

Terviseandmete taaskasutuse tõhustamine Eestis

Kerli Mooses^{1,2}, Maarja Pajusalu^{1,2}, Sulev Reisberg^{1,2,3}, Raivo Kolde^{1,2}, Marek Oja^{1,2}, Riina Raudne⁴, Raili Sillart⁵, Alar Irs⁶, Silver Heinsar^{7,8,9}, Priit Pauklin^{6,7}, Priit Kampus^{7,8}, Jaak Vilo^{1,2,3}

Eesti Arst 2026;
105(2):86–90

Saabunud toimetusse:
08.12.2025
Avaldamiseks vastu võetud:
07.01.2026
Avaldatud internetis:
23.02.2026

¹ Tartu Ülikooli arvutiteaduse instituut,
² OHDSI Eesti,
³ STACC,
⁴ Tartu Ülikooli genoomika instituut,
⁵ Sotsiaalministeerium,
⁶ Tartu Ülikooli Kliinikumi südamekliinik,
⁷ Tartu Ülikooli kliinilise meditsiini instituudi südamekliinik,
⁸ Põhja-Eesti Regionaalhaigla sisekliiniku kardioloogiakeskus,
⁹ Queenslandi Ülikooli molekulaarse bioteaduse instituut

Kirjavahetajaautor:
Kerli Mooses
kerli.mooses@ut.ee

Võtmesõnad:
terviseandmete teisene kasutamine, OMOP, Est-Health-30, andmete väärindamine

Eestis koguneb igapäevase tervishoiuteenuse osutamise käigus suures mahus andmeid, mis sisaldavad väärtuslikku infot inimeste tervises seisundi, käitumise, ravi ja selle tulemuslikkuse kohta. Seoses tehniliste, õiguslike ja korralduslike probleemidega on andmete kasutus nii ravipraktika hindamisel, tervishoiupoliitika kujundamisel kui ka teadusuuringutes liiga tagasihoidlik.

Artiklis on kirjeldatud, kuidas tõhustada olemasolevate terviseandmete taaskasutust esindusliku ning lingitud andmestiku abil, mis koondab mitme riikliku tervise-registri andmed ühtseks tervikuks. Kvaliteetsed, standardiseeritud ja analüüsiks ette valmistatud andmed võimaldavad teha teadus- ja rakendusuuringuid ressursisäästlikult, kiiresti ja patsientide privaatsust austavalt. Varem loodud andmetike ja analüüsikoodide taaskasutamine aitab kiirendada uuringute läbiviimist ning edendab nii tõenduspõhist meditsiini kui ka Euroopa terviseandmeruumi põhimõtteid senisest ulatuslikumalt ja kvaliteetsemalt.

Eesti terviseandmekogudes talletatakse igapäevaselt suur hulk infot inimeste tervises seisundi, tehtud uuringute ja analüüsides ning nende tulemuste kohta. Selle teabe laialdasem taaskasutus teadus- ja arendustegevuseks aitaks luua täiesti uue väärtuse – see võimaldaks hinnata ravi kvaliteeti ja tõhusust, jälgida ravijuhendite rakendamist ning planeerida tervishoiuteenuseid üleriigiliste andmete põhjal. Saadud teadmisi saab kasutada ennetustegevuse planeerimiseks, patsiendikesksete tervishoiuteenuste osutamiseks, parema ravijärgimuse tagamiseks, kliiniliste rakendusuuringute planeerimiseks ning tervishoiusüsteemi efektiivsemaks korraldamiseks.

PEAMISED PROBLEEMID

1. Puudub võimalus saada põhjalikku ülevaadet olukorrast

Arstidel ja tervishoiukorraldajatel puudub põhjalik ülevaade erinevate haiguste levimusest ja esmashaigestumisest erinevates soo- ja vanuserühmades ning erinevate kaasuvate haiguste korral. Arstidel on tihti ülevaade oma patsientidest, kuid sageli puudub võimalus analüüsida süsteemi laiemalt – asutuse piires, piirkondlikult, üle-eestiliselt ning miks mitte ka võrdluses

teiste riikidega –, võttes samal ajal arvesse ka patsientide liikumist tervishoiuteenuse osutajate vahel. Seetõttu puudub enamiku haiguste puhul teadmine, kui palju tegelik diagnoosi- ja ravipraktika erineb kehtivatest kliinilistest juhenditest, milline on ravimijärgimus, millises etapis langeb patsient raviteekonnalt välja, kas esineb piirkondlikke erinevusi raviteekonnas, kuidas liiguvad patsiendid tervishoiusüsteemis teiste kaasuvate hädadega ning kuidas kõik need tegurid mõjutavad tervisetulemeid ja tervishoiukululusid.

