**UURINGU EETILISE HINDAMISE TAOTLUS EESTI BIOEETIKA JA INIMUURINGUTE NÕUKOGULE**

Taotlusele 05.12.2024 lisatud täiendused on märgitud kollase värviga.

|  |  |  |
| --- | --- | --- |
| **1. Uuringu nimetus (ingliskeelsete taotluse puhul tuleb uuringu nimetus ära tuua ka eesti keeles)** | | |
| Personaliseeritud terviseriskide andmepõhiste mudelite arendamine ja valideerimine Eesti personaalmeditsiini programmi eesmärkide täitmiseks | | |
| **2. Uuringu põhieesmärk kuni 450 tähemärki (0,25 lk) (ingliskeelsete taotluse puhul tuleb uuringu põhieesmärk ära tuua ka eesti keeles)** | | |
| Andmepõhiste riskimudelite, sh geneetiliste riskiskooride ja terviseandmete põhiste otsustustoe algoritmide teaduslik arendus on maailmas oluliselt hoogustunud seoses meetodite arengu ning andmete kättesaadavuse paranemisega. Välja on pakutud mitmeid sadu erinevaid geneetilisi riskiskoore, sajakonna eri haiguse ja riskide ennustamiseks; samuti on Eesti juurutanud arstidele suunatud esimesed otsusetoe süsteemid. Samas on nende mudelite kasutuselevõtt kliinilises praktikas reeglina takistatud, sest ennustavad tarkvaramudelid liigituvad meditsiiniseadmeteks. Selliste seadmete arendusele, valideerimisele, juurutamisele ja tootmisele kehtivad nõuded ja standardid on oluliselt karmimad võrreldes teaduspublikatsioonide tavapärase põgusa eelretsenseerimisega (*peer review*).  Uuringu eesmärk on töötada välja metoodika ja taristu geneetiliste ja terviseandmete põhiste riskimudelite arendamiseks ja valideerimiseks (*in silico*), et neid oleks võimalik hiljem kergema vaevaga meditsiiniseadmeks edasi arendada ja/või kasutusele võtta. | | |
| **3. Vastutava(d) uurija(d) ning tema (nende) kontaktandmed** | | |
| Eesnimi: Jaak  Perekonnanimi: Vilo  Ametikoht: professor, andmeteaduse õppetooli juhataja, instituudi juhataja  Organisatsioon: Tartu Ülikool, arvutiteaduse instituut  Telefon: +372 5049365  e-post: [vilo@ut.ee](mailto:vilo@ut.ee)  Skype: jaakvilo | | |
| **4. Uuringu läbiviijad (lisada juurde vajalik arv ridu)** | | |
| 1. Eesnimi: Raivo   Perekonnanimi: Kolde  Ametikoht: terviseinformaatika kaasprofessor  Organisatsioon: Tartu Ülikool, arvutiteaduse instituut   1. Eesnimi: Sulev   Perekonnanimi: Reisberg  Ametikoht: terviseinformaatika teadur  Organisatsioon: Tartu Ülikool, arvutiteaduse instituut   1. Eesnimi: Sven   Perekonnanimi: Laur  Ametikoht: kaasprofessor  Organisatsioon: Tartu Ülikool, arvutiteaduse instituut   1. Eesnimi: Elena   Perekonnanimi: Sügis  Ametikoht: bioinformaatika teadur  Organisatsioon: Tartu Ülikool, arvutiteaduse instituut   1. Eesnimi: Ida Maria   Perekonnanimi: Orula  Ametikoht: nooremteadur  Organisatsioon: Tartu Ülikool, arvutiteaduse instituut   1. Eesnimi: Marek   Perekonnanimi: Oja  Ametikoht: terviseinformaatika teadur  Organisatsioon: Tartu Ülikool, arvutiteaduse instituut   1. Eesnimi: Sirli   Perekonnanimi: Tamm  Ametikoht: programmeerija  Organisatsioon: Tartu Ülikool, arvutiteaduse instituut   1. Eesnimi: Harry-Anton   Perekonnanimi: Talvik  Ametikoht: terviseinformaatika nooremteadur  Organisatsioon: Tartu Ülikool, arvutiteaduse instituut   1. Eesnimi: Lili   Perekonnanimi: Milani  Ametikoht: asedirektor, epi- ja farmakogenoomika professor  Organisatsioon: Tartu Ülikool, genoomika instituut   1. Eesnimi: Kristi   Perekonnanimi: Krebs  Ametikoht: farmakogenoomika teadur  Organisatsioon: Tartu Ülikool, genoomika instituut   1. Eesnimi: Kristi   Perekonnanimi: Läll  Ametikoht: statistilise geneetika teadur  Organisatsioon: Tartu Ülikool, genoomika instituut   1. Eesnimi: Reedik   Perekonnanimi: Mägi  Ametikoht: bioinformaatika professor  Organisatsioon: Tartu Ülikool, genoomika instituut   1. Eesnimi: Kristjan   Perekonnanimi: Metsalu  Ametikoht: IT arendusjuht  Organisatsioon: Tartu Ülikool, genoomika instituut   1. Eesnimi: Sirje   Perekonnanimi: Kolde  Ametikoht: andmeanalüütik  Organisatsioon: Tartu Ülikool, genoomika instituut   1. Eesnimi: Hendrik   Perekonnanimi: Šuvalov  Ametikoht: terviseinformaatika nooremteadur  Organisatsioon: Tartu Ülikool, arvutiteaduse instituut   1. Eesnimi: Markus   Perekonnanimi: Haug  Ametikoht: terviseinformaatika nooremteadur  Organisatsioon: Tartu Ülikool, arvutiteaduse instituut   1. Eesnimi: Maarja   Perekonnanimi: Pajusalu  Ametikoht: terviseinformaatika nooremteadur  Organisatsioon: Tartu Ülikool, arvutiteaduse instituut   1. Eesnimi: Dage   Perekonnanimi: Särg  Ametikoht: doktorant  Organisatsioon: Tartu Ülikool, arvutiteaduse instituut   1. Eesnimi: Õie Renata   Perekonnanimi: Siimon  Ametikoht: terviseinformaatika nooremteadur  Organisatsioon: Tartu Ülikool, arvutiteaduse instituut   1. Eesnimi: Mai-Liis   Perekonnanimi: Vahemäe  Ametikoht: magistrant (juhendaja Sven Laur)  Organisatsioon: Tartu Ülikool, arvutiteaduse instituut   1. Eesnimi: Anett   Perekonnanimi: Sandberg  Ametikoht: magistrant (juhendaja Raivo Kolde)  Organisatsioon: Tartu Ülikool, arvutiteaduse instituut   1. Eesnimi: Teele   Perekonnanimi: Kuri  Ametikoht: magistrant (juhendaja Raivo Kolde)  Organisatsioon: Tartu Ülikool, arvutiteaduse instituut   1. Eesnimi: Selene Margaret   Perekonnanimi: Pruuden  Ametikoht: magistrant (juhendaja Raivo Kolde)  Organisatsioon: Tartu Ülikool, arvutiteaduse instituut   1. Eesnimi: Kerli   Perekonnanimi: Mooses  Ametikoht: teadur  Organisatsioon: Tartu Ülikool, arvutiteaduse instituut   1. Eesnimi: Kunnar   Perekonnanimi: Kukk  Ametikoht: nooremteadur  Organisatsioon: Tartu Ülikool, arvutiteaduse instituut   1. Eesnimi: Kaire   Perekonnanimi: Koljal  Ametikoht: programmeerija (juhendaja Sven Laur)  Organisatsioon: Tartu Ülikool, arvutiteaduse instituut   1. Eesnimi: Telver   Perekonnanimi: Objärtel  Ametikoht: spetsialist (juhendaja Raivo Kolde)  Organisatsioon: Tartu Ülikool, genoomika instituut   1. Eesnimi: Hanna   Perekonnanimi: Sõnajalg  Ametikoht: statistik  Organisatsioon: Tartu Ülikool, Eesti Geenivaramu teaduskeskus   1. Eesnimi: Maris   Perekonnanimi: Vojdani  Ametikoht: magistrant (juhendaja Lili Milani)  Organisatsioon: Tartu Ülikool, eesti ja üldkeeleteaduse instituut   1. Eesnimi: Ami   Perekonnanimi: Sild  Ametikoht: programmeerija  Organisatsioon: Tartu Ülikool, arvutiteaduse instituut   1. Eesnimi: Kermo   Perekonnanimi: Saarse  Ametikoht: nooremteadur  Organisatsioon: Tartu Ülikool, arvutiteaduse instituut   1. Eesnimi: Anton   Perekonnanimi: Vykhovanets  Ametikoht: nooremteadur  Organisatsioon: Tartu Ülikool, arvutiteaduse instituut   1. Eesnimi: Maria   Perekonnanimi: Malk  Ametikoht: nooremteadur  Organisatsioon: Tartu Ülikool, arvutiteaduse instituut   1. Eesnimi: Neeme   Perekonnanimi: Ilves  Ametikoht: spetsialist  Organisatsioon: Tartu Ülikool, arvutiteaduse instituut   1. Eesnimi: Laura   Perekonnanimi: Lõo  Ametikoht: nooremteadur  Organisatsioon: Tartu Ülikool, arvutiteaduse instituut   1. Eesnimi: Kristjan   Perekonnanimi: Laht  Ametikoht: magistrant  Organisatsioon: Tartu Ülikool, arvutiteaduse instituut | | |
| **5. Uuringu finantseerimine** | | |
| **Finantseerimise allikad** | Eesti riiklik Personaalmeditsiini IT taristu arendamise projekt GenMed (vastavalt ERF TAT tingimustele) 2019-31.12.2022.  PRG1095 "Algoritmid ja tehisintellekt digitaalsete terviseandmete analüüsis (1.01.2021−31.12.2021)", Jaak Vilo  PRG1447 - taotlus 2022- 31.12.2026, (Jaak Vilo)  IMI OPTIMA projekt tehisintellekti arendus kolme vähi (prostata, rinna- ja kopsuvähi) terviseandmete ja ravijuhiste põhiste otsustustoe meetodite tehisintellekti arendus. (2021-2025, Andres Metspalu ja Jaak Vilo)  Õppejõudude, teadlaste ja doktorantide põhipalk, stipendiumid ja töölepingud. | |
| **Uuringu üldmaksumus (summa)** | ca 300 000 eurot aastas, 5 aastat; ning lisaks töötajad ja doktorandid oma põhitöö raames. | |
| **Uuritavale kompensatsiooni maksmine (jah, ei, põhjendus ja summa)** | Ei, kuna uuringus kasutatakse varem kogutud andmeid, geenidoonorite tagasikodeerimist ei toimu. | |
| **Uuritavate kindlustus (jah, ei, kindlustaja ja poliis)** | Ei | |
| **6. Uuringu läbiviimise aeg (algus ja lõpp kuu ja aasta täpsusega)** | | |
| Mai 2021 - Detsember 2026 | | |
| **7. Teave sama uuringu projekti varasema või samaaegse hindamise kohta (sh teistes riikides)** | | |
| Uuringut pole varasemalt ega samaaegselt hinnatud. | | |
| **8. Lühiülevaade siiani samal teemal tehtud uuringutest (kuni 900 tähemärki, 0,5 lk)** | | |
| Tartu Ülikoolis on teadustöö raames välja töötatud arvukaid geneetilisi riskimudeleid, sh polügeenseid riskiskoore (PGS) (nt Läll 2017, Läll 2019, jt), töövoog farmakogeneetilise fenotüübi tuvastamiseks (Reisberg 2019) jne. Kõigil neil on potentsiaali kasutamiseks kliinilises praktikas (Reisberg 2019). Eestis on hetkel käimas südame-veresoonkonna ning rinna- ja munasarjavähi personaliseeritud riskimudelite kliinilised uuringud (<https://www.kliinikum.ee/teadus-ja-oppetoo/arendusprojektid/>). Maailmas on loodud palju erinevaid teaduslikke riskiskoore ning isegi andmebaas vastavate skooride kogumiseks ja levitamiseks. Näiteks The Polygenic Score Catalog (<https://www.pgscatalog.org/>) raporteerib üle 730 eri skoori üle 200 tingimuse jaoks ning koondab üle 140 publikatsiooni andmed.  Tartu Ülikool osaleb ka käimasolevas GenMed (nn ERF TAT) projektis, mille eesmärk on luua riiklik IT-taristu nende mudelite kasutuselevõtuks tervishoiusüsteemis. Samas on geneetiliste riskimudelite kasutuselevõtt seotud mitmete probleemidega, mh on lahendamata küsimus, kuidas hinnata, kas riskimudel konkreetsesse populatsiooni üldse sobib (Reisberg 2017) ning millist mõju selle rakendamine meditsiinisüsteemile tähendaks. Samuti puudub ühtne metoodika uute riskiskooride loomiseks sellisel kujul, et nende meditsiiniseadmeteks edasi arendamine oleks algusest peale olnud vastava arenduse puhul arvesse võetud.  Turule on tulnud esimesed PGS põhised ise-deklareeritud (*self-declared*) valiidsusega geenitestid (nt Antegenes OÜ) ning oodata on ka rahvusvaheliste seadmete sisenemist Eesti turule. Praegu puudub paraku võimekus ja valmidus neid valideerida Eesti personaalmeditsiini programmi kontekstis. | | |
| **9. Planeeritava uuringu põhjendus ning uurimisküsimused ja/või hüpoteesid (kuni 1800 tähemärki, 1 lk)** | | |
| Kuigi nii Tartu Ülikoolis kui maailmas laiemalt on loodud mitmeid riskimudeleid (vt nt PGS Catalog) ning Eestis on arendamisel riiklik taristu nende mudelite tervishoiusüsteemis kasutuselevõtuks, ei ole võimalik olemasolevaid mudeleid sellisel kujul kliinilisse praktikasse juurutada, kuna mudelid ei pruugi olla universaalselt kasutatavad (Reisberg 2017), lisaks liigituvad need mudelid õiguslikult meditsiiniseadmeks, mille väljatöötamise protsessile rakenduvad ranged nõuded juba alates mudeli algsest loomisest. Kuigi Eestis on hetkel käimas kahe riskimudeli kliinilise käsitluse arendusele keskenduvad nn RITA uuringud, pole nende mudelite loomisel lähtutud nõutavatest kvaliteediprotsessidest ega vajadusest riskiskooride arendamiseks CE märgistatavateks meditsiiniseadmeteks. Seetõttu tuleb kliinilisse praktikasse juurutamiseks alustada nende mudelite taasloomist algusest peale, jälgides siis juba kõiki vajalikke riske ning tingimusi.  