this WP.

WP4: High-throughput genomic approaches

The main objectives of WP4 was to complete patient phenotyping with high-throughput "omic" technologies, including (i) metagenomics of the stool samples; (ii) metabolomics of serum and urine samples and (iii) transcriptomics of host immune cells and adipose tissue in patient subsets. This work aimed to sequence over 2700 stool samples for the gut microbiota and comprised three populations: (i) The MetaCardis cohort, (ii) a bariatric cohort, and (iii) a sub-cohort of patients from the MetaHIT project. This resulting work would provide what is currently the largest microbiome characterized CMD cohort with deep phenotyping. Similarly, metabolomics was performed in over 2000 urine and serum samples by untargeted ¹H nuclear magnetic resonance (NMR), gas chromatography coupled to mass spectroscopy (GC-MS), and ultra-performance liquid chromatography coupled to mass spectroscopy (UPLC-MS). Metabolomics was to provide absolute concentrations of lipoproteins and other metabolites as well as serum metabolites more recently linked with the progression of cardiovascular diseases including methylamine derivatives. To provide a link with host gene expression, monocytes from 250 individuals representing different patients groups and healthy controls were also analyzed. Gene expression in 250 adipose tissue samples from subcutaneous and visceral fat was to be examined in bariatric surgery patients before intervention. Highly quality-controlled data were then planned for transfer to WP5 to link omics analyses with numerous clinical phenotypes.

WP5: Data integration and functional modelling

Workpackage 5 aimed to capitalize on efforts from WP1, WP3, and WP4 by integrating high-dimensional clinical, biological, and omics data and to translate this to clinically applicable tools and findings. The ultimate goal of this work was to provide results that would serve for the basis of personalized medicine approaches in subjects with obesity, diabetes and cardiovascular diseases. As such, WP5 was dedicated to developing a datahub at the European Molecular Biology Laboratories in Germany, which was also mirrored in Paris, to

store data from WPs 1, 3, and 4. Work initially centered on properly receiving, storing, and processing the data, and making this data available to each network participant.

In a second step, participating partners were tasked with developing analytical pipelines, which was done with the help from work on WP1 cohorts. Groundwork was laid in metabolic modelling, gut-specific annotations, dietary and other metadata representations, and machine learning techniques to allow MetaCardis to provide a unified understanding of the states of health and cardiometabolic diseases. These efforts were linked together with a repertoire of analytical techniques and algorithms for data integration, exploration, evaluation, and visualization, which have been evaluated, benchmarked and documented. Data was error-checked and shared throughout the consortium, and a framework for team-based analysis of particular sub-questions was established. The project focused on answering clinically-driven novel hypotheses focused on 1) the transition from the healthy state to obesity and diabetes and how the microbiota and metabolites associate with disease progression, 2) the escalation from metabolic disease to cardiovascular diseases, 3) aiming to elucidate therapies influencing the gut microbiota in addition to their biological effects and 4) determining what constitutes a healthy microbiota, its link with circulating metabolites, and factors influencing microbiota composition in healthy controls

WP6: Dissemination, training and exploitation

The main objectives of WP6 were to organize trainings for young researchers of the consortium, enable efficient dissemination of the project's results through scientific publications, lay public articles, press releases, and an international MetaCardis conference. The work package sought to support the exploitation of the results, and clarify any eventual Intellectual Property issues.

Broad ranging expertise in the consortium provided opportunities for meeting interdisciplinary training and education needs in various areas (eg. clinical science, bioinformatics and functional genomics), thus contributing to train the next generation of researchers in the emerging fields of microbiota, metabolism, and systems biology. We aimed to integrate personal development and research training through educational instruments, which will increase statistical and computational skills of wet-lab scientists and physicians, deepen the understanding of bioinformaticians in clinical and fundamental research, and support planning and implementation of research projects.

As the project was expected to generate exploitable foreground and more specifically biomarkers for CMD, business developments and opportunities were monitored carefully during the course of the project in this WP and a final exploitation plans was produced.

The dissemination activities of the project ensured international visibility and acknowledgment of the MetaCardis project.

WP7 Project Management and coordination of the consortium

The goal of WP8 was to frame the scientific work of MetaCardis for all financial, legal, and coordination issues.

WP8: Regulatory and ethical issues

The aim of the WP8 was to ensure that the clinical trial was implemented in compliance with the regulatory requirements and to apply for clinical trial authorization to review boards. Additionally, the workpackage aimed to deal with ethical questions, specifically focusing on the use of big data use in the consortium and for personalized medicine, and communication of results to patients. This work also aimed to involve patients and patient associations to better understand how to communicate these results to patients.

Results and foreground

WP1: CMD risk biomarkers: gut microbiome and metabolome-derived target identification

Using data and biobanks from previously constructed cohorts, WP1 allowed the exploration of hundreds of metabolites and lipids in blood and urine (metabolomes) from hundreds of Europeans. From this work, groups have identified a number of metagenomic and metabolomic signatures associated with specific characteristics of obesity, type 2 diabetes, and ischemic heart diseases.

