

UURIMISTÖÖ AVALDUS KOOSKÕLASTUSE SAAMISEKS TARTU ÜLIKOOLI INIMUURINGUTE EETIKA KOMITEELT

1. UURIMISTÖÖ ANDMED

Uurimistöö täielik nimetus: E-sotsiaalandmete kasulikus südame-veresoonkonna haigustekke ennustamisel.

Uurimistöö toimumiskoht: Tartu Ülikooli peremeditsiini ja rahvatervishoiu instituudis (kusjuures andmeid lingitakse ja analüüsitakse Tartu Ülikooli teadusarvutuste keskus.)

2. UURIMISTÖÖ TEOSTAJAD JA UURINGUKESKUSED

2.1. Vastutav uurija

ees- ja perekonnanimi: Taavi Tillmann
teaduslik kraad: PhD
amet: rahvatervishoiu kaasprofessor
töökoht: Tartu Ülikool, peremeditsiini ja rahvatervishoiu instituut
töökoha aadress: Ravila 19, 50411 Tartu, Eesti
telefoninumber: 5342 6963
e-post: taavi.tillmann@ut.ee
allkiri: /lisatakse digitaalselt/

2.2. Kaastöötajad

ees- ja perekonnanimi: Raivo Kolde
teaduslik kraad: PhD
amet: terviseinformaatika kaasprofessor
töökoht: Tartu Ülikool, arvutiteaduse instituut
töökoha aadress: Narva mnt 18, 51009 Tartu, Eesti
telefoninumber: 506 7961
e-post: raivo.kolde@ut.ee
allkiri: /lisatakse digitaalselt/

ees- ja perekonnanimi: Anneli Uusküla
teaduslik kraad: PhD
amet: epidemioloogia professor
töökoht: Tartu Ülikool, peremeditsiini ja rahvatervishoiu instituut
töökoha aadress: Ravila 19, 50411 Tartu, Eesti
telefoninumber: 737 4195; 518 3552 (4195)
e-post: anneli.uuskula@ut.ee
allkiri: /lisatakse digitaalselt/

ees- ja perekonnanimi: Nikita Umov
teaduslik kraad: Arstiteaduse integreeritud BSc ja MSc
amet: doktorant / kliiniline nooremteadur
töökoht: Tartu Ülikool, peremeditsiini ja rahvatervishoiu instituut
töökoha aadress: Ravila 19, 50411 Tartu, Eesti
telefoninumber: 5682 8380
e-post: nikita.umov@gmail.com Alates 01.09.2029: nikita.umov@ut.ee
allkiri: /lisatakse digitaalselt/

ees- ja perekonnanimi: Laura Ilves
teaduslik kraad: BSc

amet: magistrant / spetsialist
töökoht: Tartu Ülikool, peremeditiini ja rahvatervishoiu instituut
töökoha aadress: Ravila 19, 50411 Tartu, Eesti
telefoninumber: 737 4190
e-post: laura.ilves.99@gmail.com
allkiri: /lisatakse digitaalselt/

ees- ja perekonnanimi: Eva-Maarja Mand
teaduslik kraad: BSc
amet: magistrant / spetsialist
töökoht: Tartu Ülikool, peremeditiini ja rahvatervishoiu instituut
töökoha aadress: Ravila 19, 50411 Tartu, Eesti
telefoninumber: 737 4190
e-post: evamaarjamand@hotmail.com
allkiri: /lisatakse digitaalselt/

ees- ja perekonnanimi: Markus Haug
teaduslik kraad: MSc
amet: terviseinformaatika nooremteadur
töökoht: Tartu Ülikool, arvutiteaduse instituut
töökoha aadress: Narva mnt 18, 51009 Tartu, Eesti
telefoninumber: 56886522
e-post: markus.haug@ut.ee
allkiri: /lisatakse digitaalselt/

ees- ja perekonnanimi: Marek Oja
teaduslik kraad: PhD
amet: terviseinformaatika teadur
töökoht: Tartu Ülikool, arvutiteaduse instituut
töökoha aadress: Narva mnt 18, 51009 Tartu, Eesti
telefoninumber: 737 5445
e-post: marek.oja@ut.ee
allkiri: /lisatakse digitaalselt/

ees- ja perekonnanimi: Sirli Tamm
teaduslik kraad: MSc
amet: programmeerija
töökoht: Tartu Ülikool, arvutiteaduse instituut
töökoha aadress: Narva mnt 18, 51009 Tartu, Eesti
telefoninumber: 737 5445
e-post: sirli.tamm@ut.ee
allkiri: /lisatakse digitaalselt/

ees- ja perekonnanimi: Elena Sügis
teaduslik kraad: PhD
amet: terviseinformaatika lektor
töökoht: Tartu Ülikool, arvutiteaduse instituut
töökoha aadress: Narva mnt 18, 51009 Tartu, Eesti
telefoninumber: 737 6437
e-post: elena.sugis@ut.ee
allkiri: /lisatakse digitaalselt/

ees- ja perekonnanimi: Hendrik Šuvalov
teaduslik kraad: MSc
amet: terviseinformaatika nooremteadur

töökoht: Tartu Ülikool, arvutiteaduse instituut
töökoha aadress: Narva mnt 18, 51009 Tartu, Eesti
e-post: hendrik.suvalov@ut.ee@ut.ee
allkiri: /lisatakse digitaalselt/

2.3. Vastutava uurija asutuse juhataja või tema kohusetäitja on uurimistöö korraldamisega nõus:

ees- ja perekonnanimi: Kaja-Triin Laisaar
amet: instituudi juhataja
töökoht: Tartu Ülikool, peremeditsiini ja rahvatervishoiu instituut
allkiri: /lisatakse digitaalselt/

3. UURIMISTÖÖ FINANTSEERIMINE

Allikas 1: Saadud stardigrant (PSG809), rahastus periood 01.01.2023–31.08.2028. Koostööleping Tartu Ülikooli ja ETAG vahel on sõlmitud.
Allika eelarve: 378 620 EUR
Töötasude jaotus: Vähemalt 75% rahast kulub käesoleva projekti tööaja osaliseks katteks (peamiselt Taavi palk).

Allikas 2: Saadud TÜ riigieelarveline doktorantuuri koht (meditsiiniteaduste valdkond, “Meditsiin ja Sport” programm, arstiteaduse eriala), pealkirjaga “Südameveresoonkonnahaiguste personaalne ennetus ja digilahendused”. Rahastus periood 01.09.2023–31.08.2027.
Allika eelarve: ca 115 000 EUR
Töötasude jaotus: Kogu raha kulub Nikita palga katteks.

Allikas 3: Saadud rühmagrant (PRG1844), rahastus periood 01.01.2023–31.12.2027. Koostööleping Tartu Ülikooli ja ETAG vahel on sõlmitud.
Allika eelarve: 270 000 EUR
Töötasude jaotus: Ligikaudselt 10% rahast kulub käesoleva projekti tööaja osaliseks katteks (sh. Raivo, Marek, Sirli, Markus ja Hendriku palk).

Uuritavatele kompensatsiooni maksmine: Ei

4. KAVANDATAVA UURIMISTÖÖ TEADUSLIK EESMÄRK JA PÕHJENDUS

4.1. Kirjanduse ülevaade

All kirjeldame laiemat avalikku huvi antud probleemi vastu, siis meie laiemat strateegiat kuidas plaanime seda mitmete projektide üleselt lahendada ja viimaseks käesolevat projekti mis on esimene samm selle teostamisel.

4.1.1. Suur haiguskoorem ja avalik huvi

Südame-veresoonkonna haigused (SVHs) on Eestis juhtiv haiglaravi põhjus ja juhtiv surmapõhjus.¹ Täpsemini, aastal 2022 suri vereringeelundite haigustesse (I00-I99) 8102 inimest millest 39% olid haigused kus on eriti tugev teaduspõhisisus statiinravimite tõhususe kohta: südame isheemiatõved

¹ TAI. Surmad põhjuse, soo ja vanuserühma järgi. 2022.

https://statistika.tai.ee/pxweb/et/Andmebaas/Andmebaas_01Rahvastik_04Surmad/SD21.px/

(I20-I25) ja peaaajuveresoonte haigused (I60-I69).^{2,3} Käesolev projekt on esimene samm laiemast ambitsioonist luua Eestisse maailma esimene kulutõhus SVHi ennetav sõeluuringu teenus. Antud teenust on eriti otstarbekas Eestis arendada ja juurutada, kuna antud haigus on Eestis 3x sagedasem kui maailma keskmine ja 2x sagedasem kui Euroopa keskmine.⁴ Tulenevalt on teenuse arendamine Eestis statistiliselt tõhusam ja Eesti ühiskonna avalik huvi teenuse kasude vastu on erakordselt suur. Lisaks on antud teenusel teoreetiline võimekus langetada riigisiseseid tervise ebavõrdsusi erinevate sotsiaalmajanduslike gruppide vahel, mis on üks kolmest „*Eesti rahvastiku tervise arengukavaga*“ tipp-prioriteetidest.⁵ Näiteks elavad kõrgharidusega mehed Eestis 10,6 aastat kauem kui põhiharidusega mehed.⁶ Seda vahet mõjutab SVH (sh infarkt ja insult) rohkem kui teised haigused.⁷

4.1.2. Jätku-projektides arendatava sõeluuring vajadused

SVH tekkeriski saab ennustada riskikalkulaatoritega. Need on sisendiks ennetusotsuste tegemisel. Näiteks on Euroopa Kardioloogide Seltsi 2021. a ravijuhise kohaselt asümptomaatilisel patsiendil vanuses 50–69, kelle 10 a. SVH tekkerisk on >10%, näidustatud kolesterooli alandavate ravimite (statiinide) kasutamine.⁸ Uuringumeeskonnal on kogemus SVH riskikalkulaatorite väljatöötamisel,¹⁰ sh SCORE2 riskikalkulaatoriga,⁹ mida soovib Euroopa Kardioloogide Selts. SCORE2 riskihinnang põhineb demograafilistel ja tervist kirjeldavatel teguritel (vanus, sugu, suitsetamine, vererõhk, kolesterool). SCORE2 kalkulaator ei sisalda sotsiaalmajanduslike tegureid. On olemas arvestatav teaduslik tõendus kohort uuringutest, et sotsiaalmajanduslikud tegurid on tugevad SVH tekke ennustajad.^{10, 11, 12, 13, 14, 15, 16} Näiteks vaesematel inimestel on kõrgem SVH risk. Paraku kasutavad vaesemad inimesed hoopis vähem ennetavaid ravimeid kui rikkad inimesed.¹⁷ Osaliselt võib see tuleneda sellest, et vaesus jt sarnased tegurid SVH riskikalkulaatoritest puuduvad.

² Collins R. et al. Interpretation of the evidence for the efficacy and safety of statin therapy. The Lancet. 2016.

Yusuf S, et al. Cholesterol lowering in intermediate-risk persons without cardiovascular disease. New England Journal of Medicine. 2016

³ Mihaylova B. The effects of lowering LDL cholesterol with statin therapy in people at low risk of vascular disease: meta-analysis of individual data from 27 randomised trials. The Lancet. 2012.

⁴ GBD 2019 Risk Factors Collaborators. Global burden of 87 risk factors in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet. 2020; 396: 1223-1249 <http://ihmeuw.org/66zk>

⁵ Sotsiaalministeerium. Rahvastiku tervise arengukava 2020-2030. 2021. <https://www.sm.ee/rahvastiku-tervise-arengukava-2020-2030>

⁶ Statistikaamet. Elada jäänud aastad soo ja hariduse järgi 2021. RV0454. <https://www.stat.ee/et/avasta-statistikat/valdkonnad/heaolu/tervis/oodatav-eluiga>

⁷ Currie J, et al. Life expectancy inequalities in Wales before COVID-19: an exploration of current contributions by age and cause of death and changes between 2002 and 2018. Public Health. 2021;193:48-56.

⁸ Visseren FL. 2021 ESC Guidelines on cardiovascular disease prevention in clinical practice: Developed by the Task Force for cardiovascular disease prevention in clinical practice with representatives of the European Society of Cardiology and 12 medical societies. Eur H J. 2021;42(34):3267 (Table 6).

⁹ Hageman S...Läll K, Tillmann T...Visseren F... Angelantonio ED. SCORE2 risk prediction algorithms: new models to estimate 10-year risk of cardiovascular disease in Europe. Eur H J; 2021;42(25): 2439-2454

¹⁰ Tillmann T, Vaucher J, Okbay A, ... Davey Smith G, Bobak M, Holmes MV. Education and coronary heart disease: a Mendelian randomization study. BMJ. 2017; 358:j3542.

¹¹ Tillmann T, Pikhart H, Peasey A, ... Steptoe A, Kivimäki M, Marmot M, Bobak M. Psychosocial and socioeconomic determinants of cardiovascular mortality in Eastern Europe: a multicentre prospective cohort study. PLOS Medicine. 2017; 14(12):e1002459.

¹² Tillmann T, Läll K, Dukes O, ... Fischer K, Kivimäki M, Bobak M. Development and validation of two SCORE-based cardiovascular risk prediction models for Eastern Europe: a multicohort study. Eur H J. 2020;41(35):3325-33.

¹³ Hinkle LE Jr, Whitney LH, Lehman EW, et al. Occupation, education, and coronary heart disease. Risk is influenced more by education and background than by occupational experiences, in the Bell System. Science. 1968; 358:238-46.

¹⁴ Veronesi G, Ferrario MM, Kuulasmaa K, et al. Educational class inequalities in the incidence of coronary heart disease in Europe. Heart. 2016; 358:958-65.

¹⁵ Berkman LF, Syme SL. Social networks, host resistance, and mortality: a nine-year follow-up study of Alameda County residents. American Journal of Epidemiology. 1979; 109(2):186±204

¹⁶ Barth J, Schneider S, von Känel R. Lack of social support in the etiology and the prognosis of coronary heart disease: a systematic review and meta-analysis. Psychosomatic Medicine. 2010; 72(3):229±38

¹⁷ Carter AR, Gill D, Smith GD, Taylor AE, Davies NM, Howe LD. Cross-sectional analysis of educational inequalities in primary prevention statin use in UK Biobank. Heart. 2021 Jul 26.

Tulenevalt planeerime jätku-uuringuid kus lingime käesoleva projekti andmebaasi täiendavate andmebaasidega (Tervise Infosüsteem, et tuvastada käitumise tegureid ja geenivaramu, et kontrollida loodud kalkulaatori täpsust) misjärel kaalume päriselu piloot katseid.

4.1.2. Käesoleva projekti kirjanduse ülevaade ja teadusliku tõenduse lünk

Kuigi kohort-uuringud on veenvalt näidanud, et sotsiaalmajanduslikud tegurid ennustavad SVHd,¹⁰⁻¹⁶ pole seni ükski uurimisrühm uurinud kas ka olemasolevatest e-sotsiaalandmetest mõõdetud sotsiaalmajanduslikud tegurid ennustavad SVH teket ja suremust. Defineerime ***e-sotsiaalandmeid*** kui informatsiooni uuritava maakonna, rahvuse, hariduse, ameti, tööpuuduse, sissetuleku, varade, sotsiaaltoetuste ja abielustaatus kohta, mida on võimalik hankida olemasolevatest administratiivsetest riiklikest andmebaasidest. Antud uuringu raames loeme ***sotsiaalseteks andmebaasideks*** seitset andmebaasi, mis annavad infot uuritavate demograafia, töötuse, töötamise, tulude, varade ja toetuste kohta (vt Lisa 10.1: rahvastikuregister; töötuna ja tööotsijana arvel olevate isikute ning tööturuteenuste osutamise register; töötamise register; maksukohustuslaste register; sotsiaalkaitse infosüsteem; kinnistusraamat; äriregister).

