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)

ENG: The influence of risk factors on inpatient and outpatient care of acute myocardial infarction patients in Estonia

EST: Riskifaktorite mõju ägeda müokardiinfarktiga patsientide haiglaravile ja haiglajärgsele käsitlusele Eestis

2. The main purpose of the study (up to 450 characters / 0.25 pages) (If the application is submitted in English, please also provide the main purpose in Estonian.)

ENG: The aim of this analysis is to describe the monitoring of hospital care and post-hospital management of acute myocardial infarction (AMI) patients in Estonia, compare its compliance with Estonian treatment standards, and assess factors related to differences in approaches.

EST: Käesoleva analüüsi eesmärk on kirjeldada ägeda müokardiinfarkti (ÄMI) patsientide haiglaravi ja haiglajärgse käsitluse jälgimist Eestis, võrrelda selle vastavust Eesti ravistandarditele ning hinnata lähenemisviiside erinevusega seotud tegureid.

3. Study period (the beginning and end dates MM/YYYY)

01/2026 – 12/2030

4. Principal investigator(s) and their contact details

Given name(s): Pankaj

Last name: Chejara

Position: Data Scientist, Health Data Unit

Institution: Metrosert AS

Phone: +372 5860 1799

e-mail: pankaj.chejara@metrosert.ee

5. Other researchers involved in the study (add lines as necessary)

1. Given name(s): Javier

Last name: Fernández

Position: Data Scientist, Health Data Unit

Institution: Applied Research Centre, Metrosert AS

2. Given name(s): Anti

Last name: Karumaa

Position: Data Scientist, Health Data Unit

Institution: Applied Research Centre, Metrosert AS

3. Given name(s): Toomas

Last name: Marandi

Position: Cardiology Researcher

Institution: University of Tartu Institute of Clinical Medicine, Heart Clinic

6. Financing of the study

Sources of funding	The study is a research collaboration funded by the Ravimitootjate Liit.
Total cost of the study (amount)	16,740 €
Financial compensation for the study participants (yes, no, explanation and amount)	N/A
Insurance provided for the study participants (yes, no, name of the insurance company and the certificate of insurance (COI))	N/A

7. Information about previous or parallel evaluation of the same study project (incl in other countries)

N/A

8. Brief overview of previous studies on the same topic (up to 900 characters / 0.5 pages)

Studies of post-hospital care of patients with myocardial infarction (MI) have analyzed how well real-world care is consistent with guideline recommendations, such as the use of medications (beta-blockers, ACE inhibitors, statins and other lipid-lowering drugs, antithrombotic drugs), lifestyle counseling and timely follow-up, and participation in a rehabilitation program. The results show that adherence to treatment recommendations reduces the risk of mortality and complications. Factors that influence treatment adherence include, for example, the patient's age, comorbidities, socioeconomic status, access to treatment, physician specialty, hospital resources, and regional healthcare systems. Health literacy, the amount of drug costs and the proportion of patient co-payments, and patient involvement in the treatment process also play an important role. In Estonia, a consensus document on post-hospital care and counseling of patients with myocardial infarction has been prepared as a result of the joint work of cardiologists, family physicians, and rehabilitation physicians, and was published in the journal Eesti Arst (2022, 101(5):324–328). In addition, the treatment pathway for post-hospital treatment of myocardial infarction patients has been newly described within the framework of the Health Insurance Fund's treatment pathway accelerator program.

9. Rationale for the planned study and research questions and / or hypotheses (up to 1800 characters, 1 page)

Currently, there is insufficient knowledge in Estonia about the extent to which treatment recommendations for patients with myocardial infarction are followed, the extent to which risk factors are taken into account in the post-hospital follow-up, rehabilitation and counselling of patients with myocardial infarction, and the extent to which risk factors may influence the variation in patients' treatment adherence and their prognosis during the follow-up period. Previous studies on the adherence to these recommendations in Estonia have only been conducted based on medical bills but have not used clinical data generated during clinical work.