The overall findings of WP1 has demonstrated the strong interplay between dietary/lifestyle patterns and metabolic parameters and gut microbiota characteristics. This also included the contribution of gut microbiota-derived metabolites. Notably, MetaCardis teams showed that a reduced calorie diet in overweight and obese individuals improved metabolic health and the gut microbiota diversity. Most importantly, physicians and researchers found a stronger increased effect in improved gut microbiota diversity in individuals that have a lower microbial diversity before diet intervention. However, this improvement was not sufficient to ameliorate metabolic health markers at the same level of improvement seen in persons with high gut microbiota richness. On the other hand, examining food intake characteristics, it was obvious that a healthy dietary pattern was also associated with higher gut microbiota richness. As well, this work led to examine the effects of increased abundance of beneficial bacteria, such as *Akkermansia Muciniphila*, which found that obese/overweight subjects with increased abundance of *Akkermansia Muciniphila* had a better metabolic health profile. This work was extended through WP5-based systems biology approaches in determining how gut microbiota populations may affect this bacteria. Through other analytical approaches in WP5 and using ancillary cohort data, the consortium also provided evidence on how nutrients and bacteria may interact to shape gut microbiota populations. Dietary interventions targeting the microbiota were also demonstrated to improve forms of genetic and general obesity in children.

As well as dietary interventions, WP1 focused on obese patients eligible for bariatric surgery and demonstrated that a majority of these individuals exhibit very low gut microbiota diversity, which was found to be associated with worsened metabolic health (higher incidence of type 2 diabetes and increased blood pressure). Microbiota diversity was increased by surgery, but the severe dysbiosis in these patient was not fully corrected after the obesity surgery. Partners have also demonstrated that weight loss from bariatric surgery changes the expression of genes in two types of adipose tissue while a further study linked adipose tissue gene expression following surgery with gut microbiota populations. To further examine the long-term effects of bariatric surgery, groups have provided long-term examination of the gut microbiota in both gastric bypass and stomach stapling procedures while efforts also developed a scoring algorithm to predict 5 year post-surgical diabetes status or relapse of diabetes remission. This is critical and thanks to this cohort recruitment, physicians and researchers developed a clinical tool to detect subjects with a decreased chance of diabetes improvement after bariatric surgery. This tool is currently used in clinical departments with patients entering into bariatric surgery programs. In this context, the precise role of gut microbiota is still being examined.

Collectively, exploration of existing cohorts led not only to novel findings but are also the launching pad for the questions and hypotheses evaluated in the multicentric MetaCardis cohort recruited in the WP3.

WP2: Characterization of gut microbiome-derived targets in experimental systems

The major aim of this work package was to bring a functional dimension to gut microbiota studies carried out by the consortium. This was done through the demonstration of causal relationships between altered gut microbiota and the development of cardiometabolic diseases. Consortium results from metagenomic and metabolomic studies were used for functional studies *in vivo* in preclinical models. To this end, we have established protocols and complementary experimental pipelines to:

- I. Identify, isolate and purify active products of metagenomic clones, which represent the genome of the human gut microbiota,
- II. Test the *in vivo* effects of human microbiota, isolated bacteria, and bacterial genes and metabolites in mouse and rat models of cardiometabolic diseases and
- III. Characterize cellular signaling mechanisms that mediate their effects.

Results from the transfer to germ free mice of microbiota from humans with or without metabolic diseases to germ-free mice have demonstrated that

- I. Subjects undergoing bariatric surgery have an altered gut microbiota and that this altered gut microbiota can, at least in part, mediate the beneficial effects of Roux-en-Y gastric bypass
- II. Transfer of microbiota from some subjects with prediabetes did not result in consistent transfer of impaired glucose metabolism
- III. Stool transfer from subjects with Prader Willi syndrome, a genetic form of severe obesity where individuals have relatively good insulin sensitivity despite being obese, exhibited an altered microbiota, which resulted in relatively improved glucose metabolism when transferred to mice.

Overall, we have been able to demonstrate that larger differences in the microbiota can result in replicating host phenotypes when transferred into germ-free mice, thus indicating that the altered gut microbiota can contribute to metabolic disease.

In addition to demonstrating the contribution of the microbiota to metabolic diseases, we found microbially-produced metabolites, which may contribute to host phenotype including:

- I. Imidazole propionate, a microbially produced histidine metabolite which is increased in plasma of subjects with type 2 diabetes. Short or long-term administration of imidazole propionate resulted in impaired glucose metabolism. This work also led us to analyze the levels of imidazole propionate in the MetaCardis cohort, where we were able to confirm the association between diabetes and this metabolite.
- II. Phenylacetate, which initiates hepatic steatosis in fatty liver disease
- III. 4-cresol, a bacterial product of tyrosine which is negatively correlated with type 2 diabetes. We demonstrated that chronic treatment of rat and mouse models of cardiometabolic diseases with 4-cresol and its co-metabolite 4-methylcatechol induced massively improved glucose tolerance, enhanced insulin secretion and pancreatic islet cell proliferation, and reduced obesity, adipocyte size and accumulation of liver fat, a symptom of non-alcoholic steatohepatitis (NASH).