Esialgsetes vestlustes teiste regulaatoritega (sh. Eesti bioetika ja inimuuringute nõukogu) tuvastasime kuidas jääb alles võimalus, et e-sotsiaalandmed ei ennusta Eestis SVH teket. Kui nii, siis pole jätku-uuringuid vaja teostada. Seega on käesoleva uuring esimene samm, kus võimalikult väikse andmebaasiga tuvastame millistel e-sotsiaalandmetel on võimekus edasi liikuda jätku-uuringutesse.

4.3.1 esimene uuringuküsimus: Milline kogum e-sotsiaalandmeid kõige paremini ennustab südame veresoontõve haiguste (SVH) 10.a. teket tervete inimeste seas?

Hüpotees: Igas seitsmes andmebaasis leidub vähemalt üks tegur mis ennustab SVH teket.

Põhjus: Projekti eesmärk on jätku-uuringutes välistada need andmebaasid mille empiiriline esmane analüüs ei vasta meie kirjandusest leitud ootustega. Näiteks, kui mõni usutav riskitegur ei ennusta SVH teket, siis loobume selle uurimisest jätku-uuringutes, kooskõlas andmete minimeerimise printsiibiga. Käesolevas faasis teostame küll piiratud aga siiski kasulikud esmased analüüsid **ilma andmete sidumist Tervise Infosüsteemiga ega geenivaramuga**, taaskord kooskõlas andmete minimeerimise printsiibiga.

4.3.2 teine uuringuküsimus: Kas ülaltoodud haiguse ennustust on sama täpne teha kasutades selleks kas 01.01.2012, 01.01.2014 või 01.01.2016. hetkel kogutud andmeid?

Hüpotees: SVH riskikalkulaator, mis luuakse 01.01.2012 andmetega, on sama täpne kui riskikalkulaator, mis luuakse 01.01.2014 või 01.01.2016 hetkeks kogutud andmetega.

Põhjus: Juhul kui hüpotees on ümber lükatud ja riskikalkulaatori täpsus muutub ajas, annab see meile olulist sisendit jätku-uuringu disainimisel ja otsustamisel. Näiteks võime leida, et kuna andmebaas-X kvaliteet ja ennustuvõime langeb ajas loobume selle uurimisest jätku-uuringutes. Või kuna andmebaas-X kvaliteet ja ennustuvõime kasvab ajas, siis vaatamata kasinatele ennustustulemustele jätkame selle uurimist jätku uuringutes kus ennustuvõime võib taas suurened.

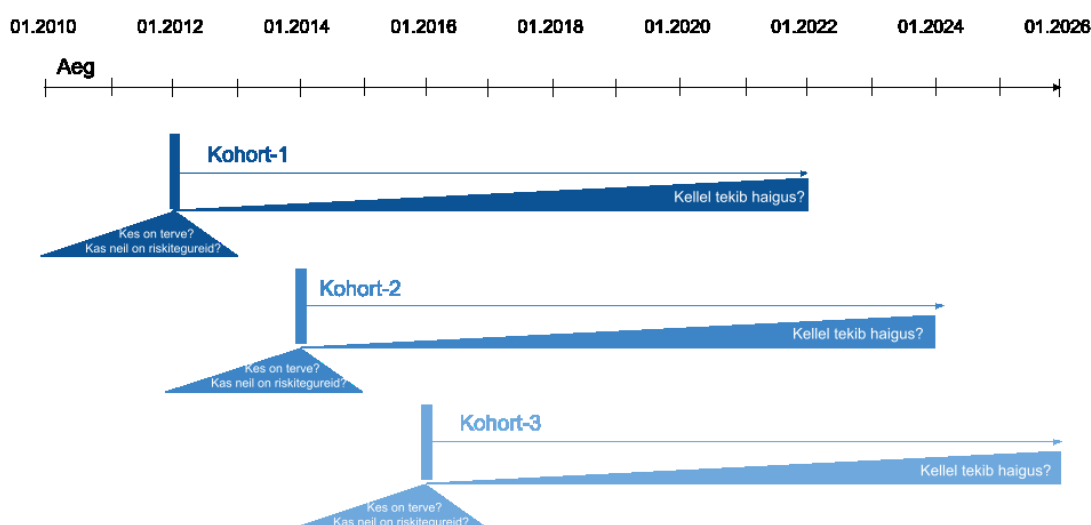
5. UURIMISTÖÖ TEOSTAMISE AEG (kuu.aasta-kuu.aasta)

10.2023–08.2028

6. LOODAVA ANDMEBAASI STRUKTUURI KIRJELDUS

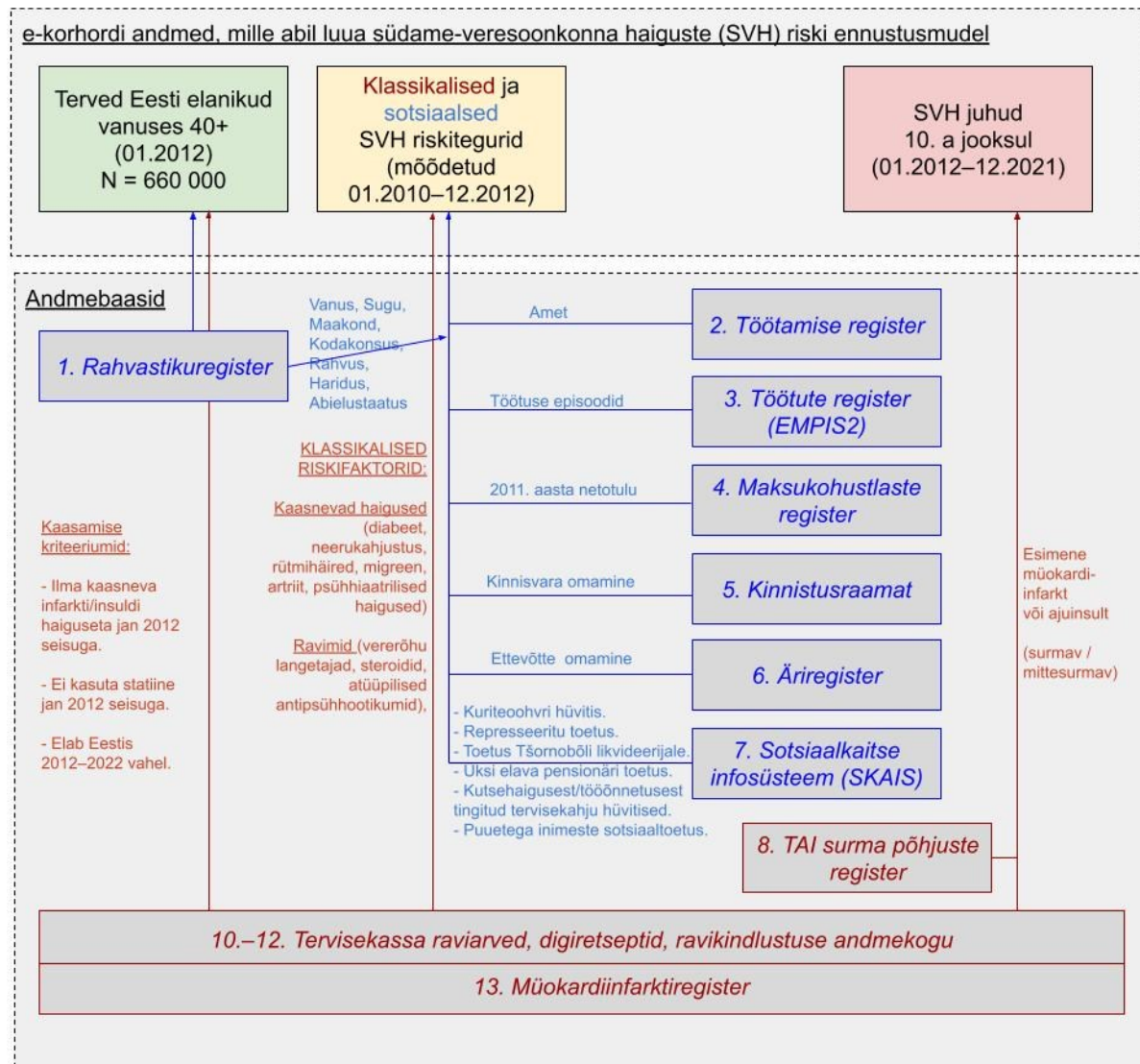
Uuringu disain

Kõnealune Eesti rahvastikupõhine uuring analüüsib andmeid riiklikest sotsiaalsetest andmebaasidest. Uuringu kolmes valimis on 40+aastased elanikud kolmel ajahetkel: 2012. a, 2014. a ja 2016. a 1. jaanuaril. Kolme kohordi liikmete seas on suur kattuvus. Uurijad jälgivad kolme kohordi liikmete seas SVH tekkeriski kuni 2026. aasta lõpuni, et tuletada ja testida SVH riskikalkulaatoreid (joonis 1).



Joonis 1. Uuringu kohortide defineerimine ja andmete linkimine SVH riski hindamiseks.

Iga kohort omakorda luuakse üheteistkümnest riiklikust andmebaasist saadavate andmeväljade linkimise teel. Andmete koosseis ja põhjendus on esitatud kokkuvõtvalt joonisel 2 (ja detailsemalt lisas 1).



Joonis 2. Kohort-1 moodustamine, analüüsi kaasatavad riskitegurid ja andmete allikad.

7. METOODIKA (VALIMI KIRJELDUS)

7.1. Valim

Kõik Eesti elanikud kes on vanuses 40+ ja elasid Rahvastikuregistri alusel kas 01.01.2012, 01.01.2014 või 01.01.2016 seisuga Eestis.

Loodud e-kohordi suurus on ligikaudselt 660 000 residenti. Siin tekib 10 a. jooksul orienteeruvalt 5000 SVH haigusjuhtu.

7.2. Statistiline võimsus ja põhjendus

Uuringu esimene küsimus keskendub sihitud riskitegurite tuvastamisele ja teine küsimus keskendub kalkulaatori täpsusele (kusjuures kalkulaatoris on mitu riskitegurit kokku lisatud). Erinevaid võimsusarvutusi on vajalik mõlemaks etapiks.

7.2.1. Esimese küsimuse võimsus (üksikud tegurid)

Koostasime võimsusarvutuse 68 riskitegurile mille puhul oli võimalik prognoosida riskiteguri levimus (ing. k. prevalence) ja võimalik riskimäärade suhe (ing. k. Hazard Ratio). See näitas, et igas andmebaasis on vähemalt üks riskitegur, mille leidmiseks on vaja vahemikus 330,000 kuni 660,000 uuritavat. Kõigis alltoodud oleme eeldanud, et 0.758% inimestel ilmneb, kas surmav või mittesurmav SVH 10 a. jälgimisaja jooksul. Igal osutatud valimil on 80% võimsus (st iga viiest riskitegurist me ei suuda leida alltoodud valimi suurusega, suurema valimi suurusega see tõenäosus kahaneb). Alfa on $P < 0.05$. Riskimäärade suhe on sageli hinnatud lihtsamas analüüsis, kus on väike loetelu kaasnevaid riskitegureid. Eeldusel, et riskitegurid on omavahel osaliselt korreleeritud, võib lõplik riskimäärade suhe osutada hinnatust väiksemaks, mis omakorda tõstab vajaliku valimi suurust (seda nii alltoodud 17 teguri puhul kui ka ülejäänud 51 teguri puhul mille valimi vajadus oli alltoodust väiksem).

Andmebaas	Riskitegur	Levimus	Riskimäärade suhe	Vajalik valimi suurus
Raviarved ja digireseptid	Migreen	2.7%	1.29 ¹⁸	475,487
	Krooniline neerupuudulikus, tase 4-5	0.2%	2.37 ¹⁸	357,500
Rahvastiku-register	Keskeriharidus (vs. Keskhariidus)	17.4% ¹⁹	0.90	400,806
	Magistri/doktorikraad (vs. Keskhariidus)	0.5% ¹⁹	0.69 ²⁰	528,540
Töötute register	Uuringule eelneval 2 aastal oli uuritav rohkem tööl kui töötu	10%	1.15 ²¹	516,657

¹⁸ Hippisley-Cox J, Coupland C, Brindle P. Development and validation of QRISK3 risk prediction algorithms to estimate future risk of cardiovascular disease: prospective cohort study. BMJ. 2017;357.

¹⁹ Statistikaamet. Rahva ja eluruumide loendus 2000: Rahvastik haridustaseme ja rahvuse järgi https://andmed.stat.ee/et/stat/rahvaloendus_rel2000_haridus/RL302

²⁰ Tillmann T, Vaucher J, Okbay A, Pikhart H, Peasey A, Kubinova R, Pajak A, Tamosiunas A, Malyutina S, Hartwig FP, Fischer K. Education and coronary heart disease: mendelian randomisation study. BMJ. 2017;358.

²¹ Tillmann T, Läll K, Dukes O, Veronesi G, Pikhart H, Peasey A, Kubinova R, Kozela M, Pajak A, Nikitin Y, Malyutina S. Development and validation of two SCORE-based cardiovascular risk prediction models for Eastern Europe: a multicohort study. European Heart Journal. 2020;41(35):3325-33.

Sotsiaalkaitse infosüsteem	Saab kutsehaiguse tervisekahju hüvitist	2.0%	1.24 ²²	617,915
	Saab represseeritu toetust	1.0%	1.50	464,400
Töötamise register	Tehnik või keskastme spetsialist (vs. ametnik või kontoritöötaja)	12%	0.90 ²³	392,272
	Tippspetsialist (vs. ametnik või kontoritöötaja)	19%	0.90 ²³	449,346
	Juhtimisorgani liige (vs. tööleping)	3%	0.78 ²⁴	639,580
	Vabatahtlik töö (vs. tööleping)	3%	0.76 ²⁵	534,671
Maksukohustuslaste register	Kuupalk neto 700-799 EUR (vs. 400-499 ehk tollane mediaan)	8%	0.91	515,450
Kinnistusraamat	Omab „kahe korteriga elamut“	2.0%	1.15	367,470
	Omab kinnisvara suurusega 61-80m2 (vs >101m2)	16.5% ¹⁹	1.10 ²⁶	455,535
Äreregister	Omab ettevõttest 100% (vs. ei oma ettevõtteid)	10%	0.85	406,305
	Omatud ettevõttel on maksuvõlg	2%	1.20	332,277
	Omatud ettevõttel on 2-10 töötajat (vs. 0-1 töötaja)	5%	0.90	443,184

7.2.2. Teise küsimuse võimsus (koondkalkulaator)

Meetodid alles arenevad, mis aitavad hinnata sobivat valimi suurust igale riskikalkulaatori loomise väljakutsele. Tuletame sobiva valimi suuruse trianguleerides argumente kolmest vaatenurgast:

7.2.2.1. *Tavapraktika* sarnaste riskiennustus kalkulaatorite väljatöötamisel

Riskikalkulaatorite metodoloogilise soovitusel kohaselt on suurematel valimitel loodud kalkulaatorite ennustusvõime täpsem. Väiksemate valimitega tekib oht, et kalkulaator näib alguses toimivat täpsena, aga tegelikult on mudel ülekohandatud ja seega on selle täpsushinnang **eksitavalt optimistlik**.^{27, 28} Vajaduselt saab seda optimismi osaliselt tuvastada ja selle alusel loodud koefitsiente nulli suunas allapoole korrigeerida. Kuna vastavad korrektsioonid on ebatäpsed,²⁹ soovitatakse mudeli loojatel ülesobitamist ja optimismi võimalikult palju vältida, kasutades võimalikult suurt valimit, mis seeläbi suurendab loodud kalkulaatori täpsust.