Käesolevas uurimisprojektis arendatakse geneetiliste riskimudelite loomise meetodeid eesmärgiga võimaldada selliste meetodite alusel loodud mudelite lihtsamat kasutuselevõttu kliinilises praktikas. Eesmärk on katta sama andmetaristuga valmidus valideerida geeni- ja terviseandmeid kasutavad otsusetoe põhised meditsiiniseadmed enne lõplikku juurutust Eesti personaalmeditsiini taristule (Technology readiness level - TRL 7-8 testid), arendada metoodikad, kuidas viia esmane teadustulemus arendamise protsessis tasemele, et saaks üldse hakata taotlema CE märgistamist (TRL 4-7) ning tagada võimalused arendustööks alusteadusest kuni seadme prototüübi loomise vajaduseni välja (TRL1-4).  Aluseks on siin võetud EL TRL tasemed:   1. TRL 1 – Basic principles observed 2. TRL 2 – Technology concept formulated 3. TRL 3 – Experimental proof of concept 4. TRL 4 – Technology validated in lab 5. TRL 5 – Technology validated in relevant environment (industrially relevant environment in the case of key enabling technologies) 6. TRL 6 – Technology demonstrated in relevant environment (industrially relevant environment in the case of key enabling technologies) 7. TRL 7 – System prototype demonstration in operational environment 8. TRL 8 – System complete and qualified 9. TRL 9 – Actual system proven in operational environment   Täpsemad uurimisküsimused on järgmised:   1. **Riskimudelite kliinilises praktikas juurutamise eelne rakendatavuse ja valideerimise taristu ning valideerimismeetodite arendus (TRL 7-8), sh**    1. Geeniriskide valideerimine terviseandmete kontekstis vajab terviseandmeid standardsel kujul, rahvusvaheliselt on selleks standardiks kujunemas OHDSI.org loodav OMOP ühtne andmemudel, mis võimaldab jagada ka tarkvara eri projektide vahel. Käesoleva taotluse käigus tuleb esmalt luua Geenivaramu eri andmekogudest pärit terviseandmeid kaasav valideerimisraamistik (OMOP ehk *The Observational Medical Outcomes Partnership* andmemudeli rakendamine Eesti terviseandmetel) patsiendigruppide statistika ning indiviidipõhiste ennustusmudelite loomiseks ja valideerimiseks.    2. Meetodite arendus erinevate geneetiliste riskimudelite ja meditsiiniseadmete valideerimiseks Eesti andmetel ning Eesti personaalmeditsiini taristul juurutamise ettevalmistuseks. 2. **Tartu Ülikoolis arendatud ja teaduskirjanduses avaldatud geneetiliste riskimudelite edasiarendamine ja nende kliinilises praktikas kasutamiseks ettevalmistamine (TRL 4-7)**    1. Riskimudelite ning haiguste vaheliste seoste valideerimine, polügeensete riskimudelitega teiste haigustega esineda võivate seoste otsimise uuringumetoodika edasiarendus, näiteks PheWAS uuringud rahvusvaheliselt mujal arendatud ja publitseeritud PGS-de ja nende eri kombinatsioonide jaoks, kasutades geeni- ja standardsel OMOP CDM (common data model) kujul terviseandmeid.    2. Riskiskooride ning erinevate haigustrajektooride hindamine, sealhulgas täpsustavate või välistavate tingimuste otsimine ja täiendavate kriteeriumite lisamine ennustusmudelitele, riskiskooride valideerimine *in silico* ning peaaegu valmis riskimudelite ettevalmistamine CE märgistamise protsessiks. 3. **Uute meetodite arendus ühtse terviseandmete kuju loomiseks, riskimudelite ja terviseseisundite vaheliste seoste analüüsiks, teadustöö riskimudelite ja otsusetoe algoritmide arenduseks uudsete algoritmiliste ning masinõppe ja tehisintellekti meetodite abil (TRL 1-5)**    1. Standardiseeritud OMOP CDM põhise terviseandmete andmestiku loomine Geenivaramu doonorite andmetele, taristu ettevalmistus ja haldus OMOP kujul andmeid kasutavate mudelite ja statistiliste uuringute läbiviimiseks.    2. Patsiendigruppide ja tingimuste defineerimine, haigustrajektooride ja ko-morbiidsuste analüüs, sarnaste patsiendigruppide automaatne koostamine ning selle baasil võimalike stsenaariumide kirjeldamine ja võrdlus.    3. Masinõppe meetodite arendus indiviidipõhiste ennustusmeetodite arendamiseks vastavate ennustusmeetodite agregeeritud kvaliteedihinnangute saamiseks.    4. Seletava tehisintellekti meetodite arendamine (*explainable AI, xAI*) ennustuste tutvustamine arstile koos võimalike põhjendustega soovituste kasuks. Seletusi kasutavate AI mudelite loomine on eelduseks edasisteks aruteludeks arstide, Haigekassa jt institutsioonidega, et otsustada võimalused ja vajadused nende meetodite viimiseks arstideni.    5. Ravijuhiste põhiste lähenemiste valideerimine ning rahvusvaheliste analoogsete ravijuhiste OMOP põhiste valideerimispakettide jooksutamine Eesti andmetel.    6. Ravijuhiste laiendamine juhistest välja jäävate juhtumite analüüsimieetodite arendamise ning võimalike eri stsenaariumide võrdlusmeetodide loomise abil.   Käesoleva arenduse eesmärk on teha otseselt ettevalmistusi Eesti riikliku personaalmeditsiini programmi juurutamiseks, valmistudes ette riskiskooride valideerimiseks ning teiseks arendada välja parem metoodika ja valmidus uute vajalike otsusetoe algoritmide loomiseks polügeensete riskide ning terviseandmete kombineerimise ja *in silico* valideerimise abil ning laiemalt uudsete masinõppe ja seletava tehisintellekti meetodite arendamine. Loodav andmetaristu koos arvutuskeskkonnaga peaks olema lisaks võimalikult sarnane Eesti personaalmeditsiini programmi IT taristu kontekstile, et nii selle taristu peal Eestis loodavad kui mujalt kaasatavad lahendused oleks võimalik valideerida ja juurutada sujuvamalt. | | |
| **10. Uurimismetoodika (kuni 1800 tähemärki, 1 lk)** | | |
| Uurimisprojektis kasutatakse Eesti geenivaramu doonorite andmeid. Uuritakse doonorite geeni- ja terviseandmete vahelisi seoseid. Uurimisprojekt jaguneb järgmisteks alamülesanneteks:   1. Tartu Ülikoolis loodud geneetiliste riskimudelite edasiarendamiseks ning meditsiiniseadmeni jõudmiseks raamistiku väljatöötamine. Aluseks võetakse Euroopas 2022. a kehtima hakkav *in vitro* meditsiiniseadmete määrus, mille rakendamiseks luuakse geneetiliste riskimudelite kontekstis konkreetne raamistik. Selle tuumaks on ühelt poolt standarditud ja valideeritud andmed, kuid teisalt kvaliteediprotsessi juurutamist võimaldava tehisintellekti ja tarkvaraarenduse protsessi rakendamine meditsiiniseadme loomisel. Käesoleva projekti käigus on eesmärk muu hulgas lahendada vajalike andmete ja arenduse töövoogude dokumenteerimine ja versioneerimine nii, et vastavaid arvutusi saaks auditeerida, korrata ja valideerida (*reproducibility*). 2. Terviseandmete viimine OMOP kujule. Uurimismeeskonnal on pikaaegne kogemus Eesti terviseandmete analüüsimisel, sh erinevate Eesti terviseandmekogude viimisel ühtsele maailmas levinud OMOP CDM andmekujule. Sellisele kujule viidud andmed võimaldaksid efektiivsemalt luua nii uusi riskimudelid kui ka kontrollida (valideerida) olemasolevate riskimudelite täpsust nii Eestis kui teistes riikides. Käesolevas uurimisprojektis arendatakse edasi raamistikku Eesti terviseandmete viimiseks OMOP kujule geneetiliste riskimudelite arendamise eesmärgil. 3. Olemasolevate riskimudelite ja terviseseisundite vaheliste uute seoste uurimine. Selleks rakendatakse kõigi terviseandmete ülese assotsiatsiooniuuringute meetodit uudsel viisil. Kui seni on uuritud üksikute haigustega seotud geenimutatsioonide seoseid teiste haigustega (Reisberg 2019), siis eesmärk on parandada mudelite ennustusvõimsust ning hinnata ka seda, kas ja kuidas mõne konkreetse haiguse tarbeks loodud geneetilised riskiskoorid on seotud ka muude terviseseisunditega kas üksinda või omavahelistes kombinatsioonides. Selleks testitakse riskiskoore ja nende erievaid kombinatsioone kõikvõimalike erinevate terviseseisundeid kirjeldavate tunnuste ja/või nende kombinatsioonide vastu. | | |
| **11. Uuritavate valim ja värbamise viisi kirjeldus.**  **Uuritavate informeerimise ja nõusoleku vormid, ankeetide, küsitluste ja testide vormid esitada taotluse lisadena.** | | |
| **Valimi suurus ja kontrollgruppide olemasolu** | Kõik Eesti geenivaramu (EGV) geenidoonorid (taotluse esitamise seisuga ca 202 300), kuid see võib käesoleva uuringu kestel ka kasvada.  Seoses sellega, et Geenivaramu andmestik on populatsioonipõhine, on iga konkreetse kohordi jaoks geenidoonorite hulgast andmepäringuga võimalik kokku panna eeldatavalt piisavalt esinduslik statistiliselt sobitatud kontrollgrupp. Seega eraldi pole vaja kaasata uut andmekogu, uuringuteks vajalik info on geenidoonorite kohordi sisene (võrdlused ja kontrollid). | |
| **Kes värbab uuritavaid ja kuidas/kus/kelle poolt võetakse informeeritud nõusolek? (kui on asjakohane)** | Uusi geenidoonoreid uuringu tarbeks ei kaasata. Eraldi informeeritud nõusolekut ei võeta. Kõik EGV geenidoonorid on juba geenidoonoriks astudes allkirjastanud geenidoonoriks saamise nõusoleku vormi.  Geenidoonoritelt informeeritud nõusoleku saamise protseduur lähtub inimgeeniuuringute seadusest (<https://www.riigiteataja.ee/akt/72581?leiaKehtiv>) ja määrusest „Geenidoonoriks saamise nõusoleku vorm” (<https://www.riigiteataja.ee/akt/117042019016?dbNotReadOnly=true>) ning määruse lisas toodud nõusoleku vormist ([https://www.riigiteataja.ee/aktilisa/1170/4201/9016/SOM\_11042019\_m39lisa.pdf#](https://www.riigiteataja.ee/aktilisa/1170/4201/9016/SOM_11042019_m39lisa.pdf)) | |
| **Kuidas ja kelle hulgast toimub uuritavate valik? Millised on uuritavate kaasamise või väljajätmise kriteeriumid?** | Eraldi valikut ei toimu. Analüüsides kasutatakse valimi alusena kogu geenidoonorite kohorti. | |
| **Sekkumiste liik (füüsiline, vaimne või andmed, sh eriliiki isikuandmed)** | Uuringus kasutatakse pseudonüümitud eriliiki andmeid (geeniandmed, terviseandmed). | |
| **Koormus uuritavale (kontaktivõtmise viisid, visiitide arv, uuringute tüüp ja arv, kutsete saatmise kordus jms)** | Uuritavatega kontakti ei võeta. | |
| **12. Koeproovide väljastamine kolmandatele osapooltele (RNA, DNA, plasma vms)** | | |
| **Mitme geenidoonori koeproove ja mis tüüpi koeproove väljastatakse?** | - (koeproove ei väljastata) | |
| **Kui palju ühe geenidoonori kohta koeproove väljastatakse?** | - | |
| **Kuhu koeproov väljastatakse (riik, asutuse nimetus, aadress)?** | - | |
| **Mida tehakse järelejäänud koeproovidega (kas ülejääk hävitatakse või saadetakse tagasi)?** | - | |
| **13. Uuringu eetiliste aspektide analüüs (3600 tähemärki, kuni 2 lk).**  Kõik uuringud, mille objektiks on inimesed, peavad olema läbi viidud, arvestades eetilisi nõudeid, eelkõige autonoomia austamise, heategemise ja kahju vältimise ning õigluse printsiipe. (<https://www.coe.int/en/web/bioethics/guide-for-research-ethics-committees-members>).  **vt ka** <https://www.etag.ee/wp-content/uploads/2020/01/Eetika_Tabel_EST_2020.pdf> | | |
| 1. Autonoomia respekteerimine   Analüüsitavad andmed pärinevad isikutelt, kes on vabatahtlikkuse alusel hakanud TÜ EGV geenidoonoriks ning kelle andmeid kasutatakse ainult pseudonümiseeritud kujul. Vabatahtlikud on enne geenidoonoriks hakkamist tutvunud geenidoonoriks saamise tingimustega, millega on tagatud autonoomia respekteerimise printsiip. Geenidoonoriks saamise nõusoleku vormi punkt 1 ütleb järgmist: „Geenivaramu eesmärk on koguda teavet Eesti rahvastiku tervise ja pärilikkuse kohta ning luua ja täiendada selle põhjal andmetega andmekogu “Geenivaramu“. Geenivaramut võib kasutada üksnes teaduslikuks uurimistööks, geenidoonori haiguste uurimiseks ja raviks, rahva tervise uurimiseks ja statistilistel eesmärkidel. Geenivaramu kasutamine muul otstarbel, eriti tsiviil- või kriminaalprotsessis tõendite kogumiseks või jälitustegevuseks, on keelatud. Geenivaramu abil tehtavad uuringud on jätkuvad ja ei ole piiratud käesoleva aja teaduse tasemega.