2. Uute uuringute läbiviimine on kulukas ja ressursimahukas

Klassikalise lähenemise korral kogutakse andmeid ainult konkreetse uuringu tarbeks ning pärast uuringu läbiviimist andmed kustutatakse. Olemasolevate terviseandmete teiseseks kasutamiseks ettevalmistamine on aja- ja ressursikulukas ning eeldab tihedat koostööd andmeinseneride ja -teadlaste ning meditsiini valdkonna ekspertide vahel. Näiteks vajab eri allikatest pärit andmete ühendamine ja esmane töötlemine peamiselt andmeteadlaste ja -inseneride IT-tehnilist pädevust. Samas, andmete puhastamisel, analüüsimisel ja tulemuste tõlgendamisel on oluline kõigi kolme

valdkonna koostöö, et tagada tulemuste õigsus. Seega on iga üksiku uuringu jaoks eraldi andmestiku loomine ebaefektiivne. Esindusliku puhastatud terviseandmestiku taaskasutus kiirendab oluliselt analüüsitulemusteni jõudmist ning vähendab uuringu kogukulu.

3. Andmete teise kasutuse potentsiaal on laiemalt kasutamata

Euroopa terviseandmeruumi (*European Health Data Space*, EHDS) määrus (1) näeb tulevikus ette terviseandmete laiapõhjalise ja turvalise teise kasutuse teadus- ja innovatsioonieesmärkidel. Praegu puuduvad Eestis veel lahendused, mis seda igapäevaselt võimaldaks. Selle asemel, et jääda ootama lahenduste valmimist, saame juba piloteerida olemasolevate lingitud andmestike taaskasutamist, järgides kõiki turvalisuse, andmekaitse ja pseudonüümimise nõudeid. See mitte ainult ei loo võimalust terviseandmeid väärintada juba praegu, vaid aitab ka luua parimaid praktikaid ja tehnilist valmisolekut, mis saavad aluseks või sisendiks EHDSi põhimõtete rakendamisele.

LAHENDUS: LINGITUD JA ESINDUSLIK TERVISEANDMESTIK EST-HEALTH-30

Tartu Ülikooli (TÜ) terviseinformaatika töörühm (2) on loonud terviseandmete teise kasutamise meetodikate, sh uute tehisar- ja teiste andmeanalüüsi meetodite arendamiseks esindusliku terviseandmestiku – Est-Health-30. Selle eripäraks on võimalus kasutada puhastatud ja standardiseeritud andmestikku ka täiendavate uuringute tegemiseks – iga uue või spetsiifilisema uurimisküsimuse jaoks tuleb sõnastada uuringu eesmärgid ja meetodika, taotleda eetikakomitee luba ning andmete vastutava töötaja nõusolek.

Est-Health-30 andmestik on loodud viisil, mis ei võimalda teadlastel otsest isikute tuvastamist, kuid pseudonüümi alusel on tagatud andmete linkimise võimekus ning seeläbi andmete pidev uuendamine. Andmestikku kuulub juhuvalimi alusel 30% Eesti rahvastikust, kelle kohta on koondatud alates aastast 2012 järgmised andmed:

- patsientide raviarved,
- laborianalüüside vastused,
- epikriisid,
- retseptiravimite väljakirjutamise ja väljaostmise info,

- surma põhjused
- vähiregistri info.

Uurimisrühm on arendanud erinevaid meetodikaid nii andmete kvaliteedi parandamiseks kui ka analüüsiprotsessi kiirendamiseks. Üheks näiteks on epikriiside vabatekstist faktilise info eraldamine nii kehakaalu ja vererõhu näitajate, aga ka PAP (Papanicolaou) ja HPV (inimese papilloomiviirus) testide, ehkardiograafia ja EKG tulemuste kohta. Nii on järjest rohkem vabatekstina salvestatud infot viidud struktureeritud ehk analüüsis kergemini kasutatavale kujule. Tegemist on esimese andmestikuga Eestis, mis võimaldab saada ülevaate näiteks ehkardiograafia ja EKG näitajatest, luues seeläbi täiesti uued võimalused südameveresoonkonna haiguste ravipraktika ja ravi tulemuslikkuse analüüsimiseks päriselu andmete põhjal. Tänu masinõppe ja suurte keelemudelite tehisar- kiirele arengule on järjest rohkem ja täpsemalt võimalik vabatekstist kätte saada vajalik info ning seeläbi andmestikku täiendada. Kuna andmestikku käideldakse turvalises piiratud ühendustega keskkonnas, rakendatakse andmete ja inimeste privaatsuse kaitseks ainult lokaalseid keelemudeleid ja andmeanalüüse.

Andmete tõhusamaks kasutamiseks on puhastatud info viidud ühtsele andmekujule (OMOP CDM, ingl *Observational Medical Outcomes Partnership Common Data Model*) (3), mis võimaldab läbi viia analüüse, mis varem olid killustunud või suisa võimatud. Näiteks saab kombineerida tervise infosüsteemist pärit laborianalüüside vastuseid Tervisekassa raviarvete ja retseptikeskuse andmetega. Ühtne standardne andmemudel võimaldab ära kasutada rahvusvaheliste jõupingutustega arendatud tarkvaralisi lahendusi ning see kiirendab oluliselt kogu analüüsi protsessi. Lisaks võimaldab see hajusanalüüsi meetodi abil teha esinduslikke rahvusvahelisi võrdlusuuringuid (4–7), tagades samal ajal lokaalsete andmete privaatsuse.