A series of eight additional microbial metabolites were also administered in preclinical models. Thus *in vivo* administration with metabolites proved to be a promising path to determine causality between gut microbiota and cardiometabolic diseases.

An original approach was also to test the biological function of gut microbial genomes. Metagenomic clones, representing bacterial genomes of the human microbiome, were selected following extensive screening on different cellular models, including human adipocytes. This proved to be a technically and highly challenging but most successful approach to isolate active genomic components of gut bacteria that were tested for their biological role in cardiometabolic diseases *in vivo* and *in vitro*.

An important aspect of research in the workpackage was to understand the molecular mechanisms that mediate the function of products of microbial activity in host organisms. The Gothenburg team in MetaCardis demonstrated the importance of a gut-microbiota derived metabolite called ImP (imidazole propionate) which is a bacterial metabolite of protein metabolism. They showed a strong link between circulating IMP

and the alteration of patients' glucose metabolism. Moreover they described how this metabolism acts mechanistically on the host.

The importance of another bacterial metabolite (4-cresol) in diabetes and insulin resistance was also identified. The teams also characterized TMA (trimethylamine, a bacterial metabolite of choline and carnitine metabolism) as a bioactive molecule having a major impact on signaling networks related to innate immunity and insulin signaling. Deep mechanistic insights were also described thanks to collaborative activity. Finally, the team identified a new biological pathway that was activated by two metagenomic clones. This involved a protein (called the kinase ALPK1) and this protein has genetic variants that appear associated in several genetic studies examining T2D, coronary diseases, and stroke. Thus a novel pathway was found paving the way to new basic research themes.

These results underline the strong cohesion between varied research activities in the workpackage and research in other workpackages and the consortium as a whole. They provide important advances in the field of gut microbiota research and opportunities for further developments.

- I. Our results represent conceptual advances in our understanding of the cross talk between the gut microbiota and the host biology. We have demonstrated that isolated components of bacterial genome and metabolites produced by bacteria have pathophysiological roles and affect host metabolic regulation with strong impact on cardiometabolic diseases. This remains to be extended to other known and as yet unknown microbial metabolites and isolated bacteria and bacterial ecosystems.
- II. Advances have been made in our understanding of the biological impact of gut microbiota from patients and control individuals through transfer to germ free mice. This is a promising research area that will develop through analyses of molecular mechanisms involved in the host following gut microbiota transfer. There are also strong prospects to apply this strategy more broadly in other disease areas. It was also important in this work to show some limits and benefits of mouse models in this context of gut microbiota transfer.
- III. The above outlined results have been obtained following implementation and optimization of methods for the administration of metabolites to germ-free and conventionally raised mice in order to test causality between microbially produced metabolites and phenotypes relevant to metabolic diseases. In a number of experiments we have established refined protocols so that we can perform more acute injections of mice (twice daily over three days) to elevate the levels of the metabolite of interest. We have also developed and validated more chronic delivery systems, either in slow releasing pellets or through osmotic minipumps allowing subcutaneous delivery of compounds. Both approaches resulted in elevation of metabolites in plasma, effects on host metabolism, while maintaining mice germ-free.
- IV. Metabolites synthesised by gut bacteria provide a source of biomarkers for disease diagnosis and prognosis, which have been demonstrated by partners of workpackage 2. A primary focus was done with imidazole propionate, phenylacetate and 4-cresol. This is an area of future developments with other candidate bacterial metabolites that will emerge from metabolome profiling studies, and will be facilitated by above mentioned experimental procedures developed and validated as part of this workpackage. These gut-microbiota derived metabolites will have to be investigated at large to very large scale in European and non European populations to determine their relevance as biomarkers for disease stratification or predictor of progression.
- V. Results from this research thus provide opportunities to develop strategies in personalised medicine that will use information from bacterial ecosystems, individual bacteria and metabolites in targeted therapies. For example, the effects of 4-cresol are being tested against those of drugs currently used in diabetic patients, but its prominent impact on pancreatic islet neogenesis and cell proliferation appear to be an almost unique advantage over these drugs.
- VI. In only a few years, following the characterisation of the pathophysiological roles of microbiota from patients and products of gut bacteria, we have reached the point where investigations into some

molecular mechanisms mediating their cellular effects become essential. We have demonstrated the central role of kinases (MAPK p38g , ALPK1, DYRK1a) in these processes. Kinases are attractive candidates that will require more systematic and deeper analyses in the context of metagenomic research.

WP3: European biobanks and phenotype databases for studies of common cardiometabolic disorders

Through WP3 efforts, the first achievement was the development of standard operating procedures (SOP) and methods for clinical visits, patient phenotyping, and biobanking. The development of these procedures has led to the production of a manual for the consortium, which is of use for future national and international studies. These SOPs are now currently used in several clinical centers for other local studies.