Lähtuvalt ülaltoodud põhimõttest on teised riskikalkulaatorid (mida teiste riikide ravijuhised soovivad) loodud äärmiselt suurte valimitega. Näiteks Inglismaa QRISK3 loodi 8 miljon residentide andmetega.²⁰ Euroopa SCORE2 riskikalkulaatori koefitsiendid määrati 677 684 valimiga pluss kalibratsiooniks analüüsiti haigusteket 11 miljonil inimesel.⁷ Leiame, et antud valim (660 000 uuritavat) on küll väiksem kui Euroopa eelkäijad, aga siiski piisavalt suur millel luua usaldusväärne riskihinnang tuvastamiseks milliseid andmevälju ja andmebaase kasutada jätku-uuringutes.

²² Boden LI, O'Leary PK, Applebaum KM, Tripodis Y. The impact of non-fatal workplace injuries and illnesses on mortality. *American journal of industrial medicine*. 2016;59(12):1061-9.

²³ Virtanen SV, Notkola V. Socioeconomic inequalities in cardiovascular mortality and the role of work: a register study of Finnish men. *International Journal of Epidemiology*. 2002;31(3):614-21.

²⁴ Toivanen S, Griep RH, Mellner C, Vinberg S, Eloranta S. Mortality differences between self-employed and paid employees: a 5-year follow-up study of the working population in Sweden. *Occupational and environmental medicine*. 2016;73(9):627-36.

²⁵ Okun MA, Yeung EW, Brown S. Volunteering by older adults and risk of mortality: a meta-analysis. *Psychology and aging*. 2013;28(2):564.

²⁶ Laaksonen M, Tarkiainen L, Martikainen P. Housing wealth and mortality: A register linkage study of the Finnish population. *Social science & medicine*. 2009;69(5):754-60.

²⁷ Babyak MA. What you see may not be what you get: a brief, nontechnical introduction to overfitting in regression-type models. *Psychosomatic medicine*. 2004 May 1;66(3):411-21.

²⁸ Steyerberg EW. Overfitting and optimism in prediction models. In *Clinical prediction models* 2019: 95-112. Springer.

²⁹ Riley RD, et al. Penalization and shrinkage methods produced unreliable clinical prediction models especially when sample size was small. *Journal of clinical epidemiology*. 2021;132:88-96.

7.2.2.2. Nüüdisaegsed statistilised arengud ja soovitused

2020.a. avaldas mainekas teadusajakiri The BMJ artikli, mis soovib järgnevat. Loodud riskiennustus kalkulaatori optimism (ja vajadus koefitsiente alla korrigeerida) jääks vahemikku 0–10%.³⁰

Käesolevas uuringus loodav riskikalkulaator määrab 164 parameetrit, selle eeldatav ennustuse täpsus on 0,019 (n-ö „adjusted_R2CS“ alusel) ja haigustekke sagedus 0,00076. Seega oleks loodud riskikalkulaatori eksitav optimism orienteeruvalt 1,3%. See paistab uurijatele optimaalne, kuna suurem valim ja väiksem optimism annaks vähe lisanduvat usaldusväärsust. Teiselt poolt väiksem valim ja suurem optimism tõstaks tõenäosust teha kahju ja langetaks tõenäosust teha kasu. Lisaks võib väiksemal valimil loodud riskikalkulaator langetada selle usaldusväärsust ja vastuvõetavust perearstide ja patsientide poolt.

7.2.2.3. Eelmise 25.a. jooksul kasutusel olev rusikareegel: „10 juhtumit tunnuse kohta“.

Kuna ülaltoodud arvutus ei pruugi olla intuiitiivselt arusaadav kõigile osapooltele pakume ka lisanduvat selgitust ja arvutust. See küll tugineb tänaseks aegunud soovitustele aastast 1998 (mis siiski sai sel perioodil üle 8000 viidet³¹). Sealset loogikat ja matemaatilist korrektsust on kergem jälgida.

Igas uuringu kohordis tekib jälgimisperioodil ligikaudselt 5000 SVH juhtumit. Nendest ligikaudselt 1670 tekib naistel ja 3330 tekib meestel. Loomes ühe kalkulaatori meeste ja teise naistele. Vaatame kui palju võimsust on naistele loodud kalkulaatoril. Üks riskitegur (nt haridus) võib vajada enam kui ühte parameetrit (nt hariduse puhul mõõdame 8 hariduse kategooriat). Iga parameeter vajab vähemalt 10 juhtumit, et sealt arvutatud riski koefitsient ei oleks liiga eksitavalt optimistlik.⁶ Eeldame, et kõik juhtumid on perfektselt jaotatud kõigi parameetrite ehk riskitegurite omavaheline korrelatsioon on null. Naiste riskikalkulaatori loomisel saame seega mudeldada $1670 / 10 = 167$ parameetrit. See kattub üllatavalt hästi meie plaaniga, kuna soovime mudeldada ligikaudselt 164 parameetrit. Kokkuvõttes, ülaltoodud eelduste ja populaarse rusikareegli alusel on Eesti naiste valimi suurus just optimaalne antud eesmärgi saavutamiseks.

Juhul kui tegurid omavahel korreleeruvad või mõnel teguril on vähem juhtumeid, siis on vaja suuremat valimit. Täpsem valimi arvutus mida kirjeldasime „7.2.2.“ sammus juba arvestas riskitegurite korrelatsioone ja on seega kõige parim hinnang, mis illustreerib valimi suuruse sobivust.

7.3. Andmete linkimine

Lühidalt, Siseministerium tuvastab valimis olevad uuritavad ja loob igale isikule projektipõhise pseudonüümi. Siseministerium saadab andmevõtme teistele andmekogude omanikele, kes tuvastavad valimis olevad uuritavad. Kõik andmekogud edastavad pseudonümiseeritud andmed Tartu Ülikoolile. Seda illustreerib joonis 3.

Detailsemalt kirjeldades, rahvastikuregistri omanik (siseministerium) pseudonümiseerib andmed (sh eemaldades nimed ja isikukoodid, asendades sünnikuupäeva vanusega), loob igale inimesele projektipõhise pseudonüümi ja väljastab pseudonümiseeritud andmed Tartu Ülikooli teadusarvutuste keskusele.

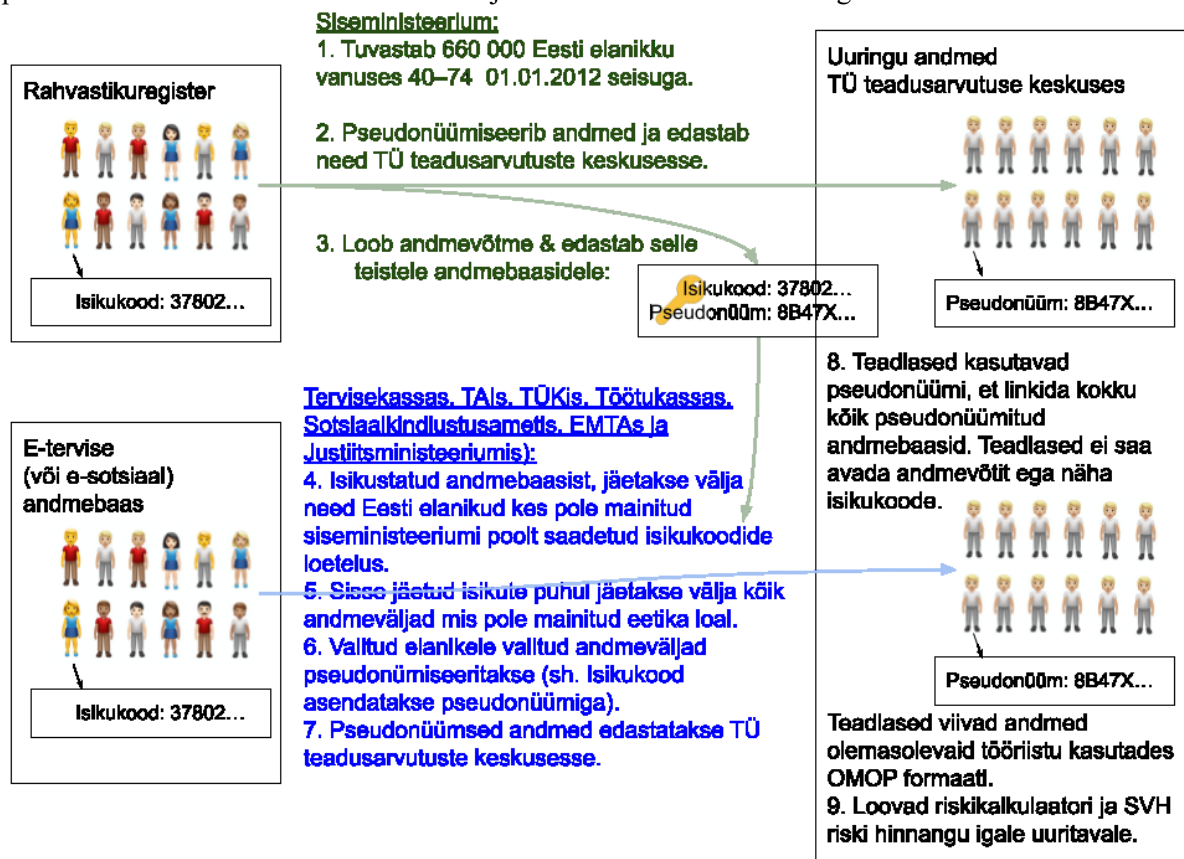
Siseministerium loob eraldi faili ehk andmevõtme. See sisaldab loetelu valimis olevate inimeste isikukoodidest ja pseudonüümidest. Andmevõti edastatakse krüpteerituna Tartu Ülikooli uurimismeeskonnale. Krüpteering on aga teostatud nii, et antud tabelit saavad dekrüpteerida ainult **seitse** nimetatud isikut üheksast riigiasutusest keda plaanime kaasata käesolevasse uuringusse ja jätku-uuringutesse (st Eesti Töötukassa, Maksu- ja Tolliamet, Sotsiaalkindlustusamet, Justiitsministerium, Tartu Ülikooli genoomika instituut, Tervisekassa, TEHIK, TAI, Tartu Ülikooli kliinikum). Peale **seitsme** isiku nime kooskõlastamist edastavad uurijad need **seitse** nime siseministeriumile. Keegi

³⁰ Riley RD, et al. Calculating the sample size required for developing a clinical prediction model. BMJ. 2020;368.

³¹ Peduzzi P, Concato J, Kemper E, Holford TR, Feinstein AR. A simulation study of the number of events per variable in logistic regression analysis. Journal of clinical epidemiology. 1996 Dec 1;49(12):1373-9.
<https://pubmed.ncbi.nlm.nih.gov/8970487/>

teine (sh antud taotluses loetletud uurijad ja muu personal TÕ teadusarvutuste keskus) andmevõtit dekrüpteerida ja avada ei saa. Seega ainsad inimesed, kes näevad ja töötlevad isikukode, on **seitse** andmebaasi omaniku esindajat.

Uurimismeeskond täidab seejärel sekretäri funktsiooni ja edastab andmevõtme teiste andmekogude omanikele ja volitatud andmetöötajatele. Seeläbi koordineerib Tartu Ülikooli uurimismeeskond andmete väljastamist erinevatest allikatest samade pseudonüümidega, ilma ise andmevõtit avamata. Andmevõtme alusel tuvastavad teised andmekogude omanikud, kes on uuringu valimis ja kes mitte. Teistest andmekogudest edastatakse vajalikud andmeväljad pseudonümiseeritud kujul (sh koos sotsiaalministeeriumi poolt edastatud pseudonüümiga, aga ilma isikukoodideta) Tartu Ülikooli teadusarvutuste keskusesse. Tartu Ülikooli uurijad kasutavad pseudonüümi, et linkida kokku kaks pseudonümiseeritud andmebaasi. TÕ uurijad ei saa avada andmevõtit ega näha isikukode.



Joonis 3. Ülevaade andmete pseudonümiseerimise ja linkimise protsessist, esimese kohordi (01.01.2012 andmete) näitel. Rahvastikuregistri tegevused on rohelisega. E-tervise- ja e-sotsiaalandmebaaside tegevused on sinisega. TÕ teadusarvutuse keskuse tegevused on musta fondiga.

7.4. Andmete analüüs Tartu Ülikoolis

Uurijad analüüsivad pseudonüümitud kujul ligikaudu 660 000 residenti. Käesolevas uuringus puudub informatsioon Tervise infosüsteemist ja geenivaramu andmebaasist.

SVH riski hindamiseks peavad iga kohordi liikmed jälgimisaja alguses olema SVH vabad (nn terved). TÕ teadusarvutuse keskkuses määravad uurijad residendi tunnuseid uuringu algul (01.01.2012), kasutades andmeid, mis pärinevad 2 aastat enne uuringu algust ja 1 aasta peale uuringu algust (st kohort-1 näitel alates 01.01.2010 kuni 31.12.2012). Nende andmete alusel eemaldatakse uuritavad, kelle on kaasnev SVH haigus (RHK-10 koodid: I20–I25, I63–I64 või G45) või kes kasutavad statiine läbilõike kuupäeva lähedal (samuti -2 aastat kuni +1 aasta).

Esimeses jälgimisaja kohordis (2012-2021) luuakse üks Cox proportional hazards mudel mis ennustab SVH surma või esmast mitte-surmavat juhtu. Mudelisse lisatakse kõik usutavad riskitegurid korraga. Oluline on lisada ka e-tervise tegurid, mis kirjanduse alusel ennustavad SVHd (nt QRISK3 kalkulaatori tegurid). Antud tegureid peetakse patsientide ja arstide silmis kõige legitiimsemateks. E-tervise tegurite mitte lisamine võib tekitada patsientide või arstide vaates mulje, nagu ainult sotsiaalteguritel põhinev hinnang ei ole piisavalt usutav või valideeritud (sõltumatu tema tehnilisest täpsusest).