Olemasoleva kvaliteetse lingitud andmestiku taaskasutamine lühendab uuringute läbiviimisele kuluvat aega mitme kuu või isegi aasta võrra (vt joonis 1). Samal ajal uuringuprotokolli kirjutamise ja eetikakomitee lubade taotlemisega on võimalik valmis kirjutada kood uuringu andmeanalüüsiks ja raporti koostamiseks. See tähendab, et peale loamenetluse lõppu saab kohe alustada analüüsiga.

Est-Health-30 andmestiku väärtus on võimalus oluliselt kiirendada uute uurin-gute läbiviimist esinduslikul ja laiapõh-jalisel terviseandmestikul. Efektiivsus tuleneb korduvate andmeväljastuse ja -puhastuse ning ettevalmistuse protsesside ärajäämisest. Kogu protsess lihtsustub lähitulevikus veelgi tänu mitmele eetika-komitee tööd reguleerivale seadusele (8, 9), mille tulemusel hakatakse kõikide riiklike terviseandmekogudega seotud uuringute taotlusi menetlema ühtses teaduseetika komitees. Est-Health-30 andmestiku tuge-vuseks on kindlasti ka selle turvalisus. Nimelt toimub kogu andmete puhastamine, eeltöötlus ja analüüs turvalises andmepuuris (SAPU), milles puudub võimalus infot välja kopeerida. Analüüsi tulemusi on võimalik turvapuurst välja tõsta ainult vastutava uurija kinnitusel spetsiaalse töövoos abil. Andmestikuga töötamiseks on rakendatud mitmeid erinevaid füüsilisi ja infotehno-loogilisi turvameetmeid, mis suurendavad oluliselt andmestiku turvalisust ja tagavad privaatsuse kaitse. Seega avab Est-Health-30 andmestik täiesti uue võimaluse tõhusamaks ja turvaliseks andmetike taaskasutamiseks teadusuuringute läbiviimisel, tagades samal ajal riikliku ja teaduseetilise kontrolli iga uurimisküsimuse spetsiifilise menetlemise teel eetikakomitee ja andmete volitatud töötajate poolt.

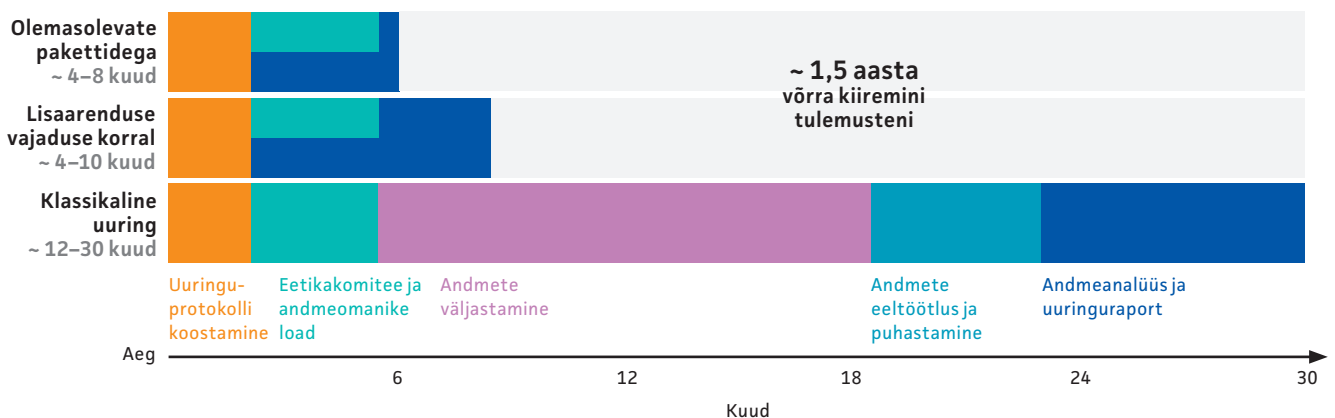
Kuigi Est-Health-30 andmestik on tervi-kuna veel komplekteerimisel, oleme juba esimese kasutusaasta jooksul esitanud koos erialaseltside ja eriarstidega mitmeid täiendavaid uurimisküsimuste taotlusi, sealhulgas põievähi, radioloogia, hüperten-

siooni jt südame-veresoonkonnahaiguste, reproduktiivtervise, infektsioonhaiguste, sünteetiliste andmete loomise, labori refe-reentsväärtuste harmoneerimise teemadel. See näitab, et huvi esindusliku, lingitud ja kvaliteetse andmestiku kasutamiseks on suur ning arstid ja teadlased näevad selles reaalselt väärtust erinevