**APPLICATION TO THE ESTONIAN COMMITTEE ON BIOETHICS AND HUMAN RESEARCH**

**FOR ETHICAL EVALUATION OF THE RESEARCH PROJECT**

|  |  |  |
| --- | --- | --- |
| **1. Name of the study (in case of an application in English, the name of the study in Estonian is required in parallel)** | | |
| Systemic Endotoxemia as a driver of chronic inflammation: Searching novel biomarkers and therapeutic targets for Arthritis (ENDOTARGET).  Süsteemne endotokseemia kui kroonilise põletiku põhjustaja – artriidi biomarkerid ja uued ravieesmärgid (ENDOTARGET). | | |
| **2. The main purpose of the study (up to 450 characters / 0.25 pages) (in case of an application in English, the main purpose of the study should be provided in Estonian, too)** | | |
| In this project, we aim to explore the relationship between gut microbiota, intestinal permeability, and systemic endotoxemia (SE). Further we aim to understand their role as drivers of disease onset and disease activity in osteoarthritis (OA), rheumatoid arthritis (RA), and spondylarthritis (SA), as well as targets of preventive and therapeutic approaches. We will study the events leading to disease onset by taking advantage of the data and blood plasma and feces samples already available in Estonian Biobank, and also collect additional data and samples to add another timepoint to the analyses and confirm our hypotheses. We will evaluate the associations between genetic variation, metabolome, microbiome, blood markers, different clinical traits and defined outcomes of interest (SE, OA, RA, SA) to identify potential predictors. This data will be further used to construct risk prediction models for studied RD diseases.  Selle projekti eesmärk on uurida seoseid soolestiku mikrobioota, soolestiku läbilaskvuse ja süsteemse endotokseemia (SE) vahel. Lisaks püüame mõista nende mõju osteoartriiti (OA), reumatoidartriiti (RA) ja spondüloartriiti haigestumisele (SA) ning haiguse aktiivusele, samuti nende mõju haiguste ennetamise ja ravimeetodite valikutele. Uurime haigestumiseni viivaid sündmusi, kasutades Eesti geenivaramus olemasolevaid vere plasma ja väljaheiteproove ning kogume ka täiendavaid andmeid ja proove, et saada teine ajapunkt ning kinnitada meie hüpoteese. Projektis tahetakse leida meid huvitavate reumaatiliste haiguste ja süsteemse endotokseemia varajasi markereid, selleks hinnatakse nende seoseid geneetiliste variatsioonide, metaboloomi, mikrobioomi, veremarkerite, erinevate kliiniliste näitajatega. Seejärel, kasutatakse neid andmeid uuritud reumaatiliste haiguste riskiennustusmudelite koostamiseks.  For the storage and research of tissue samples outside the territory of the Republic of Estonia, we request the corresponding permission from the Senate of the University of Tartu.  We submit the application in English, as many partners are from abroad and do not speak Estonian. | | |
| **3. Principal investigator(s) and their contact details** | | |
| **Given name(s): Reedik**  **Last name: Mägi**  **Position:** Professor in Bioinformatics  **Institution:** University of Tartu, Faculty of Science and Technology, Institute of Genomics (UTARTU)  **Phone:**7374045  **e-mail:** reedik.magi@ut.ee  **Skype:** reedik\_m | | |
| **4. Other researchers involved in the study (add lines as necessary)** | | |
| **Researchers from University of Tartu**   1. Given name(s): Ene   Last name: Reimann  Position: Research Fellow of Orthogenomics  Institution: University of Tartu, Faculty of Science and Technology, Institute of Genomics (UTARTU)   1. Given name(s): Kristi   Last name: Läll  Position: Research Fellow of Statistical Genetics  Institution: University of Tartu, Faculty of Science and Technology, Institute of Genomics (UTARTU)   1. Given name(s): Helene   Last name: Alavere  Position: Head of the Data Collection Office  Institution: University of Tartu, Faculty of Science and Technology, Institute of Genomics (UTARTU)   1. Given name(s): Karoliina   Last name: Kruusmaa  Position: Specialist  Institution: University of Tartu, Faculty of Science and Technology, Institute of Genomics (UTARTU)   1. Given name(s): Triin   Last name: Laisk  Position: Associate Professor of Genomics and Reproductive Genetics  Institution: University of Tartu, Faculty of Science and Technology, Institute of Genomics (UTARTU)   1. Given name(s): Kristjan   Last name: Metsalu  Position: Head of IT Development Unit  Institution: University of Tartu, Faculty of Science and Technology, Institute of Genomics (UTARTU)  9. Given name(s): Elin  Last name: Org  Position: Professor  Institution: University of Tartu, Faculty of Science and Technology, Institute of Genomics (UTARTU)  10. Given name(s): Oliver  Last name: Aasmets  Position: Research Fellow of Statistical Metagenomics  Institution: University of Tartu, Faculty of Science and Technology, Institute of Genomics (UTARTU)  11. Given name(s): Kertu-Liis  Last name: Krigul  Position: Junior Research Fellow, PhD Student  Institution: University of Tartu, Faculty of Science and Technology, Institute of Genomics (UTARTU)  12. Given name(s): Kreete  Last name: Lüll  Position: Specialist  Institution: University of Tartu, Faculty of Science and Technology, Institute of Genomics (UTARTU)  13. Given name(s): Annabel  Last name: Klemets  Position: MSc student in GI Research Group of Microbiome  Institution: University of Tartu, Faculty of Science and Technology, Institute of Genomics (UTARTU)  14. Given name(s): Kateryna  Last name: Pantiukh  Position: MSc student in GI Research Group of Microbiome  Institution: University of Tartu, Faculty of Science and Technology, Institute of Genomics (UTARTU)  15. Given name(s): Galadriel Luzia  Last name: Velazquez Silva  Position: PhD student in GI, Junior Research Fellow in Genomics  Institution: University of Tartu, Faculty of Science and Technology, Institute of Genomics (UTARTU)  16. Given name(s): Nele  Last name: Taba  Position: Research Fellow of Statistical Metabolomics  Institution: University of Tartu, Faculty of Science and Technology, Institute of Genomics (UTARTU) | | |
| **Researchers from the partner institutions:**  1. Given name(s): Aare  Last name: Märtson  Position: Orthopaedist, Professor of Orthopedics  Institution: Tartu University Hospital, Traumatology and Orthopaedics Clinic (TUH) and University of Tartu, Institute of Clinical Medicine (UTARTU)  2. Given name(s): Kaspar  Last name: Tootsi  Position: Orthopaedist, Research Fellow in Orthopedics  Institution: Tartu University Hospital, Traumatology and Orthopaedics Clinic (TUH) and University of Tartu, Institute of Clinical Medicine (UTARTU)  3. Given name(s): Egon  Last name: Puuorg  Position: Orthopaedist  Institution: Tartu University Hospital, Traumatology and Orthopaedics Clinic (TUH)  4. Given name(s): Katrin  Last name: Ulst  Position: Rheumatologist  Institution: Tartu University Hospital, Internal Medicine Clinic (TUH)  5. Given name(s): Raili  Last name: Müller  Position: Rheumatologist, Rheumatology lecturer  Institution: Tartu University Hospital, Internal Medicine Clinic (TUH) and University of Tartu, Institute of Clinical Medicine (UTARTU)  NOTE! Partners from Tartu University Hospital will support the ENDOTARGET project with the clinical know-how and do not meet the BPs or get to know their identity. They will see only the pseudonymised health records and analysis data.  6. Given name(s): Goncalo  Last name: Barreto  Position: Principal investigator  Institution: Helsinki University Hospital, Department of Rheumatology, Helsinki, Finland (HUS)  7. Given name(s): Jukka  Last name: Parantainen  Position: PhD student  Institution: Helsinki University Hospital, Department of Rheumatology, Helsinki, Finland (HUS)  8. Given name(s): Kari  Last name: Eklund  Position: Professor, rheumatologist  Institution: Helsinki University Hospital, Department of Rheumatology, Helsinki, Finland (HUS)  9. Given name(s): Johanna  Last name: Ferrero  Position: Postdoctoral researcher  Institution: Helsinki University Hospital, Department of Rheumatology, Helsinki, Finland (HUS)  10. Given name(s): Sirkku Peltonen  Last name: Peltonen  Position: Head Physician  Institution: Helsinki University Hospital, Department of Rheumatology, Helsinki, Finland (HUS)  NOTE! Partners from Helsinki University Hospital will support the ENDOTARGET project with the clinical know-how and do not meet the BPs or get to know their identity. They will also receive pseudonymized blood plasma samples for measuring the endotoxemia markers and summary statistics from the Estonian Biobank (EstBB) data analyses. They will not receive any individual level data.  11. Medical service provider. As the procurement has not been conducted yet, we cannot give the exact name of the service provider and its contact person.  