The aim of this study is to fill a gap in existing knowledge by providing a comprehensive description of the management of patients with myocardial infarction in the year 2023–2025, within one year of the index episode (hospitalization). Specifically, the study aims to answer key questions regarding the patient journey and how it relates to risk factors and treatment outcomes, such as:

- Were the prescribed treatment steps performed in a timely manner by the right specialists, and to what extent did this vary across subgroups?
- Were LDL-cholesterol and Lp(a) test results used to modify the treatment regimen (pre-MI, during hospital stay, and post-MI), including the use and dosage of lipid-lowering drugs?
- How did the subtype of MI (e.g., STEMI vs. NSTEMI) influence patient management?
- Echocardiography (LVEF value) and coronary angiography performed during hospital stay.
- Was the LVEF value during the hospital period used to refer patients for repeat testing and to determine whether to place an ICD (implantable cardioverter-defibrillator)?
- Use of revascularization procedures before MI and during hospital stay.
- Reasons for repeat hospital admissions.
- How did the use of medications (e.g., lipid-lowering drugs, beta-blockers, P2Y12 inhibitors, SGLT2 inhibitors, etc.) affect treatment outcomes?
- How did the cost of treatment differ across risk groups?
- Does the quality of treatment recommendations upon discharge from hospital correspond to that described in the consensus document?
- What was the referral to rehabilitation like in different Estonian hospitals, and did the type and length of the rehabilitation program affect the prognosis of patients?
- Could the quality indicators mentioned in the document on the treatment journey for patients with myocardial infarction after hospital treatment be helpful in improving patient outcomes and harmonizing treatment in different regions of Estonia?

The secondary objective of the study is to assess the quality of basic data in Estonian registers and identify data gaps and other challenges related to conducting analyses (e.g. obstacles to using different databases when conducting regular quality indicator queries, the need for coordination between different parties and the time taken for this, etc.).

10. Research methodology (up to 1800 characters, 1 page)

We will study the patients treatment journey and how it is affected by risk factors. The patient's journey begins when the patient has an inpatient medical bill with a primary diagnosis of I21.x or I22.x with a duration of at least two days (or several consecutive bills with a total duration of at least two days). The start date of the patient's journey (index date) is the date on which such an invoice (index invoice) was opened (in the case of several consecutive invoices, the minimum of these). Patients who die before the end of the in-hospital treatment following the index date will be excluded. The journey ends with the death of the patient or 365 days after the index date.

For each patient journey, we first characterise the patient with regard to the MI sub-type, age, biological sex, and the history of risk factors prior to the index event of the journey, such as MI associated diagnoses (including obesity, diabetes, hypertension, hypercholesteremia and chronic kidney diseases) and MI indicators based on LVEF, LDL cholesterol and Lp(a) measurements. We then quantify whether treatments in the post-hospital period (including appointments, clinical tests and medications prescribed) are in accordance with the treatment guidelines and outcomes such as additional MI events in the post-hospital. Treatment outcomes will be quantified by recording presence of subsequent MI events and/or death during the post-hospital period, as well as changes in MI risk indicators. The cost of the treatment journey will be calculated from the medical bills and prescriptions.

Statistical analyses will be carried out to summarize data and test for significant differences in outcomes and compliance with guidelines between risk groups.

The analysis of treatment costs is based on the standard price components (total price, reimbursed part, patient co-payment) included in the prescription data, which are used to assess the relationship between treatment adherence and treatment outcomes..

The data of the Health Information System are inquired only about those persons who meet the inclusion criteria (initial diagnosis I21.x or I22.x). The additional diagnostic codes listed in the data composition table are used only to describe comorbidities and the treatment pathway.

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	The total sample size will be specified based on the results of the inquiry, but it is known that the number of cases of AMI in Estonia is approximately 2,700 per year, so we estimate the total sample size of the study to be approximately 8,100 patients.
Who is responsible for recruitment? Where and how is informed consent obtained, and by whom? (if applicable)	N/A
How and from whom are the subjects selected (sampling frame)? What are inclusion or exclusion criteria of subjects?	Data will be retrieved from national registers alone; no patient will be contacted directly. Inclusion criteria are: having inpatient medical bill with a primary diagnosis of I21.x or I22.x with a duration of at least two days (or several consecutive bills with a total duration of at least two days) in the interval 01.01.2023 – 31.12.2025. Patients who die before the end of the in-hospital treatment following the index date will be excluded, as well as individuals who have opted out of data sharing for scientific purposes. The data of the Health Information System are inquired only about those persons who meet the inclusion criteria (initial diagnosis I21.x or I22.x). The additional diagnostic codes listed in the data composition table are used only to describe comorbidities and the treatment pathway.