Preparing for cohort recruitment also led to developing the electronic clinical registry form (eCRF), which was a vital tool for collecting clinical, biological, and environmental/lifestyle data in the 3 recruitment centers (Paris, Leipzig and Copenhagen). The development of a complementary biobanking standard operating procedure (SOP) manual for the consortium minimizes sample variability between centers and that samples are handled in line with international standards. Paris has also established a local version of the electronic questionnaire, which provides a means to collect the patient data in MetaCardis follow-up visits or additional cohorts with minimal effort and resources. This activity will be essential for further European activities.

In addition to these advances, clinical centers and Danone Research developed food frequency questionnaires in the 3 countries (France, Denmark and Germany) in order to capture unique dietary trends in this special population. These questionnaires are readily available to researchers to implement in future studies and the public through www.nutritools.org.

The major foreground generated in WP3, however, was the recruitment of the MetaCardis cohort which consists of over 2000 deeply phenotyped patients spanning the cardiometabolic disease spectrum from healthy volunteers to obesity and diabetes to advanced stages of cardiovascular disease and heart failure (Figure 1). In addition to the cross-sectional cohort, in group 2b, a group of patients accepted to pursue their contribution to Metacardis with post bariatric follow-up and fecal sample collection. This was a major effort from international clinicians and patient associations and Metacardis partners to recruit these patients at different stages of disease progression.

	1	1RH	2a	2b	3	4	5	6	7	8	Total
	MetS	Rather Healthy	Obese	Obese - Bariatric	Diabetic	Acute CAD	Chronic CAD without CHF	Chronic CAD with CHF	CHF without CAD	Healthy	
Patients (n)	266	48	293	222	644	159	173	114	36	295	2250
BMI (kg/m ²)	30,4	28,5	42,8	49,2	32,7	27,5	28,7	28,6	31,2	22,4	32,7
Age (y)	60,4	56,7	43,7	45	60,1	57,7	63,2	62,6	56	54,8	46,1
Male (%)	50,7	56,2	22,5	27,5	51,2	82,6	86,1	80,7	69,4	39	50,3
Female (%)	49,3	43,8	77,5	72,5	48,8	17,4	13,9	19,3	30,6	61	49,7

Table 1: The recruited MetaCardis with number of patients per group and group BMI, age, and sex distribution

Owing to the depth of cohort phenotyping, clinical, biological, lifestyle, socioeconomic and dietary data total well over 1000 variables alone, which is further enriched by multi-omics data coming from serum and urinary metabolomics, gut metagenomics, and adipose and monocyte transcriptomics in WP4. These clinical and biological variables were examined in-depth and analyzed statistically to develop cohort descriptive characteristics. In addition to the data generated from cohort recruitment, a vast array of biological samples

(serum, plasma, urine, stool, saliva, adipose tissue) collected during the study remain available through the partners' local biobanks.

Through stringent data management, the clinical centers rendered an easily usable phenotypic database with very high-quality. This has been done for an initial priority list of variables; however, further data that are relevant to clinical hypotheses are consistently being examined and process through the developed data management pipeline for partners. While these data are currently available for partners through the EMBL-housed datahub established in WP5 and mirrored in Paris, cohort clinical data will also be available for future projects and external consortia. For instance, ancillary studies involving MetaCardis consortium members have been approved and are in progress and using MetaCardis clinical data to establish potential markers of disease progression through complementary measurements in the MetaCardis cohort. However, while the clinical data is not yet available to external institutions, the consortium is in the process of developing a data access charter centered on FAIR data access principles. These data will eventually be available for external groups and consortia in the future.

Beyond cohort recruitment, WP3 efforts have also led to important findings on the MetaCardis cohort including knowledge generated on immune cells through the spectrum of cardiometabolic diseases. To this end, the Paris group grouping experts in metabolic diseases and in cardiology found that Mucosal associated invariant T (MAIT) cells, which are immune cells known to be stimulated by gut microbiota metabolites and are important mediators of the immune system, are significantly depleted with cardiometabolic disease progression and hyperglycemic episodes. Indeed, increasing glucose levels seem to affect the vitality of these important immune cells involved in the regulation of immune status.

As a further step in patient phenotyping beyond standard measures, Paris and Leipzig have focused on adipose tissue characterization of severely obese patients eligible for bariatric surgery. This work allows insight into the adipose tissue fibrosis (scarring), cellular inflammation, and histological (microscopic) characteristics in this unique population. These factors are markers of tissue alteration during metabolic disease development, which play a pivotal role in patient phenotype and treatment response. As such, this allows the clinical centers to pursue the potential correlation of adipose tissue physiology to bariatric surgery outcomes and their relationship to high-throughput omics phenotyping.

WP4: High-throughput genomic approaches

The main results generated from WP4 are the determination of omics profiles of the human gut microbiota, host and gut microbiota -derived metabolites. Moreover gene expression data from circulating inflammatory cells (monocytes) and from adipose tissue samples were also collected in patients subsets, the latter of which was done through patients accepting subcutaneous adipose tissue biopsies.