NOTE! The medical service provider to collect the samples and carry out the blood marker analyses will be chosen through a procurement process among service providers accredited by the Estonian Accreditation Centre and holding the necessary permits, licenses and certificates for the provision of the laboratory services. Since we will be recruiting BP-s for the study from across Estonia, the service provider will be chosen based on the extent of coverage of their medical clinics and ability to receive study subjects and collect samples at a collection point as close to the individuals as possible.  The service provider will support the ENDOTARGET project by collecting and transferring the blood samples, transferring the feces and saliva samples to GI, and performing RD blood marker analyses. In order to do so, the medical staff at the clinics must meet the BPs and validate their identity via passport or ID card. They will not receive any additional individual level data. | | |
| **5. Financing of the study** | | |
| **Sources of funding** | HORIZON-HLTH-2022-STAYHLTH-02-01, No: 101095084 (01.01.2023-31.12.2026) | |
| **Total cost of the study (amount)** | Total cost of the ENDOTARGET project is 6,997,891.00 €, allotment of the University of Tartu is 598,000 €  <https://www.etis.ee/Portal/Projects/Display/0d03c4f7-493a-4f03-93f0-cb16b1d842ae> | |
| **Financial compensation for the study participants (yes, no, explanation and amount)** | There will be no financial compensation offered for the study participants. The feces collection kit will be sent to participants via post, and they can leave the collected samples to the blood collection centre they prefer (e.g. closest to their home). The blood and saliva samples can be taken also at the blood collection centre. | |
| **Insurance provided for the study participants (yes, no, name of the insurance company and the certificate of insurance (COI))** | No | |
| **6. Study period (the beginning and end dates (MM/YYYY))** | | |
| 01.01.2024 – 31.12.2030  The study period is longer than the period of ENDOTARGET project with EU funding as the data analysis will take place also after the official end of the ENDOTARGET project. | | |
| **7. Information about previous or parallel evaluation of the same study project (incl in other countries)** | | |
| The project has been evaluated by EBIN and the decision has been given in 21.st of December 2023 (No 1.1-12/4564). | | |
| **8. Brief overview of previous studies on the same topic (up to 900 characters / 0.5 pages)** | | |
| Rheumatic diseases (RD) affect more than 5 % of Europe's population and cause significant disability, pain, and reduced lifespan. Rheumatoid arthritis (RA), spondyloarthritis (SA) and osteoarthritis (OA) are the most common causes of articular destruction in Western societies.1,2 Although there has been during recent years significant progress in the treatment of RA and AS still significant proportion of patients do not respond to treatment, which can lead into significant morbidity, loss of function, inability to work and human suffering. There are no effective targeted drugs available for treatment of OA and only efficacious treatment is the surgical replacement of the arthritic joints. The pathogenesis of these diseases is incompletely understood. Dysbiosis of intestinal microbiota i.e. deleterious alterations of the composition and/or function of the gut microbiota has been implicated in the pathogenesis RA, AS and OA.3 However, the mechanisms through which intestinal dysbiosis contributes to disease pathogenesis are largely unknown. Leakage of harmful compounds from intestinal bacteria in particular immunologically highly active bacterial wall lipopolysaccharides (LPS) is a significant source of systemic inflammation posing a high proinflammatory burden on the body.  High systemic levels of LPS are referred to as “systemic endotoxemia (SE)”. It is promoted by intestinal dysbiosis, compromised intestinal barrier function and lifestyle factors. Increased gut permeability facilitates the translocation of gut bacteria and bacterial effector moleculs across the gut-blood-barrier to reach systemic circulation. LPS stimulates the innate immune system and cells of the target organs via the activation of Toll like receptors (TLRs). They can also promote autoimmunity by activating adaptive immunity. Our recent study showed that blood LPS biological activity correlates with the activity of RA and is also a significant independent predictor of the treatment response.4  1 https://cordis.europa.eu/article/id/97231-ep-calls-to-recognise-the-extraordinary-burden-of-rheumatism-and-arthritis  2 European Alliance Of Associations For Rheumatology (EULAR), position paper, November 2011 (H2020 Framework Action: (eular.org)  3 Wang Y, Wei J, Zhang W, Doherty M, Zhang Y, Xie H, Li W, Wang N, Lei G, Zeng C. Gut dysbiosis in rheumatic diseases: A systematic review and meta-analysis of 92 observational studies. EBioMedicine. 2022 Jun;80: Epub 2022 May 17.  4 Parantainen J, Barreto G, Koivuniemi R, Kautiainen H, Nordström D, Moilanen E, Hämäläinen M, Leirisalo-Repo M, Nurmi K, Eklund KK. The biological activity of serum bacterial lipopolysaccharides associates with disease activity and likelihood of achieving remission in patients with rheumatoid arthritis. Arthritis Res Ther. 2022 Nov 21;24(1):256. | | |
| **9. Rationale for the planned study and research questions and / or hypotheses (up to 1800 characters, 1 page)** | | |
| In this project, we will explore the role of chronic systemic inflammation caused by intestinal microbiota derived immunologically active compounds, as a driver in the transition from health to disease with a special focus on three rheumatic diseases (RDs); osteoarthritis (OA), rheumatoid arthritis (RA), and spondylarthritis (SpA).  ***RDs affect more than 5 % of Europe's population and cause significant disability, pain, and reduced lifespan.3 As per EULAR, in Europe, rheumatic and musculoskeletal conditions represent an economic burden of estimated 240B€/year.4*** Dysbiosis of intestinal microbiota i.e. deleterious alterations of the composition and/or function of the gut microbiota has been implicated in the pathogenesis of a number of diseases. ***However, the mechanisms through which intestinal dysbiosis contributes to disease pathogenesis are largely unknown.*** Leakage of harmful compounds from intestinal bacteria in particular immunologically highly active bacterial wall lipopolysaccharides (LPS) is a significant source of systemic inflammation posing a high proinflammatory burden on the body. High systemic levels of LPS are referred to as “***systemic endotoxemia*** (***SE)***”. It is promoted by intestinal dysbiosis, compromised intestinal barrier function and lifestyle factors. Increased gut permeability facilitates the translocation of gut bacteria and specific bacterial structures across the gut-blood-barrier to reach systemic circulation.  **The more general idea of ENDOTARGET project** is to use the extensive data collected fromlarge cohort studies on the composition of intestinal microbiota, degree of intestinal permeability, SE, genetic risk profiles, and biomarkers from mechanistic studies involving the target tissue; and from interventional studies to validate predictive artificial intelligence (AI) algorithms in order to understand the role of microbiota in intestinal permeability and SE, and in the risk of developing the target RDs. By combining all this data, machine learning **(**ML) and AI-informed **rheumatic disease prediction tool (RDPT)** will be developed for clinicians to help them identify patients with increased risk of developing the target diseases. *It will thus assist in the choice of personalized blueprint intervention to reduce the risk of these diseases and disease activity in RA and SpA and to slow down the progression of OA.* Using similar approaches we will develop recommendations, which will enable people to assess their own personal disease risk and provide them with the means to reduce their personal risk by modification of diet and/or other lifestyle factors.  **The role of EstBB in this project is to participate in** (1) Population cohort analysis to explore biomarkers & lifestyle factors influencing health to disease transition by running whole genome association analyses for rheumatic diseases using EstBB cohort; (2) Developing Novel Strategies to prevent and control arthritis by interfering with the gut-joint axis (3) Data management, multiomics modelling and AI by perform unsupervised and supervised analysis combining standardized and harmonised clinical, genomic, metabolomic and microbiome data.  \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_  3 https://cordis.europa.eu/article/id/97231-ep-calls-to-recognise-the-extraordinary-burden-of-rheumatism-and-arthritis  4 European Alliance Of Associations For Rheumatology (EULAR), position paper, November 2011 (H2020 Framework Action: (eular.org) | | |
| **10. Research methodology (up to 1800 characters, 1 page)** | | |
| Toidentify drivers of chronic inflammation that determine the transition from health to pre-symptomatic and early stages of RDs, we will explore systemic endotoxemia in population level together with host genetic and gut microbiome analysis. ENDOTARGET will combine different molecular phenotypes to model the arthritis. For that we have defined three different cohorts we would use in ENDOTARGET study.   1. **Whole Estonian Biobank cohort (n>200,000 biobank participants, BPs)**   We will identify individuals from the whole Estonian Biobank cohort (n>200,000) who have been diagnosed with RA, OA or SpA or individuals with no such diagnosis as healthy controls (Please see the detailed inclusion and exclusion criteria from Appendix 1). Using the EstBB whole cohort, we will generate the risk prediction models for RDs and calculate the polygenic risk scores for RA, OA and SpA. For that the genotype and phenotype data already available in EstBB will be applied.   1. **Cohort6000 (n=6,000 BPs)**   From the whole EstBB cohort, we will choose up to 6000 BPs with or without the arthritis diagnosis for the further endotoxemia biomarker analyses (Appendix 2.). We will choose these 6000 BPs based on the availability of the plasma samples in EstBB and the criteria listed in Appendix 1. We will include BPs into Cohort6000 based on three additional criteria: (1) BPs with high or low risk for RDs (considering the risk scores calculated on whole Estonian Biobank cohort), but with no such diagnosis or (2) have the diagnosis during recruitment into EstBB (prevalent cases) or (3) BPs, who have received the RDs diagnosis after joining the EstBB (incident cases). For that the genotype and phenotype data and plasma samples already available in EstBB will be applied.   1. **Cohort1000 or Call-Back Cohort (n=1,000 BPs)**   This cohort will be chosen based on the endotoxemia biomarker analyses (Appendix 2.) results from the Cohort6000. The call-back cohort will be chosen amongst the Whole EstBB Cohort (n>200,000), the BPs will be with or without the arthritis diagnosis. We would call-back up to 1000 BPs to participate in ENDOTARGET study to collect additional samples (venous blood, feces and saliva samples) and data (Please see Appendix 6 for the study questionnaire). Blood samples will be used for additional analyses (Appendix 2) and feces and saliva samples for gut and oral metagenome (microbiome) analysis.  With the data and samples from these three cohorts, we plan to conduct the following analyses:   1. **Genome-wide association study (GWAS) between molecular traits and clinical outcomes** **using regression-based models.** For GWAS genome-wide genetic information is needed along with metabolomics data and clinical data. We already have all this information for over 200 000 BPs in EstBB, thus we do not need to collect additional data or samples for these analyses. We would include all the OA, RA and SpA (currently ~70 500 BPs) cases from EstBB cohort and the rest will be applied as controls to achieve the maximum power for the calculations (Appendix 1). 2. **Analysis of endotoxemia markers from blood plasma.** Blood markers will be measured from blood plasma (List of markers in Appendix 2). We will use blood plasma from BPs chosen into Cohort6000 and Cohort1000. We will analyse the blood plasma collected during the recruitment into EstBB and/or during the recruitment into EstMB cohort\* and/or during recruitment into ENDOTARGET study Cohort1000. Thus, for some BPs we might analyse plasma samples taken from different timepoints. 3. **Analysis of rheumatic diseases markers from blood samples.** RD markers will be measured from blood samples (List of markers in Appendix 2) from BPs chosen into Cohort1000. Blood samples are collected during recruitment into ENDOTARGET study. 4. **Risk prediction models using genomic, metabolomic, lifestyle and other predictive factors for SE.**We will use EstBB dataset and GWAS meta-analysis summary statistics from the leave-one-out meta-analysis to generate genetic risk scores (GRS) for the outcomes of interest. For this, we will use the GRS pipeline in place at the EstBB, which incorporates PRScs software, a Python-based tool that infers SNP effect sizes using GWAS summary statistics and an external LD reference panel. We will test the generated GRSs for associations with outcomes of interest and will measure the prediction accuracy using the *R2* parameter. We will then move on to test the joint effect of best-performing genetic, metabolomic and other predictive factors. PRS will be calculated using a variety of methods (such as prscs, ldpred2, megaprs, bayesR or prsice2), to best accommodate different possible genetic architectures. We will also consider different base inputs (gwases from pan ukb) or finngen project for selecting weights of GRSs. We will make use of available pipelines and scripts from genopred to aid analysis of the PRS validation and predictive ability. 5. **Gut and Oral Metagenome data analysis.** Gut and oral microbiome composition and functionality will be characterized using shotgun metagenomic sequencing (>3 Gb data per sample). We will use community clusters as a tool to assess disease risks and identify subphenotypes. All models will be additionally adjusted for other relevant covariates (BMI, age, gender, stool quality and history of antibiotic usage). The metagenomics analyses will be conducted using bacterial DNA extracted from the stool and saliva samples collected from Cohort1000 BPs. Additionally, we would use the existing metagenomics data gained from the EstMB\* cohort study.   \* EstMB This population-based cohort that was established in 2017 with the aim of enriching Estonian Biobank with microbiome data (The Estonian Committee on Bioethics and Human Research license No 266/T-10 (16.01.2027)). The data collected during the MB project have been merged with the EstBB´s database. Stool, oral, and blood samples were collected from 2509 EstBB participants (1764 females and 745 males), aged 23-89 years. The gut metagenome data was analysed with paired-end metagenomic shotgun sequencing using Illumina, Novaseq 6000 platform (on average 4.62 ± 0.44 Gb of data per sample). All subjects from EstMB cohort have rich clinical data (ICD-10 diagnosis and drug prescriptions from electronic health records, EHRs) and lifestyle data (dietary, mental health etc.) available from special questionnaires. | | |