Type of interventions (physical, mental or data, including special categories of personal data)	N/A
Participant burden (e.g. frequency of contact, number of visits or procedures, repeated invitations, etc.)	N/A
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	N/A
The amount of tissue samples to be issued per one gene donor	N/A
The entity to whom tissue samples will be issued (country, institution, address)?	N/A
What will be done with the residue samples (will the residue samples be destroyed or sent back to Biobank)?	N/A
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.** **Please see also** https://ec.europa.eu/info/funding-tenders/opportunities/docs/2021-2027/common/guidance/how-to-complete-your-ethics-self-assessment_en.pdf; https://etag.ee/en/activities/research-integrity/	

Nature and proportionality of the research.
The study uses pseudonymised individual-level data from four national registries: the Estonian Health Information System (Digilugu), EMIR, RETS and the Estonian Health Insurance Fund (EHIF) claims database. These are complemented by data from the Causes of Death Registry. The aim is to assess compliance with myocardial infarction (MI) treatment guidelines and to identify variations in outcomes between risk groups within one year after the acute event. The study serves a clear public health interest. No direct contact with participants takes place and no biological samples are collected. Data processing is limited to the minimum necessary variables for the defined research purpose in accordance with the data minimisation principle.
Legal basis and data subject rights.
Data are processed for scientific research purposes under paragraph 6(1) of the Estonian Personal Data Protection Act (IKS). Individuals who have opted out from the use of their data for research are excluded. The rights of data subjects, including access, objection, and information rights, are ensured through the procedures of the respective data controllers, namely TEHIK, EHIF and other registries. Individual notification is not proportionate or required for register-based research; The rights of data subjects are ensured through national notification procedures and an opt-out mechanism.
Roles and responsibilities.
TEHIK, in cooperation with the Estonian Health Insurance Fund, performs record linkage and pseudonymisation and acts as the data controller for the merged dataset. Analyses are conducted within the secure SAPU environment at the University of Tartu. Only aggregated and fully anonymised results are exported from the secure environment after verification by TEHIK.
Risks and mitigation measures.
The main potential risk concerns a breach of confidentiality or the possibility of indirect re-identification. These risks are mitigated by ensuring that record linkage and pseudonymisation are performed exclusively by registry controllers. The analysis is carried out in a secure and network-isolated SAPU environment where all activity is logged, and data export is strictly controlled. Dates and location data are truncated to reduce identifiability. Only aggregated results are published, and small cell counts are avoided. Access to the data is restricted to authorised analysts named in the ethics approval. There are no physical, psychological, or social risks to individuals, as participants are not contacted.
Principles of beneficence, non-maleficence and justice.
The study supports beneficence by producing evidence that can improve treatment quality, reduce regional disparities, and strengthen data-driven health system planning. No harm to individuals or groups is expected. Results will be reported only at group or institutional level in order to prevent any form of stigmatization or identifiability of small populations.
Data retention and disposal.
Pseudonymised individual-level data will be deleted after the completion of the project. Only aggregated and fully anonymised analytical outputs will be retained for further scientific use. The ethical analysis follows the principles of the Declaration of Helsinki and the WHO guidance on the ethical use of health data for research

13 a Human subjects

Support questions	No	Yes
Are people the object of research?		The study analyses pseudonymised registry data of individuals. There is no direct contact, no active participation, and no intervention of any kind.
Are the study participants vulnerable individuals or groups?	The dataset includes only adults, and there is no recruitment or interaction with individuals.	