First, the project aimed at generating gut metagenomics profiles for 2000 cohort stool samples. In a first step, samples were sequenced using SOLiD technology while generating the profiles using a 3.3 million gene catalogue. This resulted in determining the relative abundance of the genes for each individual, and data are stored on the project data base hub constructed in WP5. However, during the course of the project, technological advances in the field allowed researchers a better sequence the gut microbiota with lower error rates compared to previous technology. Additionally, extending the patient inclusion period and adding longitudinal follow-up of bariatric patients resulted in the number of samples increasing to 2700 total stool samples. The entirety of the 2700 samples were sequenced with Ion proton technology, which allows for more user-friendly data while maintaining concordance with previous (90% correlation with SOLiD results) and other current (Illumina) sequencing technology. This was performed for all Metacardis patient groups, including bariatric surgery patients before and after the intervention.

In an effort to provide more in-depth insight into the microbiota populations, a new reference gene catalogue, mapping 9.9 million microbial genes, was used to generate the gene profiles. Not only was the

amount of sequence information thus increased 2-fold, but the number of genes profiled was increased 3-fold as well. Linked to the MetaCardis sequencing, INRA (Paris) also developed analytical pipelines and took advantage of automated IT infrastructure to treat this data. Beyond gene profiles, the metagenomics species profiles were generated. For this, the reference microbial gene catalogue was organised into approximately 1500 metagenomics species (MGS), using software developed in a previous European project, MetaHIT and improved at INRA during MetaCardis timeline. An MGS is a cluster of genes encoded by a bacterial species, most of which were never cultivated. MGS profiles were transferred to the data hub and became the main staple for correlating microbiome profiles with clinical, biological (glucose metabolism and low-grade inflammation), nutritional, lifestyle and metabolomic readouts.

During the course of the project, the landscape of metagenomics analyses also shifted with the proposal of quantitative gut metagenomics, which are derived from stool bacterial cell counts. As this was developed by the Flemish research group within the consortium, MetaCardis is also the very first large-scale cohort to benefit from this representation of metagenomics.

For metabolomics, the project measured the metabolites of 2000 serum and urine samples performed in London. Through three methods of metabolomics, this work resulted in 2,162 urine and 2,251 serum samples profiled and transferred to the data hub. Similar to metagenomics, MetaCardis led to developments in experimental and computational approaches for metabolomics including new algorithms and methods to determine absolute concentrations of metabolites in the circulation and urine. The project also measured specific metabolites that were more recently discovered as important marker candidates for cardiovascular and metabolic diseases. Together, this work resulted in over 112 lipoproteins, 27 serum and 49 urinary metabolites through an initial technique coupled with semi-targeted 102 serum metabolites in a second technique and 15 CVD-associated metabolites. All metabolomics data were transferred to the project data hub, as expected, for data analysis and integration.

To complement data mostly from microbial origin and extend patient phenotype characterization, partners produced gene expression profiles in host blood monocytes and adipocytes of a patient subset. Selection was designed to include 15-30 individuals per patient group (total n=210) and 40 sex- and age-matched controls. In parallel, adipose tissue gene transcription analysis were carried out in paired visceral and subcutaneous adipose tissue samples from patients who underwent bariatric surgery in Paris and Leipzig. This resulted in the production of gene expression profiles for these patients, which are being linked to clinical and metagenomics data in an effort to respond to novel clinically-driven hypotheses.

WP5: Data integration and functional modeling

Workpackage 5 led to the creation of many analytical models and tools coming from data generated in WPs 1, 2, 3, and 4. Within the project objectives, the major development in WP5 was the creation of the MetaCardis data hub, which is physically housed on a server at the European Molecular Biology Labs (EMBL) in Germany and mirrored in Paris. The hub is a persistent, secure, access-controlled depository of safely anonymized clinical and high-throughput data from the MetaCardis cohort. Collectively, the hub forms the basis for all analysis within the project and future use of the data for external requests or internal collaborative projects. Through work in WPs 6 and 8, the governance of the datahub will be assured by the MetaCardis FAIR data sharing charter and by partnering investigators representing the MetaCardis scientific committee. As well, access to and use of this data is governed by study sponsor's regulations and applicable law.