“ Sellega on isikuid informeeritud, et geenivaramus olevaid andmeid kasutatakse teadustööks ning need uuringud on jätkuvad ja ei ole piiratud käesoleva aja teaduse tasemega. Geenidoonoritel on ka pärast geenidoonoriks hakkamist võimalus igal ajal nõusolek tagasi võtta (Geenidoonoriks saamise nõusoleku vorm punkt 12).  Geenidoonoriks saamise nõusoleku vorm: [https://www.riigiteataja.ee/aktilisa/1170/4201/9016/SOM\_11042019\_m39lisa.pdf#](https://www.riigiteataja.ee/aktilisa/1170/4201/9016/SOM_11042019_m39lisa.pdf)   1. 2. Võrdse kohtlemise printsiip on tagatud sellega, et uuringusse kaasatakse kõikide geenidoonorite andmed ja seega võetakse arvesse kõikide uuritavate võimalikke riskitegureid.   3. Mittekahjustamise printsiip  Analüüsitavad andmed pärinevad isikutelt, kes on vabatahtlikkuse alusel hakanud TÜ EGV geenidoonoriks ning kelle andmeid kasutatakse ainult pseudonümiseeritud kujul. Planeeritav uuring on kooskõlas mittekahjustamise printsiibiga, kuna uuring ei koorma geenidoonoreid ning ei põhjusta riske nende tervisele.  4. Heategemise printsiip  Antud uuringu fundamentaaluuringu komponendist saadav kasu on uuritavatele pigem kaudne, kuna otsene kliinilises praktikas kasutatav kasu võib tekkida alles mitme aasta pärast. Seega saavad antud tööst kasu pigem tulevased patsiendid, samuti on uurimistöö tulemustest kasu tervetel inimestel haiguste ennetamisel ja varajasel tuvastamisel.  Käesoleva uurimistöö riskiskooride valideerimise uuringutest, mis on pigem TRL tasemetel 7-8 on aga kasu potentsiaalselt üsna kiire. Nimelt võimaldab uus taristu paremini läbi viia uute turule tulevate geneetilisi riskiskoore kasutavate meditsiiniseadmete valideerimist vahetult enne nende juurutamisele eelnevat otsustamist. Valideerimise eesmärk on ennekõike hinnata nende seadmete sobivust Eesti rahvastiku geneetilise kompositsiooniga ning vajadusel tuvastada ka piirväärtusi, mille puhul vastavad skoorid ei oleks enam kasutatavad. Eesti geenidoonorite andmete kasutamine on sisuliselt ainus realistlik võimalus selliste mujal arendatud skooride ja seadmete turule toomise eelseks valideerimiseks. Sobivad seadmed (tarkvaralised algoritmid) saab selle teadmuse baasil juurutada personaalmeditsiini taristule. | | |
| **13 a Inimesed** | | |
| **Abiküsimused** | **Ei** | **Jah** |
| **Kas uurimisobjektiks on inimesed?**  **Jah** |  | **Vajadusel kirjeldada, kuidas tagatakse uuringus osalemise vabatahtlikkus ja välditakse uuringus osalejate mistahes lubamatut mõjutamist uuringus osalemiseks.**  Kõik geenidoonorid on astunud geenidoonoriks vabatahtlikult ning andnud informeeritud nõusoleku, milles muuhulgas annavad nõusoleku oma andmete kasutamiseks teadustöös. |
| **Kas uurimisobjektiks on haavatavad isikud või isikute grupid?** | **Ei** | 1. **Nimetada, millisesse haavatavate gruppi uuritavad kuuluvad ning milles seisneb nende haavatavus.** 2. **Vajadusel, st kui uuring viiakse läbi tuginedes isiku nõusolekule, kirjeldada informeeritud nõusoleku saamise protseduuri. Kui uuring tugineb nõusolekule, lisada nõusoleku vorm. Need tegevused peavad tagama, et isikud saavad aru uuringus osalemisega kaasnevatest riskidest.** |
| **Kas uurimisobjektiks on isikud, kes ei saa ise anda teadlikku nõusolekut uuringus osalemiseks (sh piiratud teovõimega isikud)?** | **Ei** | **Kirjeldada, kuidas informeeritakse ja saadakse eestkostjalt või seaduslikult esindajalt nõusolek piiratud teovõimega, sh laste uuringus osalemiseks.** |
| **Kas uurimisobjektiks on alaealised?** | **Ei** | 1. **Esitada osalejate vanust puudutav info.** 2. **Kirjeldada alaealise vanematelt osalemisnõusoleku saamise protseduuri.** 3. **Kirjeldada, kuidas küsitakse alaealise nõusolekut või arvestatakse alaealise arvamust võrdeliselt tema ea ja küpsusastmega.** 4. **Kirjeldada, kuidas tagatakse alaealiste heaolu.** 5. **Selgitada, mis põhjusel kaasatakse uuringusse alaealisi.**   Geenidoonoriteks on isikud alates 18. eluaastast. |
| **Kas uurimisobjektiks on patsiendid?**  Jah |  | 1. **Selgitada, missugune haigus/seisund/puue on uuritavatel.**   Geenidoonorid ei olnud enamasti värbamise ajal patsiendi staatuses. Kuna nende hulgas on erinevate haigustega isikud, siis EGV andmebaasi täiendamisel teistest registritest võisid andmed pärineda tervishoiuteenuse osutamisega seotud perioodist ning sellel ajal olid geenidoonorid patsiendid.   1. **Kirjeldada värbamise, kaasamise ja väljajätmise kriteeriume ning informeeritud nõusoleku saamise protseduuri.**   Käesoleva uuringu tarbeks eraldi uusi geenidoonoreid ei värvata ning andmete linkimist läbi ei viida.   1. **Kirjeldada, mida tehakse juhuleiu juhusliku tulemuse korral ja kuidas informeeritakse sellest uuritavaid.**   Uuritavaid ei ole plaanis juhuleidudest informeerida. Kui see siiski osutub otstarbekaks, esitatakse uuringutaotluse täiendamiseks EBIN-le uus jätkutaotlus. |
| **Kas uurimistöös kogutakse inimestelt bioloogilisi proove? Kas inimestelt võetud bioloogiliisi proove kavatsetakse eksportida kolmandasse riiki (**[**https://www.aki.ee/et/teenused-poordumisvormid/andmete-edastamine-valisriiki**](https://www.aki.ee/et/teenused-poordumisvormid/andmete-edastamine-valisriiki)**) või importida neid teisest riigist Eestisse?** | **Ei** | 1. **Kirjeldada, missuguseid proove kogutakse.** 2. **Kirjeldada proovide kogumiseks kasutatavaid protseduure sh kas kasutatakse varem kogutud proove.** 3. **Selgitada, kuidas tagatakse uuritavate õigused.** 4. **Selgitada, mida tehakse proovidega pärast uuringu lõppemist.** |
| **13 b Isikuandmed ja andmestikud** | | |
| **Kas uurimistöö käigus kogutakse või analüüsitakse isikuandmeid, sh eriliiki isikuandmeid?**  Jah |  | Analüüsitakse tervise- ja geeniandmeid. |