Does the study include persons who cannot themselves give informed consent to participate in the research (incl. persons with limited active legal capacity)?	No consent is collected, as the legal basis for data processing is statutory authorisation for scientific research.	
Does the research involve minors as participants?	Individuals under the age of 18 are excluded from the dataset.	
Does the research involve patients as participants?	Patients are not contacted, and no study-specific data are collected. Only existing registry data are used retrospectively.	
Does the research involve collection of biological samples? Are human biological samples intended for export to a third country (https://www.aki.ee/en/guidelines-legislation/cross-border-data-protection-impact-assessment) or import them from another country to Estonia?	The study uses only electronic health and administrative data, and no biological material is processed or exchanged.	
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).** The list of all variables, their definitions, and respective data sources is provided in the annex. The dataset will include only the variables necessary to reconstruct treatment pathways, assess compliance with myocardial infarction (MI) guidelines, and evaluate one-year outcomes. 2) **Confirm that informed consent exists or is obtained before the start of the study if the study is based on consent.** The study is not based on informed consent. Data are processed under the legal basis established in Article 6(1)(e) and Article 9(2)(j) of the General Data Protection Regulation (GDPR) and §6(1) of the Estonian

		Personal Data Protection Act, which allow the processing of health data for scientific research conducted in the public interest. Individuals who have opted out from the use of their data for research are excluded by the data controllers. Individual notification is not proportionate or required for register-based research; The rights of data subjects are ensured through national notification procedures and an opt-out mechanism. 3) **Explain why all data processed are relevant and necessary (based on the principle of data minimisation).** Only variables essential to meet the research objectives will be used. These include the minimum set of identifiers required for pseudonymised linkage between registries and a limited selection of clinical and administrative variables necessary to reconstruct the treatment timeline and evaluate its compliance with guidelines. No data unrelated to the study's objectives will be processed. The analysis will be restricted to these variables only, as specified in the annex. The exact date of birth is essential for calculating age—a critical risk factor of myocardial infarction — with precision, as well as for ensuring reliable data linkage and quality control (for detection of logical inconsistencies and missing values). 4) **Are the data subjects identifiable?** Data subjects are not directly identifiable. The research team will have access only to pseudonymised datasets where all direct identifiers have been removed by TEHIK prior to data transfer. The pseudonymisation keys remain solely with the data controllers and are not accessible to researchers. a. *After the removal of the personal identifiers, the purposes of data processing are no longer achievable or would be unreasonably difficult to achieve.* For the registry linkage and temporal analysis of MI treatment journeys, pseudonymised individual-level data are necessary. Complete anonymisation before linkage would prevent the reconstruction of treatment sequences

		and the assessment of outcomes over time. b. *In the opinion of the persons conducting scientific or official statistics, there is an overriding public interest therein.* The study addresses a clear public health need by identifying variation in MI care and outcomes, supporting the development of evidence-based recommendations and improving treatment equity. This constitutes an overriding public interest recognised under the GDPR and Estonian law. c. *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.* The study does not affect the rights or obligations of the data subjects in any way. Individuals are not contacted, and no decisions or actions are taken at the personal level. Data are pseudonymised, processed only within the secure SAPU environment, and used exclusively for scientific purposes. The processing of pseudonymised registry data fully complies with the GDPR, the Estonian Personal Data Protection Act, and ISO/IEC 27001 information security principles applied in the SAPU environment
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?	The study does not involve any form of systematic monitoring, surveillance, or profiling of individuals. Data are pseudonymised registry records collected previously for administrative and clinical purposes. There is no audio, video, or geolocation data processing, and no covert data collection takes place. All analyses are performed in a secure environment under strict access control. The processing poses no risk to the rights and freedoms of data subjects, as the data are used solely for scientific research in the public interest.	

Is there a plan to analyze previously collected personal data?		**1) Explain from which database (register) or source the data are obtained.** Data will be extracted from four national health registries for patients who experienced myocardial infarction (MI) during the study period 2023–2025: the Estonian Health Information System (Digilugu), the Estonian Myocardial Infarction Registry (EMIR), the National Prescription Centre (RETS) and the Estonian Health Insurance Fund (Tervisekassa) medical billing database. The full list of variables, their definitions, and data sources is provided in the annex. These variables include treatments, appointments, and costs from billing records; medication use from RETS; and clinical and diagnostic data, including age, sex, ICD-10 codes, examinations, and laboratory results, from the Health Information System. Additional risk factor data will be retrieved from EMIR. **2) Explain how subjects are informed about their rights and the potential risks that data processing may entail.** Individuals are informed of their rights to restrict the use of their health data for scientific research through the national opt-out mechanism managed by TEHIK and described on the Estonian Health Information System website. Persons who have opted out are automatically excluded from the dataset before pseudonymisation. Since no direct contact with participants occurs, individual notifications are not feasible. The processing poses no risk to data subjects, as data are pseudonymised before analysis and handled only within a secure research environment. **3) Explain why all data processed are relevant and necessary (based on the principle of data minimization).** Only data strictly required to achieve the research objectives will be processed. These include the minimum set of variables necessary to reconstruct the treatment pathway of each MI patient and to assess compliance with national and international treatment guidelines. The principle of data minimisation will be followed, and analyses will be

		limited to the variables listed in the annex. 4) **Explain why it is not possible to study the participants in such a way that the data obtained were anonymous or pseudonymous (if applicable).** The study already uses pseudonymized data. Complete anonymization before record linkage would make it impossible to combine data across registries and to construct the individual-level treatment timeline. Pseudonymization allows the purposes of data processing to be achieved while ensuring that researchers cannot re-identify data subjects. The pseudonymization keys remain solely with the data controllers (TEHIK and the Health Insurance Fund) and are never accessible to the research team. All personal identification codes will be replaced with pseudonyms; researchers will have no access to original identification numbers. The processing of registry data fully complies with the GDPR, the Estonian Personal Data Protection Act, and ISO/IEC 27001 information security standards applied in the SAPU environment.
Is there a plan to analyze publicly available data?	The study will not analyse publicly available data. All data are obtained from national health registries that are not publicly accessible and can only be used for research under strict legal and technical safeguards. Analyses will be performed in the secure SAPU environment at the University of Tartu. Only aggregated and anonymised results may be published to inform clinical and public health practice, while no individual-level data will be made public.	
Is there an intention to transfer personal data or provide access to personal data to third countries (https://www.aki.ee/en/guidelines-legislation/cross-border-data-protection-impact-assessment)?	There is no intention to transfer personal data or to provide access to personal data to any third countries outside the EU/EEA. All data processing takes place within Estonia under the responsibility of TEHIK and the Estonian Health Insurance Fund. Analyses are conducted exclusively within	

	the secure SAPU environment at the University of Tartu, which operates entirely within the EU jurisdiction. Only aggregated and fully anonymized results may be shared publicly after ethical and data protection review. Therefore, GDPR Articles 44–50 concerning international data transfers do not apply.	
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;** All pseudonymised individual-level data will be permanently deleted at the end of the project in accordance with the data management plan approved by the data controllers. The deletion will be performed by authorised personnel within the secure SAPU environment at the University of Tartu. After the completion of the analysis and validation of outputs, only aggregated and fully anonymised statistical results will be retained. The data controllers (TEHIK and the Estonian Health Insurance Fund) will ensure that no pseudonymised records remain accessible and that the pseudonymisation keys are securely destroyed or archived under restricted conditions in compliance with the GDPR and the Estonian Personal Data Protection Act. **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 risk of indirect identification is minimised through several safeguards. Variables with high re-identification potential, such as exact dates or detailed geographic information, are generalised or truncated before analysis. Aggregation thresholds are applied when reporting results to prevent small-cell disclosure. All analyses are carried out within the SAPU secure environment, which enforces network isolation, access control, and export verification. No individual-level data leave this environment. The remaining anonymised datasets and results are reviewed to ensure that no combination

		of variables can reasonably lead to the re-identification of a data subject. Consequently, the residual risk of indirect identification is assessed as very low.
13 c Other ethical issues		

Can conducting research involve ethical risks not described above?	The study does not involve any additional ethical risks beyond those already described. It uses only pseudonymized registry data, with no contact with participants and no physical, psychological, or social intervention. All procedures follow established ethical and legal standards for secondary data use in health research, and no harm or discomfort to individuals or groups is foreseen. **Other ethical considerations:** The research serves a clear public health interest by improving the quality, safety, and equity of myocardial infarction care in Estonia. The results will be presented only in aggregated form to prevent identification of individuals or healthcare providers and to avoid any potential stigmatization of patient groups or institutions. The project follows the principles of fairness, transparency, and accountability in accordance with the Declaration of Helsinki, the WHO guidance on the ethical use of health data for research, and the European Code of Conduct for Research Integrity. Potential conflicts of interest will be declared, and the results will be disseminated responsibly through scientific and public health channels. No commercial use of the data or results is planned.	