Throughout the project, bioinformatics groups have also provided their expertise in creating analytical tools for data generated from work packages 1 – 4, which includes:

- GMM models. A database collections of bacterial gene families, which can serve to describe and quantify health-relevant gut microbial functional activity, thus empowering metagenome analysis of microbiome cohorts.
- MOCAT Toolbox. Tools for processing metagenomic sequencing data and providing microbiota composition and functional potential.
- 9.9 million microbial international gene catalogue, which is a database of microbial gene families defined at the species level. Combined with functional models and tools described above, this allows reproducible human gut microbiome analyses consistently across studies.
- mOTU tool. Software for resolving and quantifying microbial operational taxonomic units (OTU) from shotgun sequencing data. It is based on the use of marker genes extended to cover even uncultured gut microbiota.
- proGenomes tool. Database for systematic and standardized microbial genome annotation, which is a comprehensive and regularly updated repository of microbial genes and proteins with functional information
- Data integration toolbox. Covariate control toolbox that controls for the influence of treatment, disease, and risk in large-scale cross-sectional -omics analysis. Currently written up in a standardized software package
- Bayesian Network reconstruction tools, which can trace causality and directional influences of clinical, metagenomic, and metabolomics data
- CASINO toolbox that allows the Metabolic network analysis toolbox available. These resources allow for flux balance modelling of the metabolic impact including expected steady states of a composite community of host and microbiome cell entities.
- GEM models. Precise and detailed metabolic network models have been developed for 100+ central gut commensals that provide input in the CASINO toolbox
- GOMixer toolbox. Software for analysis and visualization of microbiome functional module data
- iPath viewer. Online software for differential visualization and testing of microbiome functional module data. It allows for contextual assessment of the functional role of changes in gut microbiome composition.

Through the use of these analytical approaches and interactions with work on WP1 ancillary cohorts, specifically the MetaHIT cohort, WP5 also contributed to the knowledge of the influence of the diabetic medication, metformin, on gut microbiota. Additionally, the work package also contributed to the discovery of microbially produced BCAAs as a risk factor for insulin resistance. Work package 5 was also central in producing clinically interpretable results on the MetaCardis cohort using data input from WPs 3 and 4. Work on the MetaCardis cohort focused on four principle topics, which will lead to better understand the relationship between the microbiota and cardiometabolic diseases: 1) the transition from the healthy state to obesity and diabetes and how the microbiota and metabolites associate with disease progression (2) the escalation from metabolic disease to cardiovascular diseases, and 3) the determination of the impact of polypharmacy ininfluencing the gut microbiota in addition to their biological effects in patients with cardiometabolic disoredres and 4) the determination of what constitutes a healthy microbiota, its link with circulating metabolites, and factors influencing microbiota composition. This latter work is currently examined in healthy subjects (group 8).

WP6: Dissemination, training, and exploitation

WP6 was the work package dedicated to dissemination, training and exploitation and thus no foreground has been produced in this WP. One of the main outcomes of the WP were the organized short term fellowships between the different partners allowing young researchers to acquire new methods not available in their laboratory, observe laboratory operations in a foreign country, and reinforce the cooperation between

partners. The applications for the fellowships were evaluated by an internal education committee. A total of three fellowships took place during the project's lifetime. During the project period, MetaCardis teams welcomed 9 PhD and 17 post-doc fellows in the project, who contributed to each MetaCardis meeting and presented their results. This was also particularly important during the last conference international meeting.

For these young investigators, several MetaCardis workshops on different topics/methods were also organized. These workshops focused on providing junior researchers technical advice on their analyses and experiments, presenting up to date and important results from partners, as well as fostering the multidisciplinary approach of the project through presenting and discussing biological, clinical, and analytical research topics.

At the end of the project, the MetaCardis conference was organized with over 100 attendees and half of the speakers were junior fellows who presented MetaCardis project results. Group leaders of the MetaCardis consortium also presented their work and had the opportunity to present MetaCardis consortium project results in international conferences throughout the project. The younger investigators then had the opportunity to create their own network which will be critical in the future development of their careers and for some of them as team leaders.

Dissemination was also performed through academic publications (88 publications already published and many more to come), presentations in congresses and international conferences (160 in total), and media dissemination. Many more presentations will be given after the official end of the project.

For the exploitation of results and Intellectual Property Management, a handbook and training session had been organized and distributed to the partners at the start of the project in order to inform them on these issues. An exploitation panel with the different Intellectual Management Departments of the partners had been set up in the first half of the project and several telephone calls for monitoring the intellectual property output of the projects had been organized.

In order to legally frame the establishment of the MetaCardis cohort and biobank a legal agreement had been negotiated and signed by all partners in period 3. In the last period the sponsor of the MetaCardis study initiated an agreement on FAIR sharing principles for the MetaCardis datahub and between all partners on the post-project management of the cohort and its data access.

WP8: Regulatory and ethical issues

The first major achievement for WP8 was the submission and acceptance of the clinical protocol due to its synergistic efforts between WP3 partners. Due to the emerging importance of the interpretation of omics data in clinical investigations, WP8 also focused on producing valuable ethical questions related to the project. These topics ranged from data interpretation and limitations for researchers to communication between physicians and study participants.

With the help from external ethicists, the project produced:

- Direct communication and exchanges at consortium meetings with MetaCardis members spanning biological and analytical researchers to clinical investigators
- A brochure for patients with the aim to communicate the objectives of our work and the results as well as its potential consequences (e.g. for future treatments) in lay terms.
- Training for mainly medical staff and with patient associations on the communication of the project results to patients
- A report on the ethics of using big data in MetaCardis and the reliability of statistical results, the use and selection of algorithms in personalized medicine as well as the issue of data anonymization and IP on data in personalized medicine.