Kõige vähem ennustustavad tegurid eemaldatakse ükshaaval kuni viimane alles oleva ennustava riskiteguri Hazard Ratio on kas >1.1 või tema P-väärtus on <0.05 . Kui ühes andmebaasis pole ühtegi ennustavat riskitegurit, siis see andmebaas välistatakse jätku-uuringutest. Kordame identseid analüüse teises (2014-2023) ja kolmandas (2016-2025) jälgimisaja kohordis. Ajalisel sisemisel valideerimisel katsetame, kui hästi see esimese ja teise kohordi riskikalkulaatorid ennustavad suremust teisel ja kolmandal kohordil. Peamine ennustusvõime tunnus on *C-statistic* nagu ka eelmistes uuringutes.¹⁰

8. UURIMISTÖÖS KOGUTAVATE ISIKUANDMETE KAITSE KIRJELDUS

8.1 Andmete jagamise õiguslik alus ja selle analüüs

Alloleva koostamisel saime sisendit nii Tartu Ülikooli kantselei juristi Eve Pehteri, kui ka vandeadvokaat Maret Kruusa käest. Õiguslik aluse hinnangul alustame Euroopa üldmäärusest, siis vaatame Eesti seadus, siis spetsiifiliste registrite põhimäärusi.

Esiteks, [Euroopa isikuandmete kaitse üldmäärus \(2016/679\)](#) 6. artikkel:

1. Isikuandmete töötlemine on seaduslik ainult juhul, kui on täidetud vähemalt üks järgmistest tingimustest, ning sellisel määral, nagu see tingimus on täidetud:

Leiame, et antud uuringu puhul on täidetud kaks tingimust:

- e) *isikuandmete töötlemine on vajalik avalikes huvides oleva ülesande täitmiseks* või vastutava töötleja avaliku võimu teostamiseks“
 - Jah. Tartu Ülikooli avalikes huvides olevad ülesanded on defineeritud [Tartu Ülikooli seaduses § 2 lg 3](#): „Ülikool on universaalne integreeritud teadus-, arendus-, õppe- ja kultuuriasutus. Ülikooli eesmärk on edendada teadusi ja kultuuri, pakkuda teaduse ja tehnoloogia arengule tuginevaid kõrghariduse omandamise võimalusi kõrgharidustaseme kolmel astmel humanitaar-, sotsiaal-, arsti- ning loodusteaduste alal ning osutada õppe-, teadus- ja muul loometegevusel põhinevaid avalikke teenuseid.“ Antud projekt on läbinud tiheda ETAGi PSG konkurssi, et tuvastada milliseid teadustegevusi peaks ülikool täitma. Selle käigus sai projekt meditsiinivaldkonnas kõige kõrgema punktiskoori, osutades projekti olulisusele ja vajalikkusel täitmaks Tartu Ülikooli avalikku huvi.
- „f) *isikuandmete töötlemine on vajalik vastutava töötleja või kolmanda isiku õigustatud huvi korral, välja arvatud juhul, kui sellise huvi kaaluvad üles andmesubjekti huvid või põhiõigused ja -vabadused, mille nimel tuleb kaitsta isikuandmeid.*“
 - Jah. Leiame, et Eesti elanike õigustatud huvi on antud uuringu puhul suurem kui sellega võrreldes väiksem riive andmesubjekti õiguste piiramisel (st nende mitte teavitamine seaduspärasest andmete töötlemisest). Seda analüüsime hiljem toetudes Eesti seadusele.

Teiseks, [Euroopa isikuandmete kaitse üldmäärus \(2016/679\)](#) 9. artikkel:

„Isikuandmete eriliikide töötlemine

2. [Lubatud on töödelda isikuandmeid, millest ilmneb etniline päritolu, geneetilisi andmeid, terviseandmeid] kui kehtib üks järgmistest asjaoludest:

- j) *töötlemine on vajalik avalikes huvides toimuva teadusuuringute eesmärgil vastavalt artikli 83 lõikele 1, tuginedes liidu või liikmesriigi õigusele, ning on proportsionaalne saavutatava eesmärgiga, austab isikuandmete kaitse õiguse olemust ning tagatud on sobivad ja konkreetset meetmed andmesubjekti põhiõiguste ja huvide kaitseks.*“
 - Jah. Terviseandmeid töödeldakse avalikes huvides toimuva teadusuuringu eesmärgil, on proportsionaalne, austab isikuandmekaitse õigusi ja tagab sobivad meetmed kooskõlas nii Euroopa üldmääruse kui ka Eesti allolevate õigustega.

Kolmandaks, Eesti [Isikuandmete kaitse seadus](#) § 6:

*„(1) Isikuandmeid võib andmesubjekti **nõusolekuta** teadusuuringu vajadusteks töödelda eelkõige **pseudonüümitud** kujul. Enne isikuandmete üleandmist teadusuuringu töötlemiseks asendatakse isikuandmed pseudonüümitud andmetega.“*

- Jah. Enne isikuandmete üleandmist teadusuuringu töötlemiseks andmesubjekti nõusolekuta vastutavale töötlejale (Tartu Ülikoolile), asendavad andmete omanikud isikuandmed pseudonüümidega.

*„(2) **Depseudonüümimine** või muu viis, millega isikut mittetuvastavad andmed muudetakse uuesti isikut tuvastavaks, on lubatud ainult täiendavate teadusuuringute vajadusteks. Isikuandmete töötleja määrab nimeliselt isiku, kellel on juurdepääs depseudonüümimist võimaldavatele andmetele.“*

- Jah. Käesolevas uuringus kohaldub see vaid ühel juhul:
[Töötute isikute registri põhimääruse](#) § 37 sätestab, et „(1) Registrisse kantud andmeid säilitatakse kuni kümme aastat pärast isiku suhtes alustatud tööturutoetuste menetluse lõppemist. (2) Säilitustähtaja möödumisel registrisse kantud andmed pseudonüümitakse kord aastas säilitustähtaja möödumise kalendriaasta lõpus.“
Antud uuringu raames vajame töötute isikute toetuste summasid aastast 2008. Andmete väljastamise hetkel on esimeste aastate andmed (01.01.2008-31.12.2012) juba pseudonüümitud. Töötukassa analüüsi osakonnajuhataja Margit Paulus kinnitas (22.09.2023), et nende nimetatud isik on valmis ja nõus antud andmeid uuringu tarbeks depseudonüümima. Töötukassa nimetatud isik teeb seda peale Tartu Ülikooli inimuuringute eetika komitee loa saamist. Seejärel väljastab Töötukassa nimetatud isik vastutavale töötlejale (Tartu Ülikoolile) vaid pseudonüümituid andmeid.

(Selguse huvides täpsustame, et tulevikus võime küsida, et geenivaramu nimetatud isik depseudonüümida geenidonorite andmeid jätku-uuringutes teostamiseks. Selle protsessi pädevust ja kooskõla ülaltoodud ja teiste seadustega hindab EBIN.)

*„(3) Teadusuuringu andmesubjekti **nõusolekuta** tema kohta käivate andmete **töötlemine andmesubjekti tuvastamist võimaldaval kujul** on lubatud üksnes juhul, kui on täidetud järgmised tingimused“*

- Jah. Teadusuuringu tegija hinnangul on terviseandmed eriliiki isikuandmed, mille spetsiifilisus ja unikaalsus võib teha võimalikuks, et piisavalt pädev ja võimekas isik võib andmesubjekte sealt tuvastada. Tulenevalt:

*„1) pärast tuvastamist võimaldavate andmete eemaldamist ei ole andmetöötluse eesmärgid enam saavutatavad või neid oleks **ebamõistlikult raske** saavutada“*

- Jah. Kui uuring ei kasuta rahvastikuregistrit, et saa me tuvastada valimit.
Kui uuring ei kasuta surma põhjuste- või müokardiinfarkti-registrit, et saa teadusuuringu tegijad tuvastada haigusteket.
Tulenevalt muutub võimalikuks uurigu peamine eesmärk (kontrollida kas sotsiaalsed riskitegurid ennustavad SVH).

„2) teadusuuringu tegija hinnangul on selleks **ülekaalukas avalik huvi**“

- Jah, andmetöötluseks on kahte sorti ülekaalukat avalikku huvi: haiguse ennetamine ja ebavõrdsuse vähendamine.
- Esimene avalik huvi on **südamehaiguste ennetamine** keskmise Eesti elaniku seas. SVH on Eestis **juhtiv** haiglaravi põhjus³² ja juhtiv surmapõhjus³³. Täpsemini, aastal 2022 suri vereringeelundite haigustesse (I00-I99) 8102 inimest millest pea pooled (3126 surma) olid haigused kus on eriti tugev teaduspõhisisus statiin ravimite tõhususe kohta: südame isheemiatõved (I20-I25) ja peaajuveresoonte haigused (I60-I69).^{34,35} Projekt loob vundamendi tuleviku teenustele, mis proovib seda juhtivat haigusteket ennustada ja ennetada. Käesoleva projekti ambitsioon seega kooskõlas käesoleva Vabariigi valitsuse tegevusprogrammiga³⁶ („9.1.7. Arendame proaktiivseid sotsiaal- ja tervishoiuteenuseid, mille puhul riik pakub teenust **vajaduspõhiselt andmete alusel**.“ ; „9.3.2. Uuendame tervishoiu **digilahendusi** ja arendame personaalset terviselugu millega inimene näeb terviseinfot [ja] mille kaudu on võimalik saada tervisenõustamist. Eesmärk on suunata senisest rohkem tähelepanu ja ressursse haiguste **ennetamisele**.“). Käesoleva projekti ambitsioon on muuhulgas kooskõlas ka Sotsiaalministeeriumi heaolu arengukavaga, mille alusel peaks „*Valdkondadeülest koostöötet toetavaid infosüsteeme arendama*“ [seeläbi, et] „*Rakendada tervise- ja sotsiaalsektoris abivajaduse tuvastamiseks sihtrühmapõhiselt ning algoritmidel põhinevat ühist riskihindamist*“³⁷)
- Teine avalik huvi on **tervise ebavõrdsuse langetamine**. Eestis on võrreldes Euroopa keskmisega erakordselt suured sotsiaalmajanduslikud ebavõrdsused tervisenäitajate seas. Kõrgharidusega mehed elavad Eestis 10,6 aastat kauem kui põhiharidusega mehed.³⁸ Seda vahet mõjutab SVH (sh infarkt ja insult) rohkem kui teised haigused.³⁹ Antud projekt on üks vähesi Eestis, mis proovib luua personaalseid teenuseid, et seda vahet vähendada. Ambitsioon on muuhulgas kooskõlas Eesti rahvastiku tervise arengukavaga.⁴⁰ Kava seab kolm eesmärki, millest üks on „*Ebavõrdsus tervises (sugude, piirkondade ja haridustasemetel vahel) väheneb vähemalt selliselt, et 2030. aastaks ei ole eluiga üheski maakonnas Eesti keskmisest lühem kui kaks aastat ning põhiharidusega inimeste keskmine oodatav eluiga ei jää kõrgharidusega inimeste keskmisest oodatavast elueast maha rohkem kui kaheksa eluaastat*.“ Prioriteet illustreerib ebavõrdsuse tähtsust Eesti riigi tervisepoliitikas. Lisaks, kuna iga „*sekkumiste planeerimisel tuleb arvestada, et need ei suurendaks olemasolevaid lõhesid, vaid pigem aitaks neid vähendada*“ peavad uute teenuste arendajad (sh. ülikoolid) planeerima sisse ebavõrdsust langetavaid lahendusi mida käesolev projekt teeb.

³² TAI. Aktiivravihaiglast lahkunud ravi põhjuse, vanuserühma ja patsiendi elukoha maakonna järgi. 2021.

https://statistika.tai.ee/pxweb/et/Andmebaas/Andmebaas_08TerviseInfosysteem_01TerviseInfosysteem/PH5.px/

³³ TAI. Surmad põhjuse, soo ja vanuserühma järgi. 2022.

https://statistika.tai.ee/pxweb/et/Andmebaas/Andmebaas_01Rahvastik_04Surmad/SD21.px/

³⁴ Collins R. et al. Interpretation of the evidence for the efficacy and safety of statin therapy. The Lancet. 2016.

Yusuf S, et al. Cholesterol lowering in intermediate-risk persons without cardiovascular disease. New England Journal of Medicine. 2016

³⁵ Mihaylova B. The effects of lowering LDL cholesterol with statin therapy in people at low risk of vascular disease: meta-analysis of individual data from 27 randomised trials. The Lancet. 2012.

³⁶ Vabariigi Valitsus. Tegevusprogramm 2023–2027. <https://valitsus.prelive.vportal.ee/media/6186/download>

³⁷ Sotsiaalministeerium. Heaolu arengukava 2023-2030. 2023. <https://www.sm.ee/heaolu-arengukava-2023-2030>

³⁸ Statistikaamet. Elada jäänud aastad soo ja hariduse järgi 2021. RV0454. <https://www.stat.ee/et/avasta-statistikat/valdkonnad/heaolu/tervis/oodatav-eluiga>

³⁹ Currie J, et al. Life expectancy inequalities in Wales before COVID-19: an exploration of current contributions by age and cause of death and changes between 2002 and 2018. Public Health. 2021;193:48-56.

⁴⁰ Sotsiaalministeerium. Rahvastiku tervise arengukava 2020-2030. 2021. <https://www.sm.ee/rahvastiku-tervise-arengukava-2020-2030>

Täpsuse huvides põhjendame, miks **sotsiaalsed riskitegurid on legitiimsed** osad ülaltoodud ülekaalukat avalikku huvi saavutamiseks. Antud teadusuuringu käigus töötab Tartu Ülikool välja uue „*haiguse varase avastamise programmi ja uurib riskirühme*“ ([Rahvatervise seadus §5\(5\)](#)). Haiguse avastamise programmid on täpsemad (st eesmärgipärasemad) kui nad kasutavad võimalikult palju teadaolevaid riskitegureid ehk tervise mõjureid. Rahvatervise seaduse §2(10) „**tervisemõjuri**“ definitsioon sisaldab muuhulgas ka „*sotsiaal-majanduslike, käitumuslike ja psühhosotsiaalseid tegureid, mis võivad omada rahvastiku tervisele soovitavat või ebasoovitavat mõju*“. Kirjanduse ülevaates (taotluse punkt 4.1.2.) on näidatud, kuidas kõik uuritavad riskitegurid võivad omada SVH tekkele piisavalt suurt ja usutavat mõju. Antud definitsioon on kooskõlas [Inimgeeniuringute seaduse §2\(5\)](#) „tervise seisundi kirjelduse“ definitsioon, mis sisaldab muuhulgas „eluviisi, **elukeskkonda** ja pärilikke omadusi“. Rahvatervise seaduse §2(7) omakorda defineerib elukeskkonda kui „inimesega kokkupuutuva loodusliku, tehniliku ja **sotsiaalse** keskkonna tegurite kogum, mis mõjutab või võib mõjutada inimese tervist“. Muuhulgas sisaldavad enamik sotsiaalsed tegurid (nt sissetulek ja vara) nii geneetilise pärilikkuse kui ka keskkondliku pärilikkuse komponente.⁴¹ Antud definitsioon on kooskõlas [Euroopa isikuandmete kaitse üldmääruse sissejuhatuse](#) (lõik 157) kus illustreeritakse proportsionaalset ja kasulikku andmete töötlemist. Näide sisaldab üllatavalt palju sisu, mis kattub käesoleva projekti sisuga: „*Registritest saadava teabe sidumise teel võivad teadlased saada uusi väärtuslikke teadmisi laialt levinud tervise seisundite kohta nagu **südam-veresoonehaigused**. Registrate alusel saab uuringutulemusi parendada, kuna need saadakse **suuremast valimist**. Sotsiaalteaduste valdkonnas võimaldavad registritel põhinevad teadusuuringud teadlastel saada olulisi teadmisi paljude sotsiaalsete tingimuste nagu näiteks **töö** ja **haridus**. Registrate alusel saadud uuringutulemused annavad **usaldusväärseid ja kvaliteetseid teadmisi**, mis võivad parandada paljude inimeste elukvaliteeti.*“

„3) töödeldavate isikuandmete põhjal ei muudeta **andmesubjekti kohustuste** mahtu ega kahjustata muul viisil ülemäära andmesubjekti **õigusi**.“

- Antud uuringu raames, kus andmesubjekte ei teavitata uuringu tulemustest või väljunditest, ei teki andmesubjektidele uusi kohustusi. Andmesubjektide õigused jäävad peamiselt muutumatuks, st. säilib õigus nõuda andmete parandamist ja olla unustatud (vt lisatud Isikuandmete kaitse mõjuhinnang v6.doc, osa 3.3.6).