| **11. Study sample and description of recruitment method. Information and consent forms, questionnaires and tests should be submitted as annexes to the application.** | | |
| **Sample size, inclusion of control groups** | In this study we have three different study cohorts: (1) Whole EstBB Cohort (n>200,000 BPs), (2) Cohort6000 (n=6,000 BPs), (3) Cohort1000 (n=1,000 BPs). In each cohort, there will be four study groups: BPs with RA, OA, SA and healthy controls (Appendix 1).  In case of Whole EstBB Cohort and Cohort6000 we will use plasma and genotype data already available in EstBB.  For Cohort1000 (Call-Back Cohort), we will invite up to 1000 BPs to participate in the ENDOTARGET study and collect new data (blood, fecal and saliva samples). As currently the response rate to study invitations has decreased to ~20 %, we would need to send invitations to up to 5000 BPs, to collect samples from 1000 BPs. The invitations will be sent out until we have recruited the 1000 BPs into ENDOTARGET study. For Cohort1000 we will use genotype data already available in EstBB. | |
| **Who recruits and how/where/by whom is informed consent obtained? (if applicable)** | **Cohort1000**  No new BPs will be recruited for current study; all the study subjects will be chosen from the current EstBB cohort.  Prior to participating in the study, the BPs must voluntarily sign the ENDOTARGET study consent form (Appendix 5), which will be done online. The exact guidelines will be sent to the BPs within the study invitation (Appendix 4). | |
| **How and from whom are the subjects selected (sampling frame)? What are inclusion or exclusion criteria of subjects?** | **Cohort1000**  Sampling frame is all BPs (as of the 11.06.2023 211,728 BPs).We will apply the International Statistical Classification of Diseases and Related Health Problems 10 (ICD-10) codes to decide which BPs are suitable for ENDOTARGET study. Please see the Appendix 1 for the detailed inclusion and exclusion criteria applied in the study. | |
| **Type of interventions (physical, mental or data, including special categories of personal data)** | **Cohort1000**  The following samples and data will be collected:   * Stool samples * Saliva samples * Blood samples (fasting blood) * Study questionnaire (Appendix 6) | |
| **Burden on the subject (methods of contact, number of visits, type and number of procedures, repetition of invitations, etc.)** | **Cohort1000**  The BPs will be invited to participate in the study via e-mail (Appendix 4). If we receive no answer, we send out another invitation in two weeks. We will consider the answer to our invitation as “no”, if we are not able to reach the BP via two e-mails.  If the BP is willing to participate in the study, he/she needs to sign the informed consent form and answer the study questionnaire online. Answering the questionnaire might take approximately between 16 minutes (for healthy controls) to 45 minutes (for OA patients with left and right hip and knee joints affected).  After that a feces sample collection kit will be sent to BP via post. After BP has collected the feces sample, he/she can go to the preferred blood collection centre to give the blood and saliva samples and leave the feces sample there. The blood samples in 2x 6 ml EDTA tubes and feces and saliva samples will be sent to GI by the blood collection centre. The blood samples in 2 x serum tubes and 1 x EDTA tube (altogether 10 ml) will be analysed on site in the medical laboratory of the service provider and will be discarded immediately after the analysis.  As we would need a fasting sample for our study, the BP should consume no food and drink nothing except pure water at least 9 hours before giving the blood samples. It would be good, if the blood samples are given between 8.00-12.00 as also the circadian cycle affect the blood composition. Altogether up to 22 ml of venous blood will be collected possibly with a single needle stick by professional staff at the blood collection centre . | |
| **12. Issuing of tissue samples to third parties (RNA, DNA, plasma etc)** | | |
| **The number of gene donors whose tissue samples will be issued and the types of tissue samples to be issued** | 1. **Endotoxemia markers from blood plasma:**   For measuring endotoxemia markers from blood: The blood plasma samples collected at different time points from the BPs will be analysed: (1) samples collected during the recruitment into the EstBB and/or (2) during the recruitment into the EstMB cohort (3) during recruitment into the ENDOTARGET study Cohort1000. We will use blood plasma from BPs chosen into Cohort6000 and Cohort1000. Plasma from venous blood has been extracted and stored in the Core Facility of Genomics at the University of Tartu and the blood marker analyses will be conducted in HUS (Finland).  The plasma samples will be sent to HUS two times:   1. We will send the blood plasma samples of up to 6000 BPs (all Cohort6000 BPs) to HUS for endotoxemia marker analysis. We will use the left-over plasmas from other projects (have gone through some freeze-thaw cycles) from the first ~50,000 BPs and plasmas from EstMB project. 2. Additionally, we will send blood plasma samples of up to 1000 BPs (all Cohort 1000 BPs) to HUS for endotoxemia marker analysis. We will send the plasma samples collected during ENDOTARGET project. 3. **RD markers from blood:**   From Cohort1000 we’ll collect additional blood samples (2 x serum tube and 1 x EDTA tube, altogether 10 ml of blood) for the additional blood marker analyses, which will be conducted in by medical service provider (Appendix 2)   1. **Metagenomics:**   Bacterial DNA will be extracted from the Cohort1000 (1000 BPs) stool and saliva samples in the Core Facility of Genomics at the University of Tartu. The metagenomics sequencing will be conducted in EU; however the final service provider will be decided after the samples have been collected.  In order to store and study tissue samples outside the territory of the Republic of Estonia, we request the appropriate permission from the Senate of the University of Tartu.  For sending out the plasma samples for Cohort6000 we have already received the approval from the Tartu University Senate (01.03.2024 nr 5 ). Additional approval will be applied for sending out the Cohort1000 samples. | |
| **The amount of tissue sample to be issued per one gene donor** | 1. Endotoxemia markers from blood plasma: Up to 500 ul (or as much as possible, if the available sample amount is smaller) of blood plasma per timepoint per BP will be sent to HUS (Finland). 2. RD markers from blood: 2 x serum tubes and 1 x EDTA tube (altogether up to 10 ml of venous blood) will be collected in medical service provider and analysed there immediately. These samples will not be sent to GI and the leftovers will be discarded on site after the analysis. 3. Metagenomics: Up to 2 ug of DNA per sample type (stool and saliva samples) will be sent for sequencing to service provider in EU. | |
| **The entity to whom tissue samples will be issued (country, institution, address)?** | 1. **Blood plasma samples for endotoxemia marker analysis:**   Helsinki University Hospital, Department of Rheumatology, Helsinki, Finland (HUS)  Contact persons:  Goncalo Baretto  [goncalo.barreto@hus.fi](mailto:goncalo.barreto@hus.fi)  +358458538110  Biomedicum Helsinki 1  Haartmaninkatu 8, room C403b  00290 Helsinki, Finland  Sirkku Peltonen  [sirkku.peltonen@hus.fi](mailto:sirkku.peltonen@hus.fi)  Biomedicum Helsinki 1  Haartmaninkatu 8  00290 Helsinki, Finland   1. **Blood serum and EDTA samples:**   The medical service provider will be determined through the procurement procedure.   1. **DNA from stool and saliva samples:**   The metagenomics sequencing service provider in EU will be decided after sample collection. | |
| **What will be done with the residue samples (will the residue samples be destroyed or sent back to Gene Bank)?** | The residue samples in HUS will be sent back to EstBB after conducting the analyses.  At the EstBB the remaining stool and saliva (and DNA extracted from it) and blood plasma samples will be stored indefinitely as a part of the EstBB. Long-term preservation is important as it is a valuable biological material that can be used in future research. The relevant information is provided in the consent form.  The residue samples (DNA) in metagenomics sequencing service provider in EU will be destroyed on site according to the local laboratory protocols after up to 6 months of conducting the analyses.  The blood samples collected for analyses in medical laboratory service provider will be destroyed immediately after finishing the analyses. | |
| **13. Analysis of the ethical aspects of the study (3600 characters, up to 2 pages).**  **All research involving human subjects must be carried out in compliance with ethical requirements, in particular the principles of respect for autonomy, charity and the prevention of harm, and justice.** (<https://www.coe.int/en/web/bioethics/guide-for-research-ethics-committees-members>).  The project is carried out on the basis of the Human Genes Research Act (HGRA https://www.riigiteataja.ee/en/eli/ee/Riigikogu/act/508042019001/consolide) and other legislation concerning the work of the EBB. The principles of bioethics have been taken into account when planning activities.  **1. The principle of respect for individual autonomy**  The main features of the first principle are the voluntary nature of participation in the trial, the existence of written consent and the protection of privacy and personal data. This study will include data from volunteer biobank participants (BP) of Estonian Biobank . BP have signed a consent form for joining the Estonian Biobank, allowing to use their data for different research purposes. The consent form to become a gene donor and the following execution and maintenance procedures are set out by the Minister of Health and Labour (https://www.riigiteataja.ee/akt/117042019016). According to HGRA § 12 by signing the consent a gene donor agrees to provide a tissue sample, to have a description of the state of health or the genealogy of the person prepared, to entry of the description of the state of health or the genealogy in the Estonian Biobank in pseudonymised form and to use thereof for genetic research, assessment of personal health risks and prevention of diseases, public health research and statistical purposes. Scientists analysing data will use only pseudonymised data and no data re-encoding by partners will be permitted. No genomics data of BPs will be sent abroad and will be analysed by scientists of the Institute of Genomics. Therefore, the threat to the gene donors’ privacy is low.  **2. The principle of not doing harm**  The present study has a research design developed by competent scientists and has the latest technology for performing analyzes. Implementing high-level technology ensures the accuracy and high quality of the research.  **3. The principle of charity**  From this study EstBB will provide the gene donor no direct charity.  Social charity is manifested through scientific innovation and the improvement of the health system. In the long term, the application of research results will be towards more personalized treatment, prevention and diagnosis.    **4. The principle of justice**  This research creates a prerequisite for more effective, safer and better access to healthcare in society. Estonian Biobank is a population-based biobank, where different social and age groups are represented, the distribution of gender and other demographic features in the Estonian Biobank quite well reflect the composition of the entire Estonian population. Thus, in the long run, all members of society benefit from research.  Remarks:  According to the Human Gene Research Act (HGRA) EstBB all tissue samples shall be preserved in the territory of the Republic of Estonia. The University of Tartu senate may, if good reasons therefor become evident, grant permission for tissue samples to be preserved and studied outside the territory of the Republic of Estonia if the tissue sample is issued in unpersonalised form and the controller ensures effective control over the tissue samples and that the tissue samples cannot be used in a manner prohibited by legislation.  The samples will be issued from the EstBB in accordance with the requirements established in sections of HGRA § 20 (1) and § 22 (4) respectively in pseudonymised form and as a set of data, which means they cannot be used as a basis for characterizing a particular gene donor and that they can only be used for scientific research.  An agreement will be concluded with the institution receiving the tissue samples pursuant to §18 (3) of the HGRA. The tissue samples must be issued in accordance with a regulation of the minister responsible for the area “Procedure for Issuing Tissue Samples, DNA Description and Health Description of Gene Donors” (https://www.riigiteataja.ee/akt/86997?leiaKehtiv).  **Please see also** <https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/ethics/h2020_hi_ethics-self-assess_en.pdf> | | |
| **13 a Human subjects** | | |
| **Assistance questions** | **No** | **Yes** |
| **Are people the object of research?** |  | **If applicable, describe how voluntary participation in the study is ensured and any undue influence on study participants to participate in the study is avoided.**  Yes, we will not recruit new biobank participants with this study. The participation in the Estonian Biobank is strictly voluntary and based on signing an informed consent form. The participant rights (incl. obtaining consent and consent withdrawal) are regulated by the Human Genes Research Act. The participants are aware of the use of their data for research purposes based on HGRA §16 and the informed consent. The participants may withdraw their consent of participation and the data of these individuals will not be used in future studies.  To participate in the ENDOTARGET study, gene donors invited into Cohort1000 voluntarily sign a separate study consent form. There is no fee offered to participate in the study to avoid bias. Participants have the right to withdraw from the study at any time. |
| **Are the study participants vulnerable individuals or groups?** | **No** | 1. **Indicate to which vulnerable group the study participants belong and what their vulnerability lies in (details of the type of vulnerability).** 2. **In case the inclusion in the study is based on informed consent, please describe the procedure for obtaining informed consent and include a consent form. These activities must ensure that individuals understand the risks involved in participating in the study.** |
| **Does the study include persons who cannot themselves give informed consent to participate in the research (incl. persons with limited active legal capacity)?** | **No** | **Describe how the guardian or legal representative is informed and how the consent from them is obtained for partipation of the individuals with limited active legal capacity, including children/minors.** |
| **Are the study participants children/minors?** | **No** | 1. **Please provide age range.** 2. **Describe the procedure of obtaining consent from the parents of the child.** 3. **Describe how the consent of the minor is sought or the opinion of the minor is taken into account in proportion to his or her age and degree of maturity.** 4. **Describe how the well-being of minors is ensured.** 5. **Explain the reason for including minors in the study.** |
| **Are the study participants patients?** |  | 1. **What disease/condition/disability do they have?**   The BPs have been diagnosed with OA, RA or SpA or have no such diagnoses (healthy controls). There is no doctor-patient relationship in this study.   1. **Details of the recruitment, inclusion and exclusion criteria and informed consent procedures.**   **Cohort1000**  During the recruitment into the ENDOTARGET study the BPs are not patients; however, it is possible that the EstBB database contains data, which is collected at the time the BPs were patients. The BPs with arthritis diagnosis are patients; however, the data and sample collection does not take place at their doctor’s office but at their home (self-collecting the feces samples) or at preferred blood collection centre. Thus, the doctor’s- patient’s relationship is not inclining the BP to participate in the ENDOTARGET study. The BPs are informed that their participating or not participating in ENDOTARGET study does not affect the quality of their healthcare services. Detailed informed consent form (Appendix 5), notifying the subject about his or her rights, potential risks and benefits of participation is provided together with the invitation to the study. Invitation is sent via e-mail (Appendix 4). The signed informed consent is obtained from all subjects prior to inclusion into the ENDOTARGET study.  The inclusion and exclusion criteria of the study participants are listed in Appendix 1.   1. **What is your policy on incidental findings? How the participants will be informed about incidental findings?**   Within this project, no additional genotyping nor sequencing targeting human DNA will be done and therefore there cannot be any incidental findings from genetic data. |
| **Does the research involve collection of biological samples? Are human biological samples intended for export to a third country (**[**https://www.aki.ee/et/teenused-poordumisvormid/andmete-edastamine-valisriiki**](https://www.aki.ee/et/teenused-poordumisvormid/andmete-edastamine-valisriiki)**)**  **(**[**https://www.aki.ee/en/guidelines/transfer-personal-data-foreign-country-0**](https://www.aki.ee/en/guidelines/transfer-personal-data-foreign-country-0)**) or import them from another country to Estonia?** | **Yes**  **Yes** | 1. **What type of biological samples will be collected?**   Stool, saliva and venous blood samples.   1. **What are your procedures for collecting biological samples, including use of previously collected samples?**   **Cohort6000**  The previously collected blood plasma samples will be aliquoted and transported to HUS. GDPR requirements apply in Finland.  **Cohort1000**  The BPs will collect the stool samples at home prior to the appointment at blood collection centre. For that they will be sent a sample collection kit via post, which contains all the necessary for the samples collection and detailed instructions (Appendix 7).  Drawing venous blood samples applying vacuum tubes at blood collection centres by experienced specialist.  Saliva sample (~2 ml of saliva) will be collected at blood collection centre after collecting blood samples. For that, a saliva sample collection tube (sterile 15 ml falcon tube) will be provided for the study participant and the collection method will be explained on site. The saliva sample in tube will be given to blood collection centre specialist.   1. **Explain how the rights of the subjects are ensured.**   The participation in the Estonian Biobank is strictly voluntary and based on signing an informed consent form. The participant rights (incl. obtaining consent and consent withdrawal) are regulated by the Human Genes Research Act. The participants are aware of the use of their data for research purposes based on HGRA §16 and the informed consent. The participants may withdraw their consent of participation and the data of these individuals will not be used in future studies.  **Cohort1000**  Prior to agreeing to participate in the ENDOTARGET study and signing the study consent the BP will be informed about the burden and potential risks of the study in the informed consent form online. If they have any additional questions regarding participating in the study, they can find the contact to turn to from the consent form and ask for additional information before deciding, if they want to participate in the study. In case of positive decision to participate in the study, the BP needs to voluntarily sign the IC. The BP has the right to withdraw the consent to participate by sending a simple notification without providing a reason. The procedures for termination of consent are provided in the consent form.   1. **Explain what will be done with the biological samples after the end of the study.**   The remaining plasma, stool and saliva samples (incl. DNA extracted from it) collected during the study and not sent out during the ENDOTARGET project for further analysis will be stored in EstBB indefinitely.  The leftover plasma samples in HUS will be sent back to EstBB after conducting the analyses. The remaining plasma samples will be stored in EstBB indefinitely.  The blood samples collected for analyses by medical service provider will be destroyed immediately after finishing the analyses.  The residue samples (DNA) in metagenomics sequencing service provider in EU will be destroyed on site according to the local laboratory protocols after up to 6 months of conducting the analyses.  Please see further explanation in point 12. |
| **13 b Personal data and datasets** | | |
|  | **No** | **Yes** |
| **Are personal data collected or analyzed in the study, including special categories of personal data?** |  | 1. **The full list of variables collected in the study (may be provided as an annex).**   Please see the Appendix 2, 3, 6, 7, 8.   1. **Confirm that informed consent exists or is obtained before the start of the study if the study is based on consent.**   Objects of research are participants of EstBB. All participants have signed the consent to become a gene donor before joining the EstBB.  **Cohort1000**  All BP-s, who agree to participate in the study will need to voluntarily sign the additional ENDOTARGET study consent form.   1. **Explain why all data processed are relevant and necessary (based on the principle of data minimization).**   The ENDOTARGET study applies the following data:   * genotype data to evaluate the genetic associations with different phenotypic and clinical traits; * blood markers to evaluate associations with genetic and clinical data; * health data from EstBB database to evaluate associations with genetics, other phenotypic traits and clinical data.   **Additional points for Cohort1000**   * personal data is needed to contact BPs and invite them to the study; * metagenomics data from stool and saliva samples to evaluate associations with microbiome, phenotypic traits and clinical data; * study questionnaire to evaluate associations with genetics, other phenotypic traits and clinical data;  1. **Are the data subjects identifiable?**  * All the data and tissue samples in this study is pseudonymised prior to sending to the EstBB scientists or partners. * For study scientists the data will be pseudonymised so that there is no ability to link data to individuals.   **Additional points for Cohort1000**   * The tissue samples are marked with codes. The key to the codes is kept in the EstBB coding center indefinitely and the remaining tissues will be stored as a part of EstBB. * In blood collection centres the BPs need to be identified prior to collecting their blood and saliva samples and accepting their feces samples (collected at home). Medical service provider is a licensed blood collection and analysis institute, which also collaborates with different hospitals in Estonia. It has licensed and trained specialists to work with customers personal data and samples.   **If yes, please describe how the following conditions are met:**   * 1. **after the removal of the personal identifiers, the purposes of data processing are no longer achievable or would be unreasonably difficult to achieve;??**   2. **in the opinion of the persons conducting scientific or official statistics, there is an overriding public interest therein;**   3. **the scope of obligations of the data subject is not changed based on the processed personal data or the rights of the data subject are not excessively damaged in any other manner.** |
| **Does the research involve systematic monitoring of an individual, the collection of his or her data profile, or a large-scale processing of data of special categories and /or sensitive data, or the use of (intrusive) data processing techniques in a covert way (eg survival surveys, monitoring, surveillance, audio and video recording, geolocation, etc.) or any data processing that may harm the rights and freedoms of the data subject?** |  | 1. **Explain which methods are used for surveillance, monitoring and observation of subjects.**   **Cohort1000**  After signing the IC the BPs will be asked to fill the study questionnaire online (Appendix 6). After that they will be sent a feces sample collection kit via post to their preferred location. After collecting the feces samples, they can go to preferred blood collection centre, have the blood samples collected and collect the saliva sample on site and give collected feces samples to the specialist in the centre.   1. **Explain the methods of creating the profile of the subjects.**   The BP profile will be put together from the following data sources:   * genotype data from EstBB * phenotype data from EstBB (Appendix 1, 3) * data from questionnaires including anthropometric measurements (Appendix 3) * blood plasma analyses (endotoxemia markers, Appendix 2)   Cohort1000   * Study questionnaire (Appendix 6) * Microbiome data (metagenomics data) from stool and saliva samples (Appendix 7) * Additional blood markers (Appendix 2)  1. **Explain how subjects are informed about their rights and the potential risks that data processing may entail.**   **Cohort1000**   * The BPs are fully informed of all aspects of the study during invitation (detailed informed consent form notifying the subject about his or her rights, potential risks and benefits of participation is linked to the e-mail together with the invitation). * If the BPs have additional questions prior to signing the informed consent, answering the study questionnaire, collecting the feces sample or going to the blood collection centre, they can turn to EstBB via e-mail or telephone given in the invitation and informed consent form and guidelines for collecting stool samples.  1. **Explain how data will be collected for the profile of the subjects and how the study participants will be informed about possible consequences and safeguards.**   Whole EstBB Cohort, Cohort6000 and Cohort1000   * genotype data is taken from EstBB, and has been created before the ENDOTARGET study; * phenotype data from EstBB has been collected before the ENDOTARGET study; * study questionnaire including anthropometric measurements will be filled only by Cohort1000 and online. * endotoxemia markers from blood (Cohort6000 and Cohort1000), will be measured from the blood samples collected at blood collecting centres or stored in EstBB. * RD markers from blood (Cohort1000), will be measured from the blood samples collected at blood collecting centres. * Metagenomics sequencing targeting bacterial DNA will be conducted from DNA extracted from feces and saliva samples for Cohort1000 only (feces samples are collected at home, and saliva samples at the blood collection center, transferred to GI and DNA is extracted in GI).   The BPs in Cohort1000 will be informed about possible consequences and safeguards during invitation and in informed consent form, prior to signing the consent form. |
| **Is there a plan to analyze previously collected personal data?** | Yes | 1. **Explain from which database (register) or source the data are obtained.**   Yes, previously collected data about BP already collected into EstBB servers will be used.   1. **Explain how subjects are informed about their rights and the potential risks that data processing may entail.**   **Cohort1000**  The BP will be informed about possible consequences and safeguards during the invitation and via phone or e-mail if they have additional questions.   1. **Explain why all data processed are relevant and necessary (based on the principle of data minimization).**   The ENDOTARGET study applies the following data obtained from EstBB:   * genotype data to evaluate the genetic associations with different phenotypic and clinical traits. * phenotype data to evaluate associations with genetic and clinical data (Appendix 3).   **Additional points for Cohort1000**   * personal data (name, ID number, e-mail address, home address) is needed to contact BPs and invite them to the study and send them the stool samples collection kit prior to the appointment, ID number is necessary to ensure person's identity, which will be checked in the process of signing consent form online; also the ID is needed to confirm the identity at the blood collection centre.  1. **Explain why it is not possible to study the participants in such a way that the data obtained were anonymous or pseudonymous (if applicable).**   **Cohort1000**  We need to depseudonymize the BPs as we need to invite them to the ENDOTARGET study for collecting additional samples (blood and saliva at blood collection center and stool sample collected at home) and fill the study questionnaire. Prior to making the data accessible for the study scientists it will be again pseudonymized. |
| **Is there a plan to analyze publicly available data?** | No | **Explain that the data are publicly available (open data registers and databases) and can be used freely in research.** |
| **Is there an intention to transfer personal data or provide access to personal data to third countries**  **(**[**https://www.aki.ee/et/teenused-poordumisvormid/andmete-edastamine-valisriiki**](https://www.aki.ee/et/teenused-poordumisvormid/andmete-edastamine-valisriiki)**)?**  **(https://www.aki.ee/en/guidelines/transfer-personal-data-foreign-country-0)** | No | 1. **Explain what personal data are exported or imported. If so, to which and from which countries.**   **Personal data (name, contacts etc.) will not be transferred or made accessible to third countries.** HUS will analyse the blood plasma samples, thus they will generate the analysis results, which will be pseudonymized data. They will forward the analyses results to EstBB, thus will not receive any other individual level data.  Medical service provider will collect and transfer and analyse the blood samples and transfer the feces and saliva samples to GI. For that they need to identify the BP prior to collecting the samples. BPs personal information (name, ID code) together with information about the needed samples to be collected and analyses to be conducted will be sent to blood collection center after the BP has signed the informed consent of the ENDOTARGET study and answered the questionnaire (both done online).  Service provider in EU will analyse the DNA extracted from stool and saliva samples targeting the bacterial DNA. They will generate the analysis results, which will be pseudonymized data. They will forward the analyses results to EstBB, thus will not receive any other individual level data.  Genotype data will not be made accessible for partners. It will be made available for EstBB researchers via a server of the University of Tartu.  Additionally, summary statistics gained from the analyses of individual level data will be shared with the ENDOTARGET study partners. The data will be deleted up to 10 years after receiving the data.   1. **Explain what protection measures are applied, on what grounds (contract, etc.) the data are transferred and how the rights of the subjects are ensured.** |
| **Will personal data be destroyed / anonymised at the end of the research?** |  | **In case the analysis uses data in a form which enables identification the study participants, please**  **1) describe how personal data will be destroyed / anonymised after the research has been finished and the objectives have been achieved;**  In blood collection center, the personal data (list of BPs names and ID codes and analysis results) will be destroyed after conducting the analyses and transferring all samples and data to GI (not later than 31.12.2025).  **2) add an assessment of how the possibility of indirect identification of data subjects is managed after the destruction of data enabling direct identification of a person.**  The blood plasma samples sent to HUS will be sent back after conducting the analyses (if there are any leftovers). The analysis results data (at individual level) will be sent from HUS to GI via secure servers and will be destroyed at HUS servers not later that at the end of the ENDOTARGET study (31.12.2026).  The blood samples collected and analysed in medical service provider will be destroyed immediately after finishing the analyses. The analysis results data (at individual level) will be sent to GI via secure servers and will be destroyed at after conducting the analysis and transferring all samples and data to GI (not later than 31.12.2025).  The leftovers of the DNA samples sent to partner in EU for metagenomics sequencing, will be destroyed up to 6 months after conducting the analyses. The analysis results data (at individual level) will be sent to GI via secure servers and will be destroyed at service provider servers not later that at the end of the ENDOTARGET study (31.12.2026).  The data obtained during the project and biological samples at EstBB will be added to the EstBB database and stored indefinitely. |
| **13 c Other ethical issues** | | |
| **Can conducting research involve ethical risks not described above?** | No | **Explain, if necessary, the additional ethical risks that may arise from factors such as artificial intelligence, personal medicine, involvement of military partners, new developments in neurobiology, genetic engineering, nanotechnology, human-machine interaction, android and cyborg creation, etc.** |
| **14. Complete in case the research is based on data from a database and/or register** | | |
| **Name of database and/or register** | EstBB | |
| **Purpose of the processing of personal data** | Processing of personal data is necessary to conduct a study and administrate the recalling of the EstBB donors. | |
| **List of variables and period for which data are collected (in annex if necessary)**  The data is collected between 01.01.2024 to 31.12.2030. The collected data is listed in Appendices 1, 2, 3, 6 and 7. | | |
| **15. Description of personal data protection measures, including data storage, security and erasure, including date of erasure of data and / or code key (up to 1800 characters, 1 page).** | | |
| **Describe and justify the storage of data collected for the study and the deadline for storage.** | For information on biological samples collected please refer to point 12.  At the EstBB the remaining biological samples and other collected data such as results from the blood measurements will be stored indefinitely (integrated to the gene donor data stored at the EstBB). Researchers of EstBB listed on the application will have access to the collected data until the end of the study (31.12.2030) through the secure server of University of Tartu.  The residue plasma samples sent to partners will be sent back to EstBB, if there is any leftovers from the endotoxemia marker analyses.  **Additional points for Cohort1000**  The residue DNA extracted from stool and saliva samples sent to partner will be destroyed up to 6 months after conducting the analyses.  Blood collection centre, who collects the serum and EDTA samples for RD markers analyses will destroy the samples after finishing the analyses on site. The personal data of BPs (name, ID code) will be deleted from the service providers database after conducting the analysis and transferring all samples and data to GI (not later than 31.12.2025).  The study participants in Cohort1000 have been provided with the information in the informed consent to the purposes, uses and retention of the data collected. The consent forms are stored at the EstBB coding center indefinitely, to ensure that in the event of any inquiry from the gene donor, it is possible to establish his/her declaration of intent. | |
| **Describe the process and means of pseudonymisation of personal data.** | The data collected on gene donors has been pseudonymized during their accession to the EstBB. For the purpose of the study the subjects are pseudonymized through using a numbered study code, the key of the code is kept at the EstBB coding center. | |
| **Is there a plan to de-pseudonymise gene donors’ data?** | 1. **Please specify the number of gene donors whose data will be de-pseudonymised.**   Prerequisite for de-pseudonymisation is to obtain EBIN approval. Contact with a gene donor takes place based on consent to become a gene donor and the Human Gene Research Act. Further activities will take place with the additional consent of the gene donor. Institute of Genomics hereby requests permission from EBIN to de-pseudonymise data of BPs.  Up to 5000 BPs will receive the invitation (via e-mail) to the ENDOTARGET study, for that purpose they need to be de-pseudonymized.   1. **Please explain the reason for de-pseudonymisation.**   We need to de-pseudonymise the BP’s as we need to invite them into the ENDOTARGET study to fill the study questionnaire, donate blood, feces and saliva samples. Prior to making the data accessible for the study scientists it will be again pseudonymised. | |
| **Is there a plan to transport personal data? Please describe how data protection is ensured.** | No | |
| **Describe how the data are protected against unauthorized or unlawful processing.** | Data on human participants will be generated (sample analysis results) and processed for purposes relevant to accomplish the ENDOTARGET objectives. Only necessary data that are adequate and relevant will be generated and that will not include information that can be regarded as superfluous (‘data minimization principle’).  The collection, maintenance and release of the genetic and health data at the Estonian biobank are conducted in accordance with the Human Genes Research Act (HGRA, https://www.riigiteataja. ee/en/eli/508042019001/consolide), which ensures that the identity of participants remains confidential. It also specifies the conditions for sample processing, restrictions on the use of samples and conditions for research activities. All processes have been reviewed and updated in light of the GDPR. All Estonian Biobank participants, aged 18 and up, have signed an informed consent allowing broad genetic research, as well as all activities described in the Estonian Human Genes Research Act (HGRA). For EstBB researchers who perform data analysis data from participants will only be accessible in pseudonymised form. The data are stored securely on a secure server of University of Tartu with a back-up function.  ENDOTARGET partners will take adequate measures to ensure data protection and confidentiality. The latest local, national and international rules on data protection (in particular, Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data) will be followed and no information on the personal details of the study participant will be transferred. | |
| **I confirm that all researchers are aware of the ethical and personal data protection requirements of the project.** | | |
| **Signature of the principal investigator** | **Date of application: 06.08.2024** | |
| Digitally signed |  | |
| **EBIN ID of the application**  **(fills by the assessor)** | | |
|  | | |

**List of additional documents:**

**CV of the principal investigator**

The CV of the Principal Investigator, professor Reedik Mägi, can be found on the link below:

[**https://www.etis.ee/CV/Reedik\_M%C3%A4gi/est?tabId=CV\_ENG**](https://www.etis.ee/CV/Reedik_M%C3%A4gi/est?tabId=CV_ENG)

Appendix 1. Detailed inclusion and exclusion criteria

Appendix 2. List of additional markers measured from blood

Appendix 3. List of variables and samples

Appendix 4. Study Invitation (EST)

Appendix 5. Informed Consent (EST)

Appendix 6. Study Questionnaire

Appendix 7. Guidelines for collecting the stool samples and Bristol Scale