Socio-economic impact and the wider societal implications of the project

WP1: CMD risk biomarkers: gut microbiome and metabolome-derived target identification

The immediate impact of results from WP1 are through the numerous scientific publications in the consortium. In total, the project has resulted in approximately 90 publications, a majority of which came from work on existing cohorts through WP1. These findings were also summarized in over 150 presentations to scientific audiences and press releases of partner institutions as well as highlights on websites to the general public. Within these findings, the discovery of the specific effects of diet on the microbiota and health as well as specific metabolites or specific bacterial species (such as *Akkermansia muciniphila*) markers of metabolic improvements may be key to establish biomarkers or therapeutic targets with further confirmatory or follow-up research. In addition to this, these published works pave the way for future (gut microbiota based) methods to develop strategies for patients stratification in the context of complex cardiometabolic diseases. The work performed in severe obesity (bariatric surgery condition) also led to the development of new therapeutic projects aiming at improving the gut microbiota diversity via fecal transfer in patients with the deterioration of glucose metabolism. This clinical investigation will start in 2019 in Paris. Finally, these results provide a foundation for analysis in the MetaCardis cohort while also improving knowledge for scientific and clinical communities of the implications of the microbiota in health and diseases.

WP2: Characterization of gut microbiome-derived targets in experimental systems

While most microbiome studies focus on associations between microbiota and different human diseases, WP2 has made significant progress moving beyond associations to demonstrate causal relationships between microbiota and disease followed by several mechanistic insights. Eventually, these results could be used for targets of industrial applications in novel therapeutic solutions in cardiometabolic diseases and personalized medicine.

The demonstration that an altered gut microbiota contributes to metabolic diseases may open up for industry to identify bacteria that are enriched or depleted in these common diseases. Microbiota inoculation experiments in preclinical models allowed us to select bacterial ecosystems that have a beneficial impact on the host metabolism. Such findings are excellent targets both from diagnostic as well as therapeutic targets and could therefore have an increased economic impact on industry in Europe.

Furthermore, moving beyond the bacteria themselves and focusing on the metabolites that an altered microbiota generate allows us to assess the microbiota output, rather than individual strains of bacteria, which may act differently depending on the ecological networks or the diet that is ingested by the host. These microbial metabolites represent molecular biomarkers that could be screened for detection of early signs of metabolic disorders or to follow their progression. Such metabolites can not only be considered as biomarkers, but may also constitute drug targets that can be modulated pharmacologically or by manipulating the microbiome. As such, a patent was proposed regarding a breakthrough result obtained by a MetaCardis team.

Thus in general terms, our results contribute to raise awareness of the general public on the gut microbiota mediated effects of diet on human metabolism and health. The public probably understands the role of the microbiota in human health but results from our research in WP2 can contribute to deepen awareness to specific natural molecular compounds and additives in food that may have beneficial or adverse effects on health.

WP3: European biobanks and phenotype databases for studies of common cardiometabolic disorders

From the start of the WP3 work, the development of standardized operating procedures (clinical phenotyping, biobanking) can serve to harmonize future clinical research studies in an effort to decrease inter-study discrepancies. These standardized operating procedures are currently being used by different clinical investigation centers in different projects.

As well, the development of a nutritional food frequency questionnaire for the cohort provides the research community a standardized means to assess dietary patterns in cardiometabolic disease patients, which was disseminated at the international level. This is especially important because in many studies investigating gut microbiota in different disease, lifestyle information and especially dietary data are still scarcely recorded. The development of this tool in Metacardis will help researchers in the field to assess this important lifestyle aspect. This can also enhance the understanding of nutritional habits and challenges in this population in link with an altered gut microbiota.

The development of an eCRF in local languages and English also provides accessibility for future projects to use a standard means to collect patient information with common core data sets. Indeed interoperability between large scale cohorts has become critical.

The phenotypic characterization through exploration of adipose tissue in bariatric surgery patients may provide a mean to assess specific bariatric outcomes pre-surgery via scoring and deep-learning, which builds off previous work from consortium partners. This could also improve patient care through identifying alternative care pathways for individuals with less favorable outcomes.

The largest impact of WP3, however, likely will come to fruition from novel clinically-driven tested hypotheses tested on the recruited cohort in WP5. This work should generate new biomarkers of disease stage progression as well as a better understanding of variables impacting cardiometabolic diseases and providing opportunities for precision medicine approaches in treating these diseases. Additionally, the data on the MetaCardis cohort can continue to be exploited for research questions from internal and external investigators to advance cardiometabolic disease understanding.

WP4: High-throughput genomic approaches

The work carried out in WP4 has likely generated the largest dataset of multi-omics characterization of the human microbiome and metabolome in a cardiometabolic cohort with deep phenotyping (clinical, biological, lifestyle profiling). Through analyses on this data (detailed in WP5), these datasets provide the basis for microbiome in health and disease in the cohort.