Üks andmesubjekti õigus, mida uuring osaliselt kahjustab on **õigus olla teavitatud**. Teadusuuringu tegija kaalus varianti, kus uuritavaid teavitatakse individuaalselt, et nende andmeid töödeldakse (aga aktiivset opt-in nõusolekut ei küsita). Teadusuuringu tegija otsustas, et selline lahendus eeldaks **ebaproportsionaalseid jõupingutusi**. Arvestades valimi suurt suurust tooks see endaga kaasa nii suure hulga lisatööd, et teadusuuringu tegija puudub inimressurss ja rahalised vahendid seda teostada ja uuring muutuks seejärel teostamatuks.

Kokkuvõtteks, teadusuuringu tegija leiab, et isikuandmeid töödeldakse avalikes huvides tehtud teadusuuringu eesmärgil, seda tehakse pseudonüümitud kujul, uuritavad ei saa ühtegi päriselu sekkumist ega otsest mõju ja andmetöötlamise teave tehakse avalikult kättesaadavaks (**täpsem info: taotluse osa 8.2, lk 19**). Seega on andmetöötlus proportsionaalne mis ei kahjusta ülemäära andmesubjekti õigusi. Andmetöötlus on kooskõlas nii ülaltoodud Eesti seaduse kui ka [Euroopa isikuandmete kaitse üldmääruse](#) artikkel 14, lõik 5, punkt b-ga.

„(4) Kui teadusuuring põhineb **eriliiki** isikuandmetel, siis kontrollib asjaomase **valdkonna eetikakomitee** enne käesolevas paragrahvis sätestatud tingimuste täitmist. Kui teadusvaldkonnas puudub eetikakomitee, siis kontrollib nõuete täitmist Andmekaitse Inspeksioon.“

⁴¹ Benjamin DJ, Cesarini D, Chabris CF, Glaeser EL, Laibson DI, Age, Gene/Environment Susceptibility-Reykjavik Study, Guðnason V, Harris TB, Launer LJ, Purcell S, Smith AV. The promises and pitfalls of genoecconomics. Annu. Rev. Econ. 2012 Sep 26;4(1):627-62.

- Käesolevas uuringus töödeldakse uut projektipõhist andmebaasi mis on kokku pandud kasutades nii eriliiki kui ka mitte-eriliiku isikuandmeid. **Esitame eetikakomiteele teavet kõigi nende andmete kohta, et eetikakomiteele saaks hinnata:**
 - o Kas **eriliiki** andmete jagamine on kooskõlas Isikuandmete kaitse seadus §6(4)ga, ja
 - o Kas **projektiüleselt** (arvestades nii eriliiki kui mitte-eriliiki andmeid) on:
 - piisav **avalik huvi** ja
 - kas **privaatsuse riive** on proportsionaalne.

Vastav projektipõhine koond-andmebaas allub teadusuuringu tegija hinnangul pigem „eriliiki“ andmekaitse nõuetele ja seadusele. Antud uuringu asjaomases teadusvaldkonnas (meditsiin ja tervis) eksisteerivad kaks sobilikku komiteed: TAI inimuuringute eetikakomitee ja TÜ inimuuringute eetikakomitee. Seega pole Andmekaitse Inspektsiooni selle tingimuse täitmiseks vaja kaasata. Kuna Teadusuuringu tegijal puudub huvide konflikt TÜ inimuuringute Eetikakomitee liikmete vahel, eelistab Teadusuuringu tegija, et TÜ inimuuringute Eetikakomitee kontrolliks taotluse kooskõla ülaltoodud seadusega.

*„(5) Käesoleva seaduse tähenduses loetakse teadusuuringuks ka **täidesaatva riigivõimu analüüsid** ja uuringud, mis tehakse poliitika kujundamise eesmärgil. Nende koostamiseks on täidesaatval riigivõimul õigus teha päringuid teise vastutava või volitatud töötleja andmekogusse ning töödelda saadud isikuandmeid. Andmekaitse Inspektsioon kontrollib enne nimetatud isikuandmete töötlemise algust käesolevas paragrahvis sätestatud tingimuste täitmist, välja arvatud juhul, kui poliitika kujundamiseks tehtava uuringu eesmärgid ja isikuandmete töötlemise ulatus tulenevad õigusaktist.“*

- Ei kohaldu. Antud teadusuuringut teostatakse ainult ülikoolis, mitte täidesaatva riigivõimu asutuses.

*„(6) Kui isikuandmeid töödeldakse **teadusuuringu** eesmärgil, võib vastutav või volitatud töötleja Euroopa Parlamendi ja nõukogu määruse (EL) 2016/679 artiklites 15, 16, 18 ja 21 sätestatud andmesubjekti **õigusi piirata** niivõrd, kuivõrd nende **õiguste teostamine** tõenäoliselt muudab **võimatuks** teadus- või ajaloouuringu või riikliku statistika eesmärgi saavutamise või **takistab seda oluliselt**.*

Artikkel 15 käsitleb õigus tutvuda andmetega ja saada teavet. Teadusuuring ei piira õigust tutvuda andmetega, juhul kui andmesubjekt sellise päringu esitab. Küll piirab teadusuuring osaliselt andmesubjektide õigusi saada teavet, tehes seda proportsionaalselt ja seaduslikult. Vt ülaltoodud analüüs punktile „(3) töödeldavate isikuandmete põhjal ei muudeta andmesubjekti kohustuste mahtu ega kahjustata muul viisil ülemäära andmesubjekti õigusi.“

Artikkel 16, 18 ja 21 käsitlevad andmesubjekti õigusi andmete parandamisele, töötlemise piiramisele ja õigusele esitada vastuväiteid. Teadusuuring ei piira ülaltoodud õigusi.

Neljandaks, mõningate registrite puhul on määrused täpsustanud teadustööde tegemise otstarvet ja korda. Näiteks, [töötuna ja töötajana arvel olevate isikute ning tööturuteenuste osutamise registri asutamise ja pidamise põhimääruse](#) § 1 (3) „Registri pidamise eesmärgid on [...] 5) statistika ja **analüüside tegemiseks vajalike andmete kogumine.**“

„§ 35. (1) „**Juurdepääs** registrisse kantud andmetele võimaldatakse kooskõlas **Euroopa Parlamendi ja nõukogu määrusega (EL) 2016/679** füüsiliste isikute kaitse kohta isikuandmete töötlemisel ja selliste andmete vaba liikumise kohta (isikuandmete kaitse üldmäärus) ja teiste kehtivate õigusaktidega.“

- Jah. Käesoleval uuringus analüüsitakse ülaltoodud registri andmeid, kooskõlas Euroopa üldmäärusega.

[„Sotsiaalkaitse infosüsteemi põhimäärus“](#) lõpetas kehtivuse 2019.a. Seejärel reguleerib SKAIS2 andmete kasutamist eelmainitud Euroopa üldmäärus ja Eesti isikuandmete töötlemise seadus.

Sarnaselt ülaltoodule, ka [Maksukohustuslaste registri põhimäärus \(mis reguleerib nii tulumaksu- kui töötamise registri andmeid\)](#) § 61 (2) alusel „Registri andmetele võimaldatakse juurdepääs kooskõlas **Euroopa Parlamendi ja nõukogu määruse (EL) 2016/679** füüsiliste isikute kaitse kohta isikuandmete töötlemisel ja selliste andmete vaba liikumise, avaliku teabe seaduse ja isikuandmete kaitse seadusega. **Juurdepääs võimaldatakse tolliseadustikus, maksukorralduse seaduses või muus seaduses sätestatud korras ja ulatuses.**“

- Jah. Uurijad taotlevad juurdepääsu muus seaduses sätestatud korras ja ulatuses, nimelt eelmainitud Euroopa üldmääruse ja Eesti isikuandmete töötlemise seaduse korras ja ulatuses.

„(3) Register on piiratud juurdepääsuga ametialaseks kasutamiseks, välja arvatud **õigusaktiga sätestatud erandjuhud**. Ametialase kasutamise eesmärgil on registrile juurdepääsu õigus vaid selgelt kindlaks määratud teenistus- või tööülesande täitmiseks.“

- Jah. Uurijad taotlevad juurdepääsu ülalmainitud Eesti isikuandmete töötlemise seaduses sätestatud erandjuhu alusel.

„(4) Registri kasutamise õigus väljaspool Maksu- ja Tolliametit ametialase kasutamise eesmärgil eeldab andmete töötlemise **lepingu sõlmimist** Maksu- ja Tolliametiga. Lepingus kajastatakse vähemalt andmete kasutamise eesmärk, väljastatavate andmete loetelu, maht ja väljastamise kord, andmetele juurdepääsu tehnilised lahendused ning meetmed andmete kaitsmiseks.“

- Jah, uurijad on nõus sõlmima vastavat lepingut Maksu- ja Tolliametiga.

Rõhutame, et eetikakomitee ei hinda, kas mitte-eriliiki isikuandmeid plaanitakse töödelda seaduslikult või mitte (vt lk 22).

Kinnistusraamatuseaduse § 1 (1) alusel „Kinnistusraamat on andmekogu, mille eesmärk on koguda, säilitada ja **avalikustada teavet kinnisomandi tekkimise, üleandmise** ja kinnisasja asjaõigusega koormamise, samuti kinnisasja koormava asjaõiguse üleandmise, koormamise, selle sisu muutmise või lõpetamise kohta.

- Jah. Uuringu raames soovivad uurijad kasutada avaliku teavet kinnisomandi omamise kohta.

„Äriregistri seaduse § 2 alusel: (1) Äriregister on riigi infosüsteemi kuuluv andmekogu, mille pidamise eesmärk on koguda, säilitada ja **avalikustada teavet Eesti õiguse järgi asutatud eraõiguslike juriidiliste isikute** [...] kohta. (2) E-äriregister on veebiplatvorm, mille kaudu saab [...] avalikustatud andmetega tutvuda. E-äriregistris võib peale äriregistri andmete avalikustada ka teistest andmekogudest ja teiste teenuste kaudu kättesaadavaid avalikke andmeid.“

- Jah. Uuringu raames soovivad uurijad kasutada avaliku teavet eraõiguslike juriidiliste isikute kohta.

8.2. Uuritavate ehk andmesubjektide teavitamine

Pseudonüümseid uuritavaid ei teavitata personaalselt nende õigustest ja riskidest, mida andmete töötlemine võib kaasa tuua. Pseudonüümsete isikuandmete töötlemise puhul informeerime uuritavaid (kui andmesubjekte) seeläbi, et avaldame uuringu meetodi ja andmekaitse aspekte aadressil <https://tervis.ut.ee/et/sisu/teadus> (vt Lisa 10.3). Lisaks tutvustame uuringut kui näidet teistel avalikkusele suunatud üritustel ja jagame teavet meedias, mh:

- Arenguseire Keskuse andmekonverents poliitikutele ja ametnikele 30.11.2022;
- Arvamusfestival (eeldatavasti 2024, 2026, 2028).
- Veebimeedias ja trükkajakirjades artiklite avaldamine (eeskätt ERR Novaator, kuid ka Postimees ning Päevaleht, kus vastutav uurija on varem artikleid avaldanud).
- Raadiosaated (eeskätt Kuku Raadio Kukuv õun, seejärel Vikerraadio).
- Podcastid (eeskätt Levila).
- Sotsiaalmeedia (eeskätt ülikooli ametlikud Facebook kanalid, aga ka Vastutava uurijate ja teiste uurijate erapostitused. Eeskätt eestikeelses Facebookis ning inglisekeelses Twitteris).
- Proovida ka televisiooni (eeskätt Terevisioon, aga lisaks ka venekeelset publikut läbi ETV+ kus vastutav uurija on varem rääkinud).
- Koostöös TÜ Turundus- ja kommunikatsiooniosakonnaga mõtleme ka sihitud teavitustegevusi nii eestikeelt rääkivatele kui ka mitte-eestikeelt rääkivatele uuritavatele.

8.3 Uurijad kasutavad vaid pseudonüümseid andmeid mis ei võimalda isiku tuvastamist.

Isikustatud andmeid (sh isikukode) töötlevad vaid andmete omanikud (st kas vastutavad töötlejad nagu ministeeriumid ja riigiametid või nende poolt volitatud töötlejad nagu TEHIK või TAI). Seda tehakse seepärast, et ülikoolile edastatavatest pseudonüümsetest andmetest eemaldada isikud, kes ei kuulu analüüsitava valimisse (ehk n-ö mittekuuluvad isikud). See samm aitab kaitsta andmete minimeerimise printsiipi.

Teostamiseks on vaja, et üks valimi koostaja looks andmevõtme tabeli valimisse kuuluvate isikukoodide ja vastavate pseudonüümidega. See esmane asutus edastab andmevõtme kõigile teiste andmebaasi omanikele. Praktilistel kaalutlustel on lihtsam, kui andmevõtme edastamisega tegeleb uurimismeeskond. Et aga uurimismeeskond ei pääseks ise andmevõtmele ligi, krüpteerib valimi koostaja andmevõtme nii, et seda saab avada vaid nimetatud andmebaasi nimetatud töötaja. See samm tõstab turvalisust.

Teoreetiliselt on isiku tuvastamine risk siiski võimalik aga selle tõenäosus on äärmiselt väike.

8.4. Andmebaaside omanike huvi või nõusolek projekti raames andmeid jagama

Oleme alates 09.2022 suhelnud järgmiste omanikega ja saanud esialgseid vastuseid:

SA TÜ Kliinikumi Müokardiinfarktregistri direktor Gudrun Veldre:

„Arvestada tuleks sellega, et kõigi isikustatud andmepäringute väljastamise üle teeb otsuse ja annab loa registri vastutav töötleja Sotsiaalministeerium. See võtab kindlasti aega. Infopäringu vormi, mille saate täita kui olete saanud inimuuringute eetikakomiteelt loa, leiate meie kodulehelt www.infarkt.ee“

Sotsiaalkindlustusameti andmekaitse spetsialist Ainike Alla: “Kinnitan, et kui eetiline ja/või Andmekaitse Inspektsiooni koostööst on positiivne, siis ei sea ka Sotsiaalkindlustusamet takistusi andmete väljastamiseks.”

Töötukassa analüüsiosakonna juhataja Margit Paulus: “Kinnitan, et kui eetiline ja Andmekaitse Inspektsiooni koostööst on positiivne, siis saame andmeid väljastada alates juunist mil valmib (de)pseudonüümimise tehniline lahendus.”

Justiitsministeeriumi nõunik Kadri Laud: „Küsitud kinnistusraamatu ja äriregistri andmeid saame pseudonüümiseerida ja Teile edastada.“

TEHIKu Talituse juhataja Terje Lasn ja Sotsiaalministeeriumi TIS vastutav töötaja Raili Sillart kinnitasid, et projekt on neile oluline. Nad eelistasid, et andmevahetus toimub läbi Statistikaameti. Statistikaameti vanemkonsultandid Erika Orav ja Liis Kivirand: “Puhastatud pseudonüümitud andmehulk on suurem kui nende serveri andmelagi (20GB) ja seega Statistikaamet ei saa pakkuda andmevahetus/pseudonüümimise teenust.” **Plaanime peale eetilist kooskõlastamist küsida TEHIKult kas nad on nõus andmeid Tartu Ülikooliga sel viisil jagama.**

EMTA on varem sarnaseid andmeid jaganud sarnaste projektide tarbeks. Käesoleval hetkel puudub neil tehniline võimekus andmeid jagada. Plaanime koos teiste asutustega seda üheskoos arutada, et EMTA tehnilist võimekust tõsta.

Selguse huvides, jätku-uuringute teostamiseks on meil TÜ Eesti Geenivaramu Teaduskomiteele luba soovi korral edasi liikuda EBINile jätku-taotlusega.

8.5. Kuidas andmed turvaliselt transporditakse uurijateni

Andmete edastamiseks sõlmitakse leping iga andmekogu omaniku ja Tartu Ülikooli vahel. Andmete edastus toimub krüpteeritult läbi turvalise SFTP serveri, mis on üles seatud TÜ teadusarvutuste keskuse poolt.

Analüüsiks luuakse TÜ teadusarvutuste keskuse poolt turvaline keskkond, kuhu on piiratud ligipääs vaid EBINi taotlusel loetletud uurijatele. Keskkonnas on olemas kogu vajalik analüüsitarkvara, seega puudub uurijatel vajadus töö tegemiseks andmeid alla laadida. E-terviseandmed teisaldatakse OMOP andmekujule (<https://github.com/OHDSI/CommonDataModel/wiki/Standardized-Clinical-Data-Tables>), mida saab teostada vaid TÜ teadusarvutuste keskuse pilvelahenduses. Tartu Ülikooli teadusarvutuste keskus vastab ISKE M taseme nõuetele. Seoses 2022. aasta lõpus kehtima hakanud määrusega „Võrgu- ja infosüsteemide küberturvalisuse nõuded“ ja määrusega „Eesti infoturbestandard“ on Tartu Ülikoolil nüüdsest kohustus järgida Eesti Infoturbestandardit ning regulaarselt läbi viia Eesti infoturbestandardi järgimise auditeid. Tartu Ülikool peab esimese auditi läbima vähemalt 3 aasta jooksul (ehk enne 2025. a lõppu) ning käesoleval hetkel selle nimel ka tegutsetakse.

Kõik uurijad allkirjastavad konfidentsiaalsuslepingu ja neid informeeritakse andmetele kohanduvatest andmekaitse nõuetest enne ligipääsu pseudonümiseeritud andmetele. Kõik uurijad töötlevad andmeid vaid uuringu protokollis näidatud eesmärkide saavutamiseks, rakendades vaid uuringu protokollis näidatud meetodeid. Juhul kui uurijad soovivad muuta või lisada eesmärgi, meetodeid või andmete koosseisu, küsime vastavalt olukorrale kooskõlastust kuni kolmelt asutuselt (EBIN, Tartu Ülikooli eetikakomitee, andmekaitse inspeksioon).

8.6. Vastutav töötaja

Vastutav töötaja on Tartu Ülikool (registrikood 74001073, aadress Ülikooli 18, 50090 Tartu).

8.7. Andmete koosseis ja eesmärk

Vt lisa 10.1

8.8. Andmete säilitamise aeg

TÜ uurijad säilitavad pseudonüümitud andmed kuni kuueks aastaks.

Koodivõtit säilitatakse niikaua kuni kõik pseudonüümsed andmed on edukalt Tartu Ülikooli liikunud, aga mitte kauem kui **kuni 31.08.2028**.

Hiljemalt **31.08.2028** on kõik andmed Tartu Ülikoolis kustutatud.

8.9. Andmeid ei edastata välisriikidesse.

9. TEAVE SAMA PROJEKTI VARASEMATEST VÕI SAMAAEGSETEST HINDAMISTEST VÕI KOOSKÕLASTAMISTEST MUJAL

9.1 Käesolev projekt

Lähtuvalt uurijate andmekaitsealasest mõjuhinnangust, et ole käesoleva projekti raames vaja konsulteerida Andmekaitse Inspeksiooniga.

Tartu Ülikooli inimuuringu eetika komiteelt küsime kooskõlastust käesolevate andmebaaside kasutamise kohta.

9.2. Võimalikud jätku projektid

EBINi käest oleme varem küsinud kooskõlastust. Nende tagasiside tulemusena oleme oma taotluse tagasi võtnud. Jätaku uuringutes plaanime EBINi juurde taas pöörduda kui soovime kasutada andmeid Tervise Infosüsteemist ja/või geenivaramust. Juhul kui lisanduvad andmed oluliselt muudavad andmekaitsealasest mõjuhinnangut on võimalik et jätku uuringute raames konsulteerime Andmekaitse Inspeksiooniga.

10. LISAD:

10.1 Andmetabeli struktuur

10.2 Uuringu läbiviijate CVd

10.3 Uuringu meetodi ja andmekaitseaspektid uuritavatele

Eraldi lisatud: TÜ eetika 6_Kaaskiri.doc

Eraldi lisatud: Isikuandmete kaitse mõjuhinnang v6.doc

Lisa 10.1 Andmetabeli struktuur ja koosseis

10.1.1. Kokkuvõtte kasutatavatest andmebaasidest ja eetikakomitee rollist

Registripidaja / Omanik / Vastutav töötaja	Volitatud töötaja	Andmebaasi (lühendatud) nimi	Andmete liigitus: Pseudonümiseeritud ...	Kes kontrollib, kas isikuandmete töötlemise seaduslikud nõuded ja tingimused on täidetud
Tervisekassa		Raviarved	...eriliiki isikuandmed	TÜ IEK, tulenevalt <u>isikuandmekaitse seadusest §6(4)</u>
Tervisekassa		Digiretseptid	...eriliiki isikuandmed	
Tervisekassa		Kindlustatud isikud	...eriliiki isikuandmed	
Sots. ministeerium	TAI	Surma põhjuste register	...eriliiki isikuandmed	
Sots. ministeerium	TÜK	Müokardiinfarktiregister	...eriliiki isikuandmed	
Siseministeerium		Rahvastikuregister	...eriliiki isikuandmed	
Eesti Töötukassa		Töötute isikute register	...tundlikud isikuandmed	Registripidaja ise hindab, otsustab ja selgitab Uurijatele, millise seadusel alusel andmeid tohib väljastada ja kas selle seaduse nõuded ja tingimused on täidetud.
Sots. kindlustusamet		Sotsiaalkaitse infosüsteem	...tundlikud isikuandmed	
Maksu- ja Tolliamet		Maksukohustuslaste register	...tundlikud isikuandmed	
Maksu- ja Tolliamet		Töötamise register	...isikuandmed	
Justiitsministeerium		Kinnistusraamat	...isikuandmed	
Justiitsministeerium		Äriregister	...isikuandmed	Eetikakomitee või AKI kooskõlastust pole ilmtingimata vaja.

Selguse huvides täpsustame, et tulenevalt isikuandmekaitse seadusest §6(4), ei tohi registripidajad jagada **eriliiki** isikuandmeid **ilma** eetikakomitee kooskõlastuseta. Selle saavutamiseks arvestab eetikakomitee **kõiki projekti isikuandmeid** (sh. eriliiki ja mitte-eriliiki andmeid), et hinnata projekti üleselt:

- a) Kas **terves uuringus** eksisteerib või puudub ülekaalukas **avalik huvi**, ja
- b) Kas **terves uuringus andmesubjekti õiguste** riive (sh. privaatsuse riive ja nõusoleku mitte küsimine) on proportsionaalne. Sh. kas privaatsust saaks suurendada viisil mis oluliselt ei takistaks uuringu teostamist?

Juhul kui eetikakomitee kooskõlastab, et terves uuringus eksisteerivad ülaltoodud ja teised eetilised omadused, siis annab see **võimaluse** kõigile registripidajatele andmeid uuringu tarbeks jagada, juhul kui nad seda **soovivad**. Eetikakomitee kooskõlastus **ei kohusta** andmete jagamist. Iga registripidaja (sh. nii eriliiki kui mitte-eriliiki registre pidajad) peab seejärel ise analüüsima, kas eksisteerivad teisi **spetsiifilisemaid seadusi** (väljaspool isikuandmekaitse seaduse §6(4)t), mis täpsemini sõnastavad kas ja millal antud registri andmeid tohib jagada (nt. Inimgeeniuringute seadus, Maksukohustuslaste registri põhimäärus jne) mille kohta eetikakomiteel puudub ülevaade. Kokkuvõtteks, lõplik **vastutus** andmete jagamise **seaduslikkuse** üle lasub **registripidajal**. Eetikakomitee annab soovitavaid hinnanguid vaid isikuandmekaitse seaduse §6(4) ulatuses raames.

10.1.2. Andmevajaduse kokkuvõtlik põhjendus

Tervise andmebaasides leiduvat informatsiooni on vaja, et tuvastada:

- a) südame-veresoonkonna haigestumist iga kohordi algul (01.01.2010–31.12.2016),
- b) südame-veresoonkonna haigestumist jälgimisperioodil (01.01.2012–31.12.2026) ja

Haigestumise informatsioon on kajastatud erinevates tervise andmebaasides. Eelnevale kirjandusele toetudes on alust arvata, et kui ühte tervise andmebaasi mitte kasutada, siis vähendab see tuvastatavate haiguste ja riskitegurite hulka. See omakorda

langetab loodud riskikalkulaatori täpsust (seejuures on võimatu ennustada, kas sellised riskid oleksid üle- või alahinnatud). See omakorda vähendaks statiinravi jõudmist nendeni, kellele seda tegelikult vaja on. Lisaks suurendaks see ravimite kõrvalmõjusid ja ravimitele tehtavaid kulutusi nende seas, kellel ravimit tegelikult vaja pole. Kokkuvõttes langetaks see loodava sõeluuringu teenuse tõhusust (tervise parandamisel) ja kulutõhusust (kasutada maksumaksja ja Terviskassa raha võimalikult efektiivselt ära).

Sotsiaalandmebaasides leitud informatsiooni on vaja, et tuvastada valimis olevate elanike sotsiaalmajanduslikke riskifaktoreid kolme kohordi algushetkedel: 01.01.2012, 01.01.2014 ja 01.01.2016. Väärtuse tuvastamiseks kasutatakse kirjanduses levinud põhimõtet, et tõe tuvastamiseks on vaja vaadata informatsiooni 24 kuud enne sihitud kuupäeva. Seega on e-sotsiaalandmete kasutamise periood 01.01.2010–31.12.2015. Erandjuht on töötu staatus mille puhul küsime andmeid 5 aastat tagasiulatavalt (täpsemalt all põhjendatud).

10.1.3. Detailne andmekoosseis

TERVISEANDMEBAASID

1. Terviskassa raviarved (kuluperioodi aasta järgi vahemikus 01.01.2010–31.12.2026):
 1. Tervishoiuteenuse osutamise koht (EHAK kood)
 2. TTO registrikood ja nimi
 3. Haigla liigi tunnus (kood ja nimetus)
 4. Arve esitanud arsti eriala (kood ja nimetus)
 5. Arve alustatud (raviarve alguskuupäev)
 6. Arve lõpetatud (raviarve lõpukuupäev)
 7. Järgarve (jah/ei)
 8. Arve rahastamise allika kood ja nimetus
 9. Vältimatu abi tunnus
 10. Ravitüüp (statsionaarne, ambulatoorne, perearstiabi jne)
 11. Raviarve lõpetamise põhjuse kood ja nimetus
 12. EMO tunnus (kas teenus on EMO-s tehtud)
 13. Arve diagnoosid:
 1. Diagnoosi kood ja nimetus RHK-10 järgi
 2. Diagnoosi liik (võimalikud väärtused on P – peamine ehk suurimat ravikulu kandev diagnoos; K – täiendavad diagnoosid, mis ei ole välispõhjus või peamine diagnoos ja V – välispõhjus)
 14. Tervishoiuteenused:
 1. Tervishoiuteenuse kood ja nimetus
 2. Kordi
 3. Kogus
 4. Teenuse osutamise kuupäev
 5. Teenusgrupi kood ja nimetus
 15. Protseduurid:
 1. NCSP kood ja nimetus
 2. Protseduuri kuupäev

Selgitus: diagnoosiandmeid on vaja esiteks, et võimalikult täpselt tuvastada südamehaiguste teke. Lisaks tuvastame teised haigused, mis omakorda tõstavad südamehaiguste riski (nt teist tüüpi diabeet, krooniline neerukahjustus, südame rütmihäired nagu kodade virvendusarütmia, migreen, reumatoidartriit, raskemad psühhiaatrilised haigused nagu atüüpiline skisofreenia). Teatud ravikäitumuslikud tunnused (näiteks kui patsient kasutab peamiselt vältimatut abi, aga mitte perearsti abi) annavad informatsiooni uuritava sotsiaalmajandusliku staatuse kohta. See omakorda võib teha täpsemaks kas otseselt südamehaiguste riskiennustuskalkulaatorid või siis kaudselt täpsustades puuduvate andmete väärtuste (nt kolesterooli) ennustamise ja imputeerimise algoritme.

Informatsioon selle kohta, kes haigust diagnoosis (nt kardioloog või perearst) või kus see toimus (kohalik haigla või keskhaigla), mõjutab antud diagnoosi usaldusväärsust.

Ülaltoodud andmete analüüsimata jätmine langetaks riskiennustuskalkulaatori täpsust. Ühelt poolt suurendaks see inimeste hulka, kes saavad statiinravi ilma vajaduseta. Teisalt langetaks see nende inimeste hulka, kes statiinravi ei saa, kuigi neil oleks seda tegelikult vaja.