|  |  | 1. **Täpne andmete koosseis, mis on vajalik uuringu läbiviimiseks (võib esitada lisana).**   Tulenevalt projekti eesmärgist kasutatakse uuringus kõikide geenidoonorite andmeid, mis on nende kohta EGVs olemas (fenotüübi andmed ja digitaalsed DNA andmed) ja mis lisanduvad uurimisprojekti jooksul. Andmeid väljastatakse uurijatele pseudonüümitud kujul.   1. **Kinnitada, et informeeritud nõusolek on olemas või saadakse enne uuringu algust, kui uuring põhineb nõusolekul.**   Kinnitame, et kõik geenidoonorid on geenidoonoriks astumisel allkirjastanud informeeritud nõusoleku vormi.   1. **Selgitada, miks on kõik töödeldavad andmed asjakohased ja vajalikud (lähtudes andmete minimeerimise põhimõttest).**   Käesolev projekt uurib erinevate (nii olemasolevate kui ka loodavate) riskiskooride ja terviseandmete seoseid. Riskiskoore on maailmas juba tuhandeid, ainuüksi <https://www.pgscatalog.org/> kataloogis on neid taotluse esitamise seisuga 738, paljud neist kasutavad kümneid tuhandeid erinevaid geenimarkereid üle kogu genoomi. Seetõttu ei saa käesoleva uuringu jaoks määratleda kindlat geeniregiooni, vaid taotleme andmeid kogu genoomi kohta. Uuringus otsime muuhulgas riskiskooride seni teadmata seoseid haigustega. Seetõttu on võrdluseks vajalikud ka geenidoonorite täielikud fenotüübi kirjeldused. Adekvaatse tulemuse saamiseks ei ole mõistlik piirata uuringus osalevate geenidoonorite hulka, vaid kasutada kõikide geenidoonorite andmeid. Vastasel korral võib konkreetsete seoste puhul puudu jääda statistilisest võimsusest. Kasutatavad andmed on pseudonüümitud, s.t otsest tuvastamist võimaldavaid andmeid ei väljastata (sünni- ja surmaajast väljastatakse vaid kuu ja aasta, kontaktandmeid ei väljastata, elukohast väljastatakse vaid maakond).  Lisaks on maailmas OHDSI ([Observational Health Data Sciences and Informatics](https://ohdsi.org/)) OMOP CDM ühise andmemudeli baasil loodud palju fenotüüpide ja haiguste või meditsiiniliste seisundite kohortide kirjeldusi (vt näiteks <https://data.ohdsi.org/PhenotypeLibrary/>). Need kiiresti kasvavad kirjeldused/definitsioonid võivad kasutada erinevaid otseseid või kaudseid parameetreid ning vastava info puudumine võib seada ohtu uuringute korratavuse Eesti andmetel.   1. **Kas andmesubjektid on tuvastatavad? Kui jah, siis kirjeldada, kuidas on täidetud järgmised tingimused:**    1. **pärast tuvastamist võimaldavate andmete eemaldamist ei ole andmetöötluse eesmärgid enam saavutatavad või neid oleks ebamõistlikult raske saavutada;**    2. **teadus- või riikliku statistika tegija hinnangul on selleks ülekaalukas avalik huvi;**    3. **töödeldavate isikuandmete põhjal ei muudeta andmesubjekti kohustuste mahtu ega kahjustata muul viisil ülemäära andmesubjekti õigusi.**   Andmesubjektid ei ole otseselt tuvastatavad. Kuigi uuringus kasutatakse geeniandmeid, ei ole uuringumeeskonnal üksnes geeniandmete põhjal võimalik isikuid tuvastada, sest selleks tuleks kõrvutada andmestikku mingi teise, isikustatud geeniandmekoguga, mida aga uuringus ei kasutata. Terviseandmete põhjal isiku tuvastamine on väga vähetõenäoline, kuna väljastatakse suur hulk andmeid ning konkreetse isiku tuvastamine isegi mingi harvemesineva seisundi alusel pole ilma suurte pingutusteta võimalik.  Lisaks ei väljastata terviseandmete detailseid kuupäevi (andmed väljastatakse kuni kuu täpsusega).  Andmehalduse printsiipe järgides välistame andmestike illegaalse kõrvutamise võimalused. Lisaks jälgime andmelogisid, et tuvastada võimalikke probleeme eos.  Juurdepääs väljastatud andmetele on vaid nendele uurijatele, kes on nimetatud taotluses (antud uuringu uurijad). Uute uurijate lisamiseks esitatakse taotlus EBINile.  Kõikide uurijatega sõlmitakse konfidentsiaalsusleping. |
| **Kas uurimistöö hõlmab üksikisiku süsteemset jälgimist, tema andmeprofiili kogumist või töödeldakse suures ulatuses eriliiki ja/või tundlikke andmeid või kasutatakse (sekkuvaid) andmete töötlemise meetodeid varjatud viisil (nt elulemuse uuringud, jälgimine, järelevalve, audio ja video salvestamine, geo- positsioneerimine jne) või mistahes andmete töötlemise protsessi, mis võib kahjustab uuritavate õigusi ning vabadust?**  Jah |  | 1. **Selgitada, missuguseid meetodeid kasutatakse uuritavate jälgimiseks, järelevalveks ja vaatlemiseks.**   Uuringus uuritakse seoseid geeniandmete ja fenotüüpide vahel, selleks töödeldakse suures mahus eriliiki andmeid. Otseselt üksikisiku andmed uuringus huvi ei paku. Masinõppe lahenduste puhul, kui viiakse läbi ennustamised üksiku andmesubjekti tasemel, toimub nende ennustuste analüüs siiski andmekohordi tasemel, näiteks ennustuse kvaliteedi mõõtmisel kasutatakse ROC graafikuid või visuaalseid punktipilvi, kus iga indiviid küll kontributeerib iga konkreetse “punkti”, kuid indiviid ise pole oluline teadusliku järelduse tegemisel.   1. **Selgitada uuritavate profiili loomise meetodeid.**   Uuritavate andmete põhjal luuakse geneetilisi riskimudeleid, mis jagavad uuritavad erinevatesse riskigruppidesse (näiteks geneetiline risk haigestuda mõnda haigusesse).   1. **Selgitada, kuidas informeeritakse uuritavaid nende õigustest ja võimalikest riskidest, mida andmete töötlemine võib kaasa tuua.**   Uuritavaid eraldi ei informeerita. Uuritavad on andnud informeeritud nõusoleku nende andmete kasutamiseks teadusuuringutes geenidoonoriks astumisel.   1. **Selgitada, kuidas toimub uuritavate profiili jaoks andmete kogumine ning nende teavitamine võimalikest tagajärgedest ja kaitsemeetmetest.**   Uuritavate andmeid eraldi koguma ai hakata, kasutatakse juba olemasolevaid andmeid. |