Genome-wide gene data will be used to detect functional connections between gut microbiome and host metabolome data. Of direct relevance to MetaCardis, correlations between transcriptome and metagenome data will identify functional relationships between bacterial ecosystems and/or genes and host biological pathways in organs involved in cardiometabolic diseases. In parallel, results from analysis of associations between patterns of gene transcription in fat and monocytes and metabolome signals will be mined to detect co-regulated genes and host circulating metabolites of bacterial origin (including, for example. those used for in depth functional analyses in workpackage 2), which will underline the impact of microbial metabolism on host genome expression.

Beyond the use of the WP4 generated dataset within the project eventual open use of the MetaCardis datasets with those generated by ambitious world-wide efforts in many countries (Europe, USA, China...) will contribute to the community efforts to better understand role of the microbiome in health and disease.

WP5: Data integration and functional modeling

The immediate impact of WP5 is through the collaborative results produced in WP1 and analytical toolboxes developed for the scientific community. The development of these tool boxes provide significant advancement into methods to analyze and integrate “omics” data with bioclinical variables.

To this end, MetaCardis has provided knowledge of important variables to control in studies examining the microbiota in health and disease and the means to interpret omics and clinical data. This is especially true when considering the importance of fecal bacterial load when analyzing metagenomics data. We believe this aspect, beyond these initial developments, WP5 has resulted in the creation of an immense datahub, which will serve future scientific collaborations and eventually external research groups and the public. This will undoubtedly lead to further discoveries beyond the life of the project, especially when used in coordination with other EU-funded deeply phenotyped cohorts.

Within the project, the use of WP5 analyses will contribute to the understanding of the microbiota in health and disease (cardiovascular disease, obesity and metabolic syndrome). Using data from WP4, correlating the microbiome and metabolome data could identify molecules specified by the microbiome, and thus infer functions playing a role in the evolution of the disease. This should open avenues to develop treatments for alleviation of risks for disease progression, which would be of a considerable impact on public health. Additionally, accounting for diet and polypharmacy, the analyses could reveal predictive signatures of the disease progression, aiming to identify at risk individuals via omics signatures. In addition to this, the impact of polypharmacy on gut microbiota and derived metabolomics factors will provide essential information to the pharmaceutical sector, where it is well-known that patients will respond differently to a given drug.

These results should thus immediately impact the scientific and clinical community while possibly leading to industrial applications (i.e. food-grade products, individualized responses to therapeutics, etc) with further mechanistic confirmation. As well, eventual recommendations for dietary, therapeutic, or lifestyle modifications may be derived from these findings after further confirmation in additional cohorts, longitudinal follow-up, or interventional studies. This latter aspect is already ongoing, where some Metacardis partners have already undertaken specific dietary interventions in Metacardis groups to patients to evaluate the impact of these dietary changes in metabolic health of type 2 diabetic patients.

WP6: Dissemination, training, and exploitation

The different activities of WP6 provided training and collaborative experiences for young researchers in the consortium, who were able to acquire new experimental methods and analytical approaches. These experiences also provided an opportunity for them to share and present results with internationally recognized scientists in the field of gut microbiota and with clinician involved in cardiometabolic disorders. One result of these collaborations would be a greater experience in the fields and established network leading to future career opportunities of well-trained young investigators.

Outside of the scientific community, dissemination activities targeting the lay public have improved the public’s knowledge on the gut microbiota. Importantly, this allowed researchers to interact with the public and provide scientifically or clinically-backed information, which can clarify confusion created through media and marketing campaigns. As well, press releases on the project will have an impact on the notoriety of the field of gut microbiota research in the general public.

The final conference of MetaCardis allowed to disseminate the results to an international audience and impact on future collaboration as a follow-up of the project. Overall the dissemination activities of the project ensured international visibility and acknowledgment of the MetaCardis project.

WP8: Regulatory and ethical issues

Work on ethics centered on interpretation of the results at a societal level and communication with the patients with the contribution of patient associations. Several aspects regarding impacts have to be

mentioned. The brochure produced for patients, as well as the training sessions designed for the physicians, will impact the communication with patients on i) the dissemination of information in link with the importance of gut microbiota in the pathophysiology of cardiometabolic diseases and ii) more specifically, awareness linked to the research results of the Metacardis project. The brochure and content of the training will be made available on the MetaCardis website enabling thus to broaden the impact to a general audience beyond the project. Within the scientific community, an article derived from the ethical report will be published in the future in order to disseminate more broadly the ethical reflections discussions especially regarding issues around big data, machine learning and personalized medicine in the cardiometabolic field and more broadly in other chronic disorders.

Public website address and contact details

<http://www.MetaCardis.net>

Scientific coordinator:

Prof. Karine Clément MD, PhD
Sorbonne University, INSERM UMR1269 NutriOmics research Unit
Assistance Publique Hôpitaux de Paris, Nutrition department;
Pitié-Salpêtrière Hospital,
91 boulevard de l'Hôpital
75013, Paris, France
karine.clement@inserm.fr