2. Tervisekassa digiretseptid (vahemikus 01.01.2010–31.12.2026):
 1. Retsepti koostamise aeg
 2. Retsepti kehtivusaeg
 3. Retsepti liik (SR – soodusretsept enne digiresepti, MR – meditsiiniseadme retsept, TR – tavaretsept, T2 – tavaretsept(2x), T3 – tavaretsept(3x), NR – narkootilise ravimi retsept)
 4. Arsti eriala kood
 5. TTO info (kui koostajaks perearst, siis tema arsti kood Dxxxx, muul juhul TTO äriregistri kood)
 6. Raviinfo (kõik väljad)
 1. Diagnoosi kood ja nimetus
 2. Väljakirjutatud ravimi toimeaine (toimeaine 1, toimeaine 2, toimeaine 3) nimetus
 3. Realiseeritud toimeaine (toimeaine 1, toimeaine 2, toimeaine 3) nimetus
 4. Väljakirjutatud ravimi ravimvormi kood ja selgitus
 5. Realiseeritud ravimi ravimvormi kood ja selgitus
 6. Toimeainete 1, 2, 3 sisaldus ja ühik (nt 0,4 mg 1ml)
 7. Realiseeritud toimeainete 1, 2, 3 sisaldus ja ühik (nt 0,4 mg 1ml)
 8. Välja kirjutatud ravimi ATC kood
 9. Realiseeritud ravimi ATC kood
 10. Ekstemporaalse ravimi kirjeldus (vabatekst)
 11. Ravimpreparaadi pakendikood
 12. Realiseeritud ravimpreparaadi pakendikood
 13. Ravimpreparaadi nimi
 14. Väljakirjutatud ravimi koguhulk (ühikute kogus ühes pakendis ja ühe ühiku suurus pakendis)
 15. Realiseeritud ravimi koguhulk (ühikute kogus ühes pakendis ja ühe ühiku suurus pakendis)
 16. Lubatud ravimi koguhulk (meditsiiniseadme limiidikogus) limiidiperioodi jooksul
 17. Meditsiiniseadme kaardi uuendustüüp (01 – ühekordne, 06 – poolaasta, 12 – kalendriaasta, 24 – 2 kalendriaastat)
 18. Arsti selgitus (vabatekst)
 19. Väljakirjutatud ravimpreparaatide arv (pakendite arv)
 20. Realiseeritud ravimpreparaatide arv (pakendite arv)
 21. Annustamine (ravikuuri tüüp, ravikuuri pikkus päevades, tükkide arv, tükkide ühik, kordade arv ja kordade ühik, annustamisskeemi perioodi pikkus).
 22. Retsepti staatuse kood ja nimetus
 23. Annulleerimise põhjus (kood ja seletus)
 7. Väljastaja info (müümise aeg, apteegi kood)
 1. Ravimi väljaostmise kuupäev
 2. Apteegi kood

Selgitus: ravimite andmeid on vaja esiteks, et võimalikult täpselt tuvastada südamehaiguste teket ja seotud haigusi (nt diabeeti). Mõlemal juhul on teatud ravimite kasutamine ainus indikaator, et patsiendil on antud diagnoos (juhul kui arst on diagnoosi jätnud mujale märkimata).

Lisaks on oluline välistada valimist inimesed, kes uuringu algul juba tarvitavad statiini.

Informatsioon selle kohta, kes ravimi välja kirjutas (nt kardioloog või perearst) mõjutab antud ravimiga seotud diagnoosi usaldusväärsust.

Ravimi andmete analüüsimata jätmine langetaks riskiennustuskalkulaatori täpsust. Ühelt poolt suurendaks see inimeste hulka, kes saavad statiinravi ilma vajaduseta. Teisalt langetaks see nende inimeste hulka, kes statiinravi ei saa, kuigi neile oleks seda tegelikult vaja.

3. Ravikindlustuse andmekogu (ehk Tervisekassa kindlustatud isikute register)
 1. Isiku ravikindlustuse alus (kood ja nimetus; nt töandja poolt kindlustatud isik, vanaduspensionär jne)
 2. Isiku ravikindlusega hõlmatusiga perioodid (algus- ja lõpukuupäevad) vahemikus 01.01.2010–31.12.2026
 3. Väljavõte iga isiku ajutise töövõimetuse perioodidest vahemikus 01.01.2010–31.12.2015, sh:
 1. Perioodi algus- ja lõppkuupäev
 2. Ajutise töövõimetuse alus (haigushüvitis; hooldushüvitis; muu)
 3. Ajutise töövõimetuse hüvitise suurus.
 4. Ajutise töövõimetuse hüvitise arvutamiseks kasutatud keskmine tulu.

Selgitus: ravikindlustuse puudulikkus on üks sotsiaalmajanduslikest tunnustest, mis võib ennustada südamehaiguste teket. Lisaks aitavad kindlustuse andmed tuvastada isikuid, kes jälgimisperioodil uuringust lahkuvad. Kirjanduse alusel on leitud, et haiguslehele minemine või hooldajaks saamine tõstab infarkti riski. Antud tegurite analüüsimata jätmine langetaks riskiennustuskalkulaatori täpsust. Ühelt poolt suurendaks see inimeste hulka, kes saavad statiinravi ilma vajaduseta. Teisalt langetaks see nende inimeste hulka, kes statiinravi ei saa, kuigi neile oleks seda tegelikult vaja.

4. Surma põhjuste register (surmad vahemikus 01.01.2010–31.12.2026)
 1. Surmakuupäev
 2. Diagnoosid
 1. Vahetu surmapõhjuse diagnoosid, koodid (RHK-10) ja nimetused
 2. Varasemad surmapõhjuse diagnoosid, koodid (RHK-10) ja nimetused
 3. Surma algpõhjuse diagnoosid, koodid (RHK-10) ja nimetused
 4. Surma välispõhjuse diagnoosid, koodid (RHK-10) ja nimetused
 5. II osa diagnoosid (muud olulised haigused ja seisundid), koodid (RHK-10) ja nimetused
 3. Suremispaik (kood ja nimetus; nt 1 – raviasutus, 2 – kodu jne)
 4. Surma põhjuse kindlakstegemine – nimetus ja kood (1 – lahang, 2 – meditsiinidokument, 3 – laiba ülevaatus jne)

Selgitus: surmaregistri andmeid on vaja esiteks, et võimalikult täpselt tuvastada südamehaiguste teket. Nii suremispaik ja surmapõhjuse kindlakstegemise viis mõjutavad eeldatava surmapõhjuse usaldusväärsust. Andmete analüüsimata jätmine langetaks rahvastiku üleselt riskiennustuskalkulaatori täpsust ja statiinravi tõhusust.

5. Müokardiinfarktiregister (juhtumid vahemikus 01.01.2010–31.12.2026)
 1. Andmed raviarsti ning teatise täitja kohta (raviarst, meditsiinasutus, **kliinik, osakond**, allüksus, **sisestatud teatise viimase muutmise aeg, sisestamiskuupäev, teatise täitja**)
 2. Eelnev kardiovaskulaarne anamnees
 3. Riskitegurid
 4. Atakiga seotud andmed (sh sümptomite algus)
 5. Kliiniline diagnoos, isiku surma korral surmatunnistuse diagnoosid (vastavalt RHK-10).
 1. Juhul kui on teada ja mugav väljastada, siis kas diagnoos oli põhihaigus või kaasuv haigus.

Selgitus: müokardiinfarktiregistri andmeid on vaja esmalt selleks, et võimalikult täpselt tuvastada südamehaiguste teke. Uuringud teistes riikides näitavad, et infarkti on sageli mainitud vaid ühes e-terviseandmebaasis, aga puudub kõigis teistes andmebaasides. Seda kinnitab ka vastutava uurija teaduslik kogemus geenidoonoreil leitud andmetel. Informatsioon diagnoosinud arsti eriala ja töökoha kohta mõjutab diagnoosi usaldusväärsust.

Registri andmete analüüsimata jätmine langetaks rahvastiku üleselt riskiennustuskalkulaatori täpsust ja statiinravi tõhusust. Lisaks annab riskitegurite väli uurijatele lisainformatsiooni selle kohta, kas eelnevalt kogutud riskiinformatsioon teistes andmebaasides oli piisav või ebapiisav.

SOTSIAALSED ANDMEBAASID

6. Eesti rahvastikuregister (01.01.2012, 01.01.2014 ja 01.01.2016 seisuga)
 1. Isiku sünniaasta
 2. Isiku surmakuupäev
 3. Isiku sugu
 4. Isiku elukoht (riik, maakond)
 5. Kodakondsus
 6. Rahvus
 7. Kõrgeim omandatud haridustase
 8. Abielustaatus
 9. Juhul kui isik väljus rahvastikuregistrist vahemikus 01.01.2012–31.12.2026, kuna kolis Eestist välismaale, siis sel juhul esmakordne välismaale kolimise kuupäev.

Selgitus: rahvastikuregistri andmeid on vaja esmalt selleks, et:

- 1) võimalikult täpselt tuvastada algne kohort keda jälgida. Teiste andmebaasidega loodud kohordid on reeglina haigemad, kuna terved inimesed ei kasuta tervishoiu- ja riigiteenuseid nii palju kui haigemad inimesed. Kallutatud valimiga loodud riskiennustusmudel oleks valesti kalibreeritud. See ülehindaks riski ja soovitaks statiine liiga paljudele inimestele, kuigi tegelikult on Eestis palju terveid inimesi, kes on tuvastatavad vaid rahvastikuregistri alusel.
- 2) võimalikult soodsalt määrata sotsiaalmajanduslikud riskitegurid võimalikult representatiivsel valimil. Uurijad arvestavad, et muud meetodid (nt küsimustiku saatmine, nagu tehti geenivaramus) annavad täpsemaid andmeid. Need aga on kulukamad nii riigile kui kodanikule, pluss küsimustikega tekib suur probleem, et pooled inimesed ei soovi küsimustikele vastata. Vastajatel on suremusrisk u 2 x madalam kui mittevastajatel, seepärast on kasulik mõõta sotsiaalmajanduslikke riskitegureid ka nende seas, kes muidu küsimustikele vastata ei soovi.
- 3) võimalikult täpselt tuvastada, kes lahkub riigist uuringu jälgimisperioodi jooksul. Kui see tuvastamata jätta ja uuritav saab südamehaiguse teises riigis, mille kohta uurijatel andmed puuduvad, siis loodud riskikalkulaator alahindaks riske (st soovitaks statiine liiga vähestele kodanikele).

7. Töötamise register

Väljavõte iga isiku töötamise perioodidest (kus vähemalt osa tööajast oli) vahemikus 01.01.2010–31.12.2015, sisaldades iga episoodi:

1. alguskuupäeva
2. lõppkuupäeva
3. töötamise liikide klassifikaatorit
4. ametikoha klassifikaatorit

8. Maksukohustuslaste register

Väljavõte iga isiku aasta netotulust järgmistel perioodidel:

1. 01.01.2011–31.12.2011
2. 01.01.2013–31.12.2013
3. 01.01.2015–31.12.2015

Andmeid väljastada **keskmise kuu kohta** järgmistes kategooriates (EUR): 0; 1-99; 100-199; 200-299; 300-399; 400-499; 500-599; 600-699; 700-799; 800-999; 1000-1199; 1300-1599; 1600-1999; 2000-2999; 3000+

Selgitus: 2011.a. mediaan netopalk oli 553 EUR. See võimaldab eristada mediaanist vasemat 4 gruppi, kus infarkti riski tõuseb järsult sissetuleku langemisega. Mediaanist rikkamat jaotust on raske prognoosida (ja muutub kiiremini 2012-2016 jooksul) mistõttu grupeerivad teadlased need sobivalt edasi andmete analüüsimise käigus.

Ühelt poolt on see piisav granulaarsus, et ennustada infarkti piisavalt täpselt. Teiselt poolt, täiendavate gruppide lisamine omaks küsitlevat kasu täpsusele ja seevastu nõrgendaks andmete minimeerimise printsiipi.

9. Töötuna ja töötisjana arvel olevate isikute ning tööturuteenuste osutamise register

Väljavõte iga isiku töötuna arvel olemiste perioodidest vahemikus 01.01.2008–31.12.2015, sisaldades iga episoodi:

1. alguskuupäeva
2. lõppkuupäeva
3. töötamise lõpu põhjus (8 kategooriat)
4. töötamise arvel oleku lõpetamise põhjus (6 kategooriat)
5. töötuna arvel olemisele eelneva töökoha ametikoha (klassifikaator).
6. töötutoetuse ja töötukindlustushüvitise väljamakse suurus (EUR).
7. töötutoetuse väljamakse arvutamisel kasutatud lähimineviku brutotulu (EUR).

Selgitus: Uurijate eelmised uuringud on leidnud, et isegi, kui juba on mõõdetud uuritava haridust ja abielustaatus, suurendab täiendavate andmete kogumine sissetuleku ja tööpuuduse kohta võimekust südamehaigusi veel täpsemini ennustada. Kirjanduses on varem leitud, et amet ennustab südamehaiguse teket. Kirjanduses on varem leitud, et lisateave selle kohta kas isik oli töötanud 1,2,3 või 4 aastat annab täpsemat infarkti riski ennustust. Kirjandus on näidanud, et töötuks jäämistest hulk (nt 0-6 korda) eelmisel 8. aastal annab täpsemat infarkti riski ennustust kui töötuse binaarne käsitlus. Andmete minimeerimise printsiibist küsime minevikku vaatavaid andmeid -4 aastat esimese kohordi puhul (st. 2008-2012) ja -8 aastat kolmanda kohordi puhul (st. 2008-2016). Antud andmed võimaldaks meil kontrollida (lähtuvalt uuringu esimesest küsimusest ja hüpoteesist) kas samasuguseid seoseid leiame ka Eesti andmetes. Antud andmete analüüsimata jätmise langetaks rahvastiku üleselt riskiennustuskalkulaatori täpsust ja statiinravi tõhusust.

10. Sotsiaalkaitse infosüsteem

Väljavõte iga isiku toetustest makstud perioodil 01.01.2010–31.12.2015, juhul kui isik sai ühe allolevast kuuest toetusest, sisaldades iga väljamakstud toetuse puhul:

1. Toetuse liik:
 1. kuriteoohvri hüvitis
 2. represseeritu toetus
 3. toetus Tšornobõli katastroofi tagajärgede likvideerijale
 4. üksi elava pensionäri toetus.
 5. kutsehaigusest/tööõnnetusest tingitud tervisekahju hüvitised
 6. puuetega inimeste sotsiaaltoetus
 7. töövõimetuspension
2. Toetuse perioodi (nt töövõimetuspensionihüviti puhul), selgitamaks mis aja eest on alus toetusele ja mis aja eest seda välja makstakse.
3. Toetuse summa (EUR), ümardatud lähima kümne euroni.
4. Maksekuupäev

Selgitus: üksi elamine, kutsehaigused ja puude olemasolu on teadolevad südamehaiguste riskifaktorid. Uurijad plaanivad analüüsida, kas nende lisamine riskiennustuskalkulaatorisse tõstaks selle täpsust.

Tšornobõli likvideerimine, poliitiline repressioon ja kuriteo ohvriks langemine on faktorid, mille seost südamehaigustega pole varem uuritud. Uurijad plaanivad vastata küsimusele, kas need ennustavad südamehaiguste teket. Teoreetiliselt peaks see nii olema, kuna kõik kolm sisaldavad tugevat psühholoogilist stressi. Ka mõõdukalt psühholoogilist stressi on varem seostatud südamehaiguste tekkega. Juhul kui need tegurid ennustavad südamehaiguste teket, langetaks andmete analüüsimata jätmise rahvastiku üleselt riskiennustuskalkulaatori täpsust ja statiinravi tõhusust.

11. Kinnistusraamat

Taastada kinnistusraamat ja nende reaalsad kolmel ajahetkel (01.01.2012, 01.01.2014, 01.01.2016 kuupäevadega). Kui üks nendest kolmest raamatutest sisaldab omanikke kes on mailitud Siseministeeriumi pseudonüümide loetelus, siis nende omanike kohta soovime järgnevaid andmeid:

1. Sihtotstarve (nt elamumaa 100%)
2. Aadress:
Andmete minimeerimise huvides väljastada vaid ADS tasemed 1-3 (maakond, omavalitsus, asustustüksus nt vald, linnaosa) ja mitte väljastada ADS tasemed 4-8 (nt väikekoht, tänav, maja number).
3. Omanik:
Omaniku nime ega isikukood pole vaja jagada. Asendada omaniku info omaniku pseudonüümiga.
4. Reaalsad pindala ruutmeetrites, ümardatud järgnevalt: 0-39m², 40-49m², 50-59m², 60-69m², ≥70m².

Selgitus: Kirjandus on näidanud, kuidas kinnisvara omamine ja suurema kinnisvara omamine on seotud madalama infarkti riskiga. Seni on teadmata kas kinnistu liik, sihtotstarve ja hägune aadress (nt linn/maakoh) on seotud infarkti riskiga. Uuring vastab neile uurimisküsimustele.

12. Äriregister

Analüüsida majandusaasta aruandeid kolmel kalendriaastal (2011, 2013 ja 2015). Kui üks nendest kolmest sisaldab omanikke kes on mainitud Siseministeeriumi pseudonüümide loetelus, siis nende omanike kohta soovime järgnevaid andmeid:

1. Õiguslik vorm (nt Osühing)
2. Põhitegevusala (EMTAK ja NACE kood)
3. Aadress:
Andmete minimeerimise huvides palume mitte jagada aadressi esimest rida (nt tänav, number) ja jagada vaid järgmiseid ridasid (nt linnaosa, linn, maakond, st. ADS tase 1-4).
4. Maksualane info: Maksuvõlg olemas / puudub.
5. Järgnevad andmed soovime nii aruande majandusaastal (2011, 2013 ja 2015) kui ka aruande perioodile eelneval majandusaastal (2010, 2012, 2014). Kõik summad ümardada lähima kümne tuhande euroni:
 1. Bilanssi leht:
 1. Eelmiste perioodide jaotamata kasum (kahjum) / akumulieeritud tulem (rida 660)
 2. Aruandeaasta kasum (kahjum) / tulem (rida 670)
 3. Omakapital või netovara kokku (rida 700)
 2. Rahavoogude leht:
 1. Makstud dividendid (rida 336)
6. Tasutud maksud ja töötajate arv eelmises kvartalis (st 2011, 2013 ja 2015 viimane kvartal):
 1. Riiklikud maksud (ümardatud lähima tuhande euroni).
 2. Tööjõumaksud (ümardatud lähima tuhande euroni).
 3. Töötajate arv (kui rohkem kui 7, siis ümardada lähima kümneni. Kui rohkem kui 75, siis ümardada lähima sajani. Kui rohkem kui 750, siis ümardada lähima tuhandeni.)
7. Omanik:
 1. Omaniku nime ega isikukood pole vaja jagada. Asendada omaniku info omaniku pseudonüümiga.
8. Osalus (nt 100%)

Selgitus: Kirjandus on näidanud, kuidas varade omamine (sh. dividendide saamine) on seotud madalama infarkti riskiga. Seni on teadmata kas sellele lisaks langetab infarkti riski ka subjektiivne heaolu mis tuleb siis, kui sinu poolt omatud ettevõtte pakub tööd mitmetele inimestele, maksab palju makse, ja/või on ilma makseraskusteta. Samuti uurime interaktsioone ettevõtte tegevusalaga (nt kognitiivsed vs manuaalsed tegevusalad) ja aadressiga (nt linn / väiksem asula) infarkti ennustamises. Lisaks on teadmata kas infarkti ennustab paremini tegelikult välja makstud dividendid (üle-eelmise aasta kasumi arvelt) või lähituleviku ootus järgmise aasta dividendide saamisele (eelmise aasta kasumi arvelt). Uuring vastab neile uurimisküsimustele.

Lisa 10.2.1 – Curriculum vitae – Taavi Tillmann (vastutav uurija)

Töökohad

01.01.2022–...	Tartu Ülikool, meditsiiniteaduste valdkond, peremeditsiini ja rahvatervishoiu instituut, rahvatervishoiu kaasprofessor (1,00. Koormus ajutiselt 0,60 väikelapse tõttu.)
01.09.2017–31.12.2021	UK Tervise Arengu Instituut; UK Terviseamet; linnavalitsused, rahvatervise arst-resident (0,50) UCL maailmatervise instituut, lektor (0,50)
01.09.2014–31.08.2017	UCL epidemioloogia ja rahvatervise osakond, kliiniline nooremteadur (doktorant) (1,00)
01.08.2011–01.08.2013	Londoni haiglad, nooremars (1,00)

Haridus

2014–2021	Suurbritannia rahvatervise teaduskonna liige (st residentuur lõpetatud). Sh. uuringute eetika ja andmekaitse koolitused ja kompetentsused (vt. https://www.fph.org.uk/media/3450/public-health-curriculum-v13.pdf lk 34, 75)
2014–2018	PhD, sotsiaalepidemioloogia ja rahvatervis, University College London “Psychosocial and socioeconomic factors in the development of cardiovascular disease: a study of causality, international variation, mediation, and prediction in predominantly Eastern European settings”
2013–2014	MSc rahvatervis (tervisemajanduse suund), London School of Hygiene and Tropical Medicine “Alcohol related mortality in the Baltic States”
2005–2011	MBChB, arstiteadus, University of Glasgow
2008–2009	BSc, psühholoogia, University College London

Valitud publikatsioonid

Tillmann T, Läll K, Dukes O, ... Fischer K, Kivimäki M, Bobak M. Development and validation of two SCORE-based cardiovascular risk prediction models for Eastern Europe: a multicohort study. Eur H J. 2020.

Tillmann T. New risk prediction models in England may lead to targeted PCSK9 inhibitor treatment, for patients with established cardiovascular disease. Eur J Prev Cardiol. 2020

Tillmann T. Atherosclerotic metabolites: basic science is progressing, so we need to think about clinical implications. Eur H J. 2019;40(34):2897-2898

Hageman S...Tillmann T...Visseren F... Angelantonio ED. SCORE2 risk prediction algorithms: new models to estimate 10-year risk of cardiovascular disease in Europe. Eur H J;2021;42(25):2439-2454

Tillmann T, Vaucher J, Okbay A, ... Davey Smith G, Bobak M, Holmes MV. Education and coronary heart disease: a Mendelian randomization study. BMJ. 2017;358:j3542.

Tillmann T, Pikhart H, Peasey A, ... Steptoe A, Kivimäki M, Marmot M, Bobak M. Psychosocial and socioeconomic determinants of cardiovascular mortality in Eastern Europe: a multicentre prospective cohort study. PLOS Medicine. 2017;14(12):e1002459.

Tillmann T, Gibson A, Harrison O, Scott G, Dominiczak A, Hanlon P. A systems approach to medical science may help reveal the individual nature of chronic diseases. Journal of Medical Internet Research. 2015;17(3):e64

Tillmann T, Cavanagh J, Krishnadas R, Petrides KV. Possible rheumatoid arthritis subtypes in terms of rheumatoid factor, depression, diagnostic delay and emotional expression: an exploratory case-control study. Arthritis Research & Therapy. 2013;15(2):R45

Teised jooksvad projektid

Digiravim "Activate": Kaalu langetava digiravimi prototüübi arendamine ja testimine
01.10.2021–31.12.2023 32 500 EUR Rahastaja: Activate Health OÜ
Sisu: kirjanduse ülevaade ja ärialane nõustamine. Uuritavaid ei osale.

Esmatasandi patsiendi virtuaalteenindamise tarkvara arendus ja vastuvõetavuse uuring.
01.04.2023–31.12.2023 29 999 EUR Rahastaja: Certific OÜ
Sisu: kvalitatiivne uuring kus oleme teinud intervjuusid 10 tervishoiutöötajaga. Patsiente ei uurita.

Lisa 10.2.2. Kaastöötajate CVd:

Alljärgnevatest linkidest leiate magistrikraadiga kaastöötajatele CVd. Lisaks kopeerisin ühe kõige relevantsema projekti või kogemuse millel sarnased eetilised ja/või andmekaitsealased aspektid.

Soovi korral võime terve CV taotlusesse kopeerida. Otsustasime seda käesolevas taotluses mitte teha, et dokument ei muutuks liiga pikaks.

https://www.etis.ee/CV/Raivo_Kolde/est/

[Kliiniliselt oluliste radade tuvastamine terviseandmetest](#)

https://www.etis.ee/CV/Anneli_Uusk%C3%BCla001/est/

[Aastatel 2006–2015 enesetapu sooritanute epidemioloogiline ülevaade](#)

https://www.etis.ee/CV/Markus_Haug/est/

[COVID-19 haigusjuhtumite analüüs ja riskirühmade väljaselgitamine Eestis](#)

https://www.etis.ee/CV/Marek_Oja/est/

[Eesnäärmevähi ravitrajektoorid Eestis 2012.-2019. a tervisedokumentide põhjal](#)

https://www.etis.ee/CV/Elena_S%C3%BCgis/est/

[Algoritmid ja tehisintellekt digitaalsete terviseandmete analüüsis](#)

Lisa 10.3 – Uuringu meetodi ja andmekaitsetaseme aspektid uuritavatele
(mida avaldame aadressil <https://tervis.ut.ee/et/teadus>)

Projekt: E-sotsiaalandmete kasulikus südame-veresoonkonna haigustekke ennustamisel
Teavitustekst, kuidas Tartu Ülikool selles projektis isikuandmeid töötleb

Personaalmehaanika on loonud mitmed kalkulaatoreid mis ennustavad kes võib tulevikus haigestuda. Need aga ei arvesta sotsiaalmajanduslikke tegureid nagu näiteks kas keegi on töötanud või mis on tema haridustase. Kõrgema haridusega inimesed kasutavad selliseid kalkulaatoreid rohkem. See suurendab tervise ebavõrdsust. Antud projekt soovib seda parandada.

Mis on uuringu eesmärgid?

Peamine uuringu küsimus on, kas **sotsiaalmajanduslikud tegurid, mis leiduvad olemasolevates riigi andmebaasides saavad aidata ennustada kellel tekib südame- ja veresoonkonna haigus?** Sellele vastamiseks loome uue riskikalkulaatori, mis ennustab tervete inimeste seas infarkti riski järgmise 10 aasta jooksul.

Kuidas me seda teeme?

Loomes kohort-uuringu ligikaudselt 660 000 Eesti elaniku andmetega. Iga inimese puhul lingime kokku olemasolevaid e-sotsiaalvaldkonna andmeid (sh. Eesti rahvastikuregistrit; töötuna olevate isikute registrit; töötamise registrit; maksukohustuslaste registrit; sotsiaalkaitse infosüsteemi; kinnistusraamatut; äriregistrit). Kui Teie olite 01.01.2023 kuupäevaga vähemalt 46 aastat vana ja kui Te elasite Eestis 2010. aasta esimeses pooles, siis suure tõenäosusega analüüsime ka Teie andmeid. Olemasolevatest andmebaasidest tuletame mis on Teie rahvus ja haridustase; mida Te tegite 2011. a. jooksul (sh. mis maakonnas elasite, mis oli Teie amet ja sissetulek, kas olite abielus, töötanud või saite teatud sotsiaalseid toetusi) ja mis oli Teie varaline seis (kas omasite kinnisvara või ettevõtteid). Kõik need tegurid arvatavasti ennustavad infarkti teket. Peale andmete puhastamist loome uue infarkti ennustamise riskikalkulaatori. Mõeldakse kui täpne see on. Teie andmeid töötleme pseudonüümselt kus me ei tea teie isikuandmeid. Tulemusena loodame luua uue haigust ennustava meditsiiniseadme edaspidiseks testimiseks. Uuring loodab illustreerida kuidas vähendada tervise ebavõrdsust, kui sotsiaalsed tegurid on nutikalt integreeritud uute teenuste arendamise protsessi.

Mida uuring pakub kodanikele ja kodaniku õigused

Selle uuringu jooksul ei suhtle me üldse kodanikega ega patsientidega. Küsime luba kahelt Tartu Ülikooli inimuuringu eetika komiteelt. Peale loa saamist saavad uurijad kätte pseudonüümitud andmed ligikaudselt poole Eesti elanike kohta. Teie ei pea selleks midagi tegema ega luba andma.

Kui Te ei soovi, et me Teie andmeid analüüsime, siis palun saatke lühikene avaldus vastutava uurija e-posti aadressile taavi.tillmann@ut.ee. Te võite seda teha ükskõik millal enne uuringu lõppu 31.08.2028. See, kas me Teie andmeid analüüsime või mitte, ei mõjuta Teie õigust saada mis tahes liiki tervishoiu või muid riigi teenuseid. Laiemas plaanis on Teil õigus tutvuda milliseid Teie kohta kogutud isikuandmeid me töötleme, edastada need teisele vastutavale töötajale ja/või nõuda ebaõigete isikuandmete parandamist. Samuti võite nõuda, et peatame Teie andmete töötlemist juhul, kui olete isikuandmete õigsust vaidlustanud või esitanud vastuväite, või kui vajate andmeid oma õigusnõuete esitamiseks.

Kuidas me kodanike andmeid töötleme?

Vaatame rahvastikuregistrit kes elasid Eestis 2012-2016 vahel. Kui Te elasite Eestis, siis Siseministeerium pseudonüümitab Teie andmed. See tähendab, et nad ei avalda teadlastele ei Teie nime, isikukoodi, sünnikuupäeva, täpset aadressi ega kontakt andmeid. Siseministeerium loob Teile uue pseudonime. Siseministeerium loob ka andmevõtme, ehk lühikese tabeli kus on kirjas Teie pseudonimi ja isikukood. Siseministeerium edastab andmevõtme teistele riigi andmekogude omanikele, kes tuvastavad isikukoodi alusel kas Tartu Ülikooli teadlased soovivad Teie andmeid antud uuringus analüüsida või mitte. Teiste andmekogude omanikud edastavad enda poolt pseudonüümiseeritud andmed Tartu Ülikoolile. Ülikooli Teadusarvutuse keskkuses luuakse spetsiaalne turvaline keskkond kus Teie pseudonüümitud andmed analüüsivad vaid nimetatud uurijad. Andmed hoitakse siin krüpteeritud ja protsesse pidevalt seiratakse, et tagada andmete turvalisus.

Kuidas saada rohkem infot?

Projekti vastutav täitja on rahvatervise kaasprofessor Taavi Tillmann.

Ta töötab Tartu Ülikoolis, Peremeditsiini ja rahvatervishoiu instituudis (4. Korrus, Biomeediku, Ravila 19, 50411 Tartu).

Taavi.tillmann@ut.ee + 372 5342 6963

Ootame jooksuvaid küsimusi vahemikus 2023-2028:

- Kodanikelt andmete töötlemise kohta;
- Tudengitelt ja teadlastelt koostöövõimaluste kohta, ja
- Meedialt töö tulemuste ja leidude kohta.

Kes vastutab andmete töötlemise eest?

Vastutav töötaja on Tartu Ülikool (registrikood 74001073, aadress Ülikooli 18, 50090 Tartu). Isikuandmete töötlemise ja Teie õiguste kohta võite pöörduda ülikooli andmekaitse spetsialisti poole aadressil andmekaitse@ut.ee. Kui arvate, et viis, kuidas ülikool töötleb Teie isikuandmeid on vastuolus isikuandmete seadusega, võite pöörduda Andmekaitse Inspektsiooni poole (info@aki.ee, 627 4135).