| **Kas uurimistöös analüüsitakse eelnevalt kogutud isikuandmeid?**  Jah |  | 1. **Selgitada, missugusest andmebaasist (registrist, andmekogust) või allikast andmed pärinevad.**   Andmed pärinevad Eesti geenivaramust. Kasutatakse geenidoonorite pseudonüümitud andmeid.   1. **Selgitada, kuidas informeeritakse uuritavaid nende õigustest ja võimalikest riskidest, mida andmete töötlemine võib kaasa tuua.**   Uuritavaid eraldi ei informeerita. Uuritavad on andnud laiapõhjalise informeeritud nõusoleku uuringuteks geenidoonoriks astumisel.   1. **Selgitada, miks on kõik töödeldavad andmed asjakohased ja vajalikud (lähtudes andmete minimeerimise põhimõttest).**   Käesolev projekt uurib erinevate (nii olemasolevate kui ka loodavate) riskiskooride ja terviseandmete seoseid. Riskiskoore on maailmas juba tuhandeid, ainuüksi https://www.pgscatalog.org/ kataloogis on neid taotluse esitamise seisuga 738, paljud neist kasutavad kümneid tuhandeid erinevaid geenimarkereid üle kogu genoomi. Seetõttu ei saa käesoleva uuringu jaoks määratleda kindlat geeniregiooni, vaid taotleme andmeid kogu genoomi kohta. Uuringus otsime muuhulgas riskiskooride seni teadmata seoseid haigustega. Seetõttu on võrdluseks vajalikud ka geenidoonorite täielikud fenotüübi kirjeldused. Adekvaatse tulemuse saamiseks ei ole mõistlik piirata uuringus osalevate geenidoonorite hulka, vaid kasutada kõikide geenidoonorite andmeid. Vastasel korral võib konkreetsete seoste puhul puudu jääda statistilisest võimsusest. Kasutatavad andmed on pseudonüümitud, s.t otsest tuvastamist võimaldavaid andmeid ei väljastata (sünniajast väljastatakse vaid kuu ja aasta, kontaktandmeid ei väljastata, elukohast väljastatakse vaid maakond).  Lisaks on maailmas OHDSI (Observational Health Data Sciences and Informatics) OMOP CDM ühise andmemudeli baasil loodud palju fenotüüpide ja haiguste või meditsiiniliste seisundite kohortide kirjeldusi (vt näiteks <https://data.ohdsi.org/PhenotypeLibrary/>). Need kiiresti kasvavad kirjeldused/definitsioonid võivad kasutada erinevaid otseseid või kaudseid parameetreid ning vastava info puudumine võib seada ohtu uuringute korratavuse Eesti andmetel.   1. **Selgitada, miks ei ole võimalik uurida uurimisobjekte nii, et saadud andmed oleksid anonüümsed või pseudonüümsed (kui on asjakohane).**   Andmed on pseudonüümitud. Koodivõti on ainult EGV-l kui vastutaval töötlejal. Andmete seostamist isikuga ehk depseudonüümimist ei teostata. Uurijatele ei saa väljastada anonüümseid andmeid, kuna uuring on jätkuva iseloomuga ning uurijatele väljastatakse ka uuringuperioodi jooksul lisanduvaid andmeid. |
| **Kas uurimistöös analüüsitakse avalikult kättesaadavaid andmeid?**  Ei |  | **Selgitada, kas andmed on avalikult kättesaadavad (avatud andmete registrid ja andmekogud) ja neid võib uurimistöös vabalt kasutada.**  Uurimistöös kasutatakse teiste uuringute baasil loodud riskimudeleid, millest paljud on avalikult kättesaadavad andmebaasidest. Kuid isikute/indiviidide tasemel ei ole kasutatud isikuga seotud andmeid. |
| **Kas kavatsetakse edastada isikuandmeid või võimaldada neile juurdepääs kolmandast riikidest (**[**https://www.aki.ee/et/teenused-poordumisvormid/andmete-edastamine-valisriiki**](https://www.aki.ee/et/teenused-poordumisvormid/andmete-edastamine-valisriiki)**)?** | Ei | 1. **Selgitada, missuguseid isikuandmeid eksporditakse või imporditakse. Kui jah, siis millistesse ja millistest riikidest.** 2. **Selgitada, milliseid kaitsemeetmeid rakendatakse, millisetel alustel (leping vms) andmeid edastatakse ning kuidas tagatakse uuritavate õigused.**   Isikuandmeid antud uuringu käigus ei ekspordita ega impordita. Küll aga juurutatakse mujal loodud otsusetoe ja riskiskooride algoritme, et valdieerida neid Eesti andmetel. Sama kehtib ka vastipidi, Eestis loodud uued meetodid ja riskiskoorid on vajalik valideerida ka teistel populatsioonidel. Sel juhul ei ekspordita individuaalseid andmeid Eesti isikute kohta vaid ainult agregeeritud statistilisi koondeid või masinõppe mudeleid, kus ei ole enam indiviidi tasemel infot, vaid ainult võimalikud tuletatud reeglid ja valemid. Andmeanalüüsi viiakse läbi ennekõike rahvusvaheliste võrdlusuuringute käigus EU partneritega (EU IMI projektid EHDEN, OPTIMA, ja covid andmete statistilised uuringud). |
| **Kas uurimistöö lõppedes toimub isikuandmete hävitamine/ anonüümimine?**  Jah |  | **Kui analüüs põhineb jooksvalt andmesubjekte tuvastada võimaldaval kujul**   1. **kirjeldada, kuidas pärast uuringu läbiviimist ja eesmärkide saavutamist isikuandmed hävitatakse/ anonüümitakse;**   Pärast uuringu läbiviimist isikuandmed hävitatakse. Analüüsiserveris ja juhul, kui mõnest geenitestist arendatakse edasi meditsiiniseade, siis ka meditsiiniseadme kohustuslikus projektdokumentatsioonis säilitatakse agregeeritud tulemid, mis on vajalikud teaduspublikatsioonide ja riskiskooride valiidsuse tõestamiseks 5 aastaks pärast projekti lõppu või riskiskooride kliinilises kasutuses oleku puhul nende kasutuse lõpuni. Nendest agregeeritud tulemitest üksikisikute tuvastamine pole võimalik.  Teadustöö tulemusena saadud andmed tagastab uurimisgrupp Geenivaramusse, kus need liidetakse Geenivaramu andmebaasiga ja säilitatakse tähtajatult IGUS § 18 alusel.  Eraldi tuleb lahendada küsimus võimalikust kohustusest säilitada konkreetse loodud meditsiiniseadme algoritmi treenimise aluseks olnud andmete pika-ajalisest säilitamisest juhuks, kui on vaja auditeerida võimalikke ennustusvigu (nt miks andis algoritm ohtliku hoiatuse; kas tegu on veaga algoritmi loomisel ning kas on olnud analoogseid juhtumeid kus tulem on olnud vastupidine). Nimetatud lahenduse kasutuselevõtmiseks küsitakse eelnevalt hinnang EBINilt.   1. **lisada hinnang, kuidas on välistatud andmesubjektide kaudne tuvastamine pärast isikut otseselt tuvastada võimaldavate andmete hävitamist.**   Konkreetsete uuritavatega seostatavad geeni- ja terviseseisundi andmed uuringu lõpus hävitatakse. Säilitatakse agregeeritud tulemid, mis on vajalikud teaduspublikatsioonide ja riskiskooride valiidsuse tõestamiseks, kuid millest üksikisikute tuvastamine pole võimalik (tagatud k-anonüümsus). |