14. Development, deployment and/or use of artificial intelligence (AI)-based systems or techniques **See:** https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai		
Support questions	**Ei**	**Jah**
Does this activity involve the development, deployment and/or use of Artificial Intelligence-based systems?	**x**	**1)** Explanation as to how the respect to fundamental human rights and freedoms (e.g. human autonomy, privacy and data protection) will be ensured. **2)** Detailed risk assessment accompanied by a risk mitigation plan: a) the abilities, limitations, risks and benefits of the proposed AI system/technique; b) Details on the measures taken to avoid bias in input data and algorithm design.
Could the AI based system/technique potentially stigmatise or discriminate against people?	**x**	Detailed explanation of the measures set in place to avoid potential bias, discrimination and stigmatisation
Does the AI system/technique interact, replace or influence human decision-making processes?	**x**	**1.** Detailed explanation on how humans will maintain meaningful control over the most important aspects of the decision-making process. **2.** Explanation on how the presence/role of the AI will be made clear and explicit to the affected individuals.
Does the AI system/technique have the potential to lead to negative social and/or environmental impacts?	**x**	**1.** Justification of the need for developing/using this particular technology **2.** Assessment of the ethics risks and detailed description of the measures set in place to mitigate the potential negative impacts during the research, development, deployment and post-deployment phase.

15. Complete in case the research is based on data from a database and/or register	
Name of database and/or register	The study will combine data from national registers: • Estonian Health Information System (Digilugu) • Estonian Myocardial Infarction Registry (EMIR) • Cause of Death Register (SPR) • Retseptikeskus (RETS) • Medical bills from Tervisekassa (KIRST)
Purpose of the processing of personal data	We will conduct a scientific research study of the concordance to treatment guidelines for AMI patients in the one-year follow-up period, to

	quantify risk factors and how concordance varies depending on them, and variation in health outcomes (such as risk for recurrent infarction or death) during the follow-up period. The results of the study will be disseminated as scientific publications and shared with the research community.
List of variables and period for which data are collected (in annex if necessary) Please see Annex	
16. 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.	The scientific aim of the project necessitates using personal data, but principles of data minimisation will be followed. Data records will be pseudonymised and information requested from each register is restricted to data points necessary for carrying out the analyses. Date and time information will be truncated to date. The individual level data will be deleted from storage at the end of the project.
Describe the process and means of pseudonymisation of personal data.	We have arranged that, upon receiving ethical approval from EBIN and TAIEK and permission from the data owners (SoM and Tervisekassa), the relevant records will first be extracted from the registers by TEHIK and Tervisekassa. TEHIK will then work with Tervissekassa to link the datasets (using personal identifiers) and pseudonymise the dataset, and act as a data custodian for this data during the project. All personal identification codes will be replaced with pseudonyms; researchers will have no access to original identification numbers.
Is there a plan to de-pseudonymise gene donors' data?	1) Please specify the number of gene donors whose data will be de-pseudonymised. 2) Please explain the reason for de-pseudonymisation. N/A (we do not ask for gene donor's data)
Is there a plan to transport personal data? Please describe how data protection is ensured.	The data will be processed in the University of Tartu (UT) secure processing environment (SAPU). TEHIK will be responsible for transferring the data to the secure environment and for deleting the data at the end of the project. Only outcomes of statistical analyses and other anonymous data resulting from the project will be copied out of the SAPU.
Describe technical and organizational measures used to protect data from unauthorized access or processing.	The UT SAPU is a platform for sensitive data analyses. It achieves a high level of security, having complete network isolation based on firewall rules. Access to the machine is possible only through a virtual desktop environment, accessible through encrypted connections, and only analysts named on the ethical approvals will have access to the environment. The monitoring layer and the server record all actions taken, including any copying or moving of the data. Copying any individual-level data out of the environment will be forbidden. Exporting the results of the analyses entails the analysts moving the results files to the export area of the environment, where they will be reviewed by the data custodian at TEHIK for compliance with anonymization requirements before released.

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** 10/10/2025
EBIN ID of the application **(fills by the assessor)**	

List of additional documents:

1. CV of the principal investigator

2. List of variables