

# ANDMEPÄRINGU TÄPSUSTUSED

*Riskifaktorite mõju ägeda müokardiinfarktiga patsientide haiglaravile ja haiglajärgsele käsitlusele Eestis*

Health Data Unit

Metrosert AS

Käesolevas dokumendis esitatakse täiendavaid selgitusi andmekaitespetsialisti küsimustele, mis hõlmab (1) andmete minimeerimist, (2) andmeturbe ja andmetöötuse meetmeid ning (3) andmekaitsealast mõjuhinnangut. Kui antud teema on EBIN-i taotluses adresseeritud, viidatakse vastavale peatükile ning esitatakse lühikokkuvõte.

## **1. Andmete minimeerimine – muutuja “Meditsiinisündmus” vajalikkus ning periood 2023-2025**

### **Muutuja “Meditsiinisündmus”**

Antud muutuja ei olnud algses EBIN-i taotluses esitatud, kuna me ei olnud taotluse esitamise ajal sellest teadlikud. Küsisime EMIR-i muutujaid, mida pidasime indekssündmuse iseloomustamiseks piisavaks. Alles hilisema arutelu käigus, kui olime EBIN-lt loa juba saanud, selgus, et EMIR hoiustab teatisepõhiseid andmeid ning ka detailsemat agregeeritud versiooni juhtumipõhistest andmetest. Kui teatisepõhised andmed sisaldavad mitmeid sama MI patsiendi kandeid, siis juhtumipõhised andmed on töödeldud, kus visiidid on iga MI juhu puhul agregeeritud.

Teatisepõhised andmed on vajalikud mitmel põhjusel. Üksikasjalikum haigusjuhtumipõhine informatsioon võimaldab iseloomustada ägeda MI raviteekonda, mis on käesoleva uurimuse peamiseks eesmärgiks. Lisaks, võimaldavad need hinnata EMIR-i andmete kvaliteeti, tuvastada andmelünku ja probleemkohti, mis on uurimuse teiseks eesmärgiks. Muutuja “Meditsiinisündmus” on sealjuures oluline sama MI juhu erinevate teatiste sidumiseks. Vastasel juhul ei saa teatiseid usaldusväärselt grupeerida, mis muudab uuritava haigusjuhu defineerimise ebatäpseks, võib moonutada haigestumise ja korduvhaigestumise näitajaid ning ei võimaldaks tunnuseid (nt LVEF, revaskularisatsioon, MI alatüüp) vastava episoodi või juhuga siduda.

Mis puudutab uuringu perioodi (2023-2025), siis teatisepõhised andmed on saadaval ning võimaldaksid projekti õigeaegset täitmist. Juhtumipõhised agregeeritud andmed on perioodi 2023-2024 kohta ning 2025.aasta analoogselt töödeldud andmete saamine võtaks märkimisväärselt aega, mis võib põhjustada viivitusi projekti lõpuleviimisel.

Muutuja “Meditsiinisündmus” ei sisalda kliinilist teavet, identifikaatoreid ega muud informatsiooni andmesubjekti kohta. See on üksnes sisemine grupeeriv ID, ilma milleta EMIR-i teatisepõhiseid andmeid ei oleks võimalik episoodideks või juhtudeks grupeerida. See muudaks indekssündmuse defineerimise hägusaks, kallutaks haigestumise ja korduvhaigestumise juhte ning puuduks võimalus tunnuseid (nt LVEF,

revaskularisatsioon, MI alatüüp) vastava juhuga siduda. Muutuja “Meditsiinisündmus” muudaks taotluses juba heakskiidetud EMIR-i muutujad MI uuringu jaoks tõlgendatavaks.

Oleme valmis kandma nimetatud muutuja EBIN-I taotluse lisasse kui tehnilise siduva muutujana juhul, kui Sotsiaalministeerium peaks seda pidama paremaks lahenduseks.

## **Periood 2023-2025**

Eestis on ligikaudu 2700 ägeda MI juhtu aastas. Selleks, et viia läbi planeeritud stratifitseeritud analüüsid (MI alatüüp, komorbiidsus, haiglad ning piirkond) ning arvestada 365-päevase indekseerimisjärgse jälgimisperioodiga, on vaja kolmeaastast vaatlusperioodi. Lühema vaatlusperioodi korral oleks statistiline võimsus liiga madal, et alarühmade lõikes analüüse läbi viia või piiraks jälgimisperioodi. Kolmeaastane periood on koosõlas ka Eesti MI haigla järgse ravijuhendiga (Eesti Arst, 2022) ning Tervisekassa raviteekonna kiirendiprogrammiga, mis on poliitika kujundamise seisukohalt oluline vaatlusperiood, mida antud uuring on kavandatud hindama.

## **2. Andmeturbe ja andmetöötluse meetmed**

Antud meetmed on täpsemalt kirjeldatud EBIN-I taotluse **peatükis 16** (“Description of personal data protection measures, including data storage, security and erasure”) ning **peatüki 13** vastavasisulistes alajaotustes (“Roles and responsibilities”, “Risks and mitigation measures”, “Data retention and disposal”). Peamised punktid on välja toodud allpool.

- **Andmete hoiustamine.** Andmete pseudonümiseerimist ning registritevahelist seostamist teostab TEHIK enne kui need jõuavad uurimisrühmani. TEHIK edastab pseudonümiseeritud andmed Tartu Ülikooli turvalisse töötuskeskkonda (SAPU), kus viiakse läbi kõik analüüsid.
- **Ligipääs.** Andmetele on ligipääs ainult ligipääsuõigusega uurijatel. SAPU on täielikult võrgust isoleeritud keskkond – volitatud kasutajad loovad ühenduse virtuaaltöölauaga üle krüpteeritud kanalite ning saavad andmeid töödelda vaid selles keskkonnas. SAPU-s puudub internetiühendus ja võimalus kopeerida/kleepida ning puudub võimalus väliste failide importimiseks. Ainus võimalus SAPU-st andmete väljaviimiseks on läbi kontrollitud ekspordiala, kus TEHIKu andmehaldur peab enne andmete väljastamist kinnitama nende vastavust anonümiseerimisnõuetele. Kõik keskkonnas tehtavad toimingud salvestatakse logis.
- **Säilitusperiood.** Pseudonümiseeritud indiviidi tasandi andmeid säilitatakse SAPU-s üksnes projekti kestel (01/2026–12/2030 EBIN peatükk 3). Projekti lõppemise järgselt hoitakse teaduspublikatsioonide ja korratavuse tagamise eesmärgil alles ainult koondatud ja täielikult anonümiseeritud analüüsitulemid.

- **Hävitamine.** Kõik pseudonümiseeritud indiviidi tasandi andmed kustutatakse projekti lõpus jäädavalt vastavalt andmevastutajate poolt heaks kiidetud andmehaldusplaanile. Andmete hävitamise viivad läbi Tartu Ülikooli turvalise SAPU keskkonna volitatud töötajad. Pärast analüüside läbiviimist ning tulemuste valideerimist säilitatakse üksnes agregeeritud ja täielikult anonümiseeritud statistilisi tulemusi. Andmevastutajad (TEHIK ja Tervisekassa) tagavad, et ühtegi pseudonümiseeritud kirjet ei jääks kättesaadavaks ning et pseudonümiseerimisvõtmed hävitatakse turvaliselt või arhiveeritakse piiratud tingimustel vastavalt isikuandmete kaitse üldmäärusele (GDPR) ja Eesti isikuandmete kaitse seadusele

Andmetöötlus SAPU keskkonnas on kooskõlas GDPR-i, Eesti isikuandmete kaitse seaduse ning ISO/IEC 27001 infoturbe põhimõtetega.

### 3. Andmekaitsealane mõjuhindang

Andmekaitsealast mõjuhindangut on kirjeldatud EBIN-i taotluse **peatükis 13** (“Analysis of the ethical aspects of the study”) ja **16** (“Personal data protection measures”), käsitledes proportsionaalsust, õiguslikku alust, riske ja nende maandamise meetmeid, jääkriske ning andmesubjekti õiguseid. Peamised punktid on välja toodud allpool.

- **Uurimuse olemus ja ulatus.** Retrospektiivne registripõhine uuring hõlmates ligikaudu 8100 ägeda MI patsienti (2023-2025) ning mille käigus kogutakse pseudonümiseeritud andmeid registritest TIS/Digilugu, EMIR, SPR, RETS ja KIRST.
- **Õiguslik alus.** Andmeid töödeldakse teadusuuringu eesmärgil isikuandmete kaitse seaduse (IKS) paragrahv 6 lõige 1 alusel. Isikud, kes on keeldunud oma andmete kasutamisest uuringu tarbeks, jäetakse valimist välja. Andmesubjektide õigused, sealhulgas õigus andmetega tutvuda, esitada vastuväiteid ja saada teavet, tagatakse vastavate andmevastutajate protseduuride kaudu (TEHIK, EHIF ja teised registrid).
- **Vajalikkus ja proportsionaalsus.** Muutujate hulk on piiratud minimaalse vajalikkuga, et rekonstrueerida iga patsiendi raviteekond ja hinnata selle vastavust riiklike ravijuhistega (vt lisatud muutujate loendit ning peatükki 1 eespool). Anonümiseerimine enne andmete sidumist ei ole teostatav, kuna registritevaheline andmete sidumine eeldab ühist pseudonümimise võtit ning ravi ajaline järjestamine eeldab indiviidi tasemel kronoloogiliselt järjestatud andmeid.
- **Riskid:** konfidentsiaalsuse rikkumine andmeedastusel; kaudne tuvastamine (andmete uuesti isikustamine) muutujate kombineerimise kaudu; volitamata juurdepääs; väikese sagedusega väärtuste esinemine tulemustes (kaudne tuvastamine).

- **Riskide maandamine:** pseudonümiseerimist teostavad üksnes andmevastutajad; andmete otseedastus SAPU keskkonda ilma vahekoopiateta; SAPU täielik võrgisolatsioon, krüpteeritud juurdepääs ja toimingute logimine; kuupäeva- ja asukohamuutujate üldistamine; andmete agregeerimise lävede rakendamine ning väikese sagedusega väärtuste maskeerimine kõigis väljundites, mida TEHIK sealjuures kontrollib enne nende väljastamist; indiviidi tasandi andmete ja pseudonümiseerimisvõtmete kustutamine projekti lõppedes.
- **Jääkriske** on hinnatud **madalaks:** puudub otsene kontakt uuritavatega, sekkumisi ei toimu, puudub äriiline eesmärk, andmeid ei edastata väljapoole EL-i/EMP-d ning indiviidi tasandi andmeid ei viida turvalisest andmetöötluskeskkonnast välja.