| **13 c Teised eetilised küsimused** | | |
| **Kas uurimistöö läbiviimine võib kaasa tuua eelpool kirjeldamata eetilisi riske?** | Ei | **Vajadusel selgitada eetilisi lisariske, mis võivad tuleneda asjaoludest nagu tehisintellekt, personaalmeditsiin, sõjaliste partnerite kaasamine, uued arengud neurobioloogias, geenitehnoloogias, nanotehnoloogias, inimese-masina suhtluses, androidide ja küborgide loomisel jne.**  Personaalmeditsiini ja tehisintellekti riskid tulenevad peamiselt kolmest asjaolust - kaasatusest (kas igaüks kes soovib saada vastavat teenust saab liituda geenidoonoriks?), kas samad teenused on saadaval ka mitte-geenidoonoritele? (riiklik programm loob eelduse ka mitte-geenidoonorite kaasamiseks), ning kolmandaks tehisintellekti enda arusaadavus ja kallutamatus.  Kallutatuse näiteks on algoritmid mis baseeruvad kallutatud andmetel. Eestis näiteks baseeruvad enamus algoritmid eeldatavalt kohaliku geneetilise kompositsiooniga inimeste andmetel, seetõttu võivad teise geneetilise taustaga isikud saada teoorias ennustamatuid tulemusi. Sellise kallutatavuse tuvastamine, riskide hindamine ja ennetamine on olulised. Kuid selliseks ennetustööks on selgelt vajalik kõikide andmete kaasamine maksimaalselt kallutamatute ja tervisele ohutute ennustuste tedemiseks või ennetavaks andmesubjektide välistamiseks. |
| **14. Täita, kui uuring põhineb andmekogu ja/või andmeallika andmetel.** | | |
| **Andmekogu ja/või andmeallika nimetus:** Eesti geenivaramu | | |
| **Isikuandmete töötlemise eesmärk:** Teadus- ja arendusuuring | | |
| **Andmekoosseis ja periood, mille kohta andmed kogutakse (vajadusel lisana)**  Uuringu andmekoosseisu moodustab kõikide geenidoonorite kohta (ka uuringuperioodil lisandunud) Eesti Geenivaramus olemasolev info järgmiste andmeliikide kohta:   * Doonori demograafilised andmed (sünniaasta, sugu, haridus, elukoht maakonna tasemel, surma kuupäev) * Doonori kõik haigused (RHK-koodid ja nimetused koos kuupäevaga, info kirje allika kohta (ise raporteeritud, ambulatoorne epikriis, statsionaarne epikriis, raviarve, retsept vms)) * Doonori kõik ravimid ja retseptid (allikas, RHK-kood, nii välja kirjutatud kui välja ostetud ravimi kohta: kuupäev, ATC kood, kogus, pakendite arv, ravimi pakendi kood, annustamise info) * Doonorile tehtud meditsiinilised protseduurid, operatsioonid (Haigekassa teenusekood, NOMESCO (NCSP) kood (olemasolu korral muud koodid), kuupäev) * Doonori laborianalüüside vastused (info kirje allika kohta, kuupäev, LOINC kood, analüüt, mõõteväärtus, ühik, referentsvahemik) * Doonori ambulatoorsete ja statsionaarsete raviepisoodide üldinfo (info kirje allika kohta, aeg, liik) * Doonori allergiate info koos kuupäevaga * Doonori üldiste mõõtmiste info: kehakaal, pikkus, kehamassiindeks, kas suitsetab (kõik koos kuupäevaga) * Doonori mikrokiipidega mõõdetud geeniandmed koos imputeeritud andmetega   Uuringus kasutatava andmestiku uuendamine/värskendamine sõltub andmete üldise uuendamise rütmist Eesti Geenivaramus. Eeldatavalt poole seda mõistlik teha sagedamini kui kord poolaastas ega ka harvemini kui kord aastas. Seega kokkuvõtvalt keskmiselt 6-12 kuu intervalliga. | | |
| **15. Isikuandmete kaitse meetmete kirjeldus, sealhulgas andmete hoidmise, säilitamise, turvalisuse ja kustutamise kohta, sh andmete ja/või koodivõtme kustutamise kuupäev (kuni 1800 tähemärki, 1 lk).** | | |
| **Kirjeldada ja põhjendada uuringu vajaduseks kogutud andmete säilitamist ja tähtaega.** | Andmeid säilitatakse kuni uuringu lõppemiseni (detsember 2026). Pärast uuringu lõppu säilitatakse vaid agregeeritud kujul andmeid, mis on vajalikud teaduspublikatsioonide või riskiskooride valiidsuse tõestamise säilitamiseks.  Teadustöö tulemusena saadud andmed tagastab uurimisgrupp Geenivaramusse, kus need liidetakse Geenivaramu andmebaasiga ja säilitatakse tähtajatult IGUS § 18 alusel. | |
| **Kirjeldada isikuandmete pseudonüümimise protsessi ja vahendeid.** | Geenidoonorite andmed on pseudonüümitud juba geenivaramuga liitumisel. | |
| **Kas kavatsetakse geenidoonorite isikuandmeid depseudonüümida?**  Ei | 1. **Täpsustada, mitme geenidoonori andmed depseudonüümitakse.** 2. **Selgitada, mis otstarbel andmed depseudonüümitakse.** | |
| **Kas toimub isikuandmete transportimine ning kirjeldada, kuidas on tagatud andmete turvalisus.** | Andmed edastatakse uuringumeeskonnale Tartu Ülikooli teadusarvutuskeskuse kaudu. Geenivaramu valmistab ette väljastuspaketi ning paigutab selle turvalisse teadusarvutuskeskuse pilve. Uuringumeeskond töötleb andmeid Teadusarvutuskeskusesse paigaldatud spetsiaalses ainult uuringumeeskonnale kättesaadavas uurimisserveris. | |
| **Kirjeldada, kuidas on andmed kaitstud loata või ebaseadusliku töötlemise eest.** | Andmed paigaldatakse Tartu Ülikooli teadusarvutuskeskusesse (HPC) spetsiaalselt selleks loodud ainult uuringumeeskonnale kättesaadavasse uurimisserverisse. | |
| **Kinnitan, et kõik uuringu läbiviijad on teadlikud projekti läbiviimisega kaasnevatest eetilistest ja isikuandmete kaitsega kaasnevatest nõuetest.** | | |
| **Vastutava uurija allkiri**  **/*digiallkiri*/** | **Taotluse esitamise kuupäev**  23.03.2021  Taotlust täiendatud ja esitatud 06.09.2024 | |
|  | | |
| **Taotluse EBIN ID**  **(täidab hindaja)** | | |
|  | | |

**Lisadokumentide loetelu:**

1. **Vastutava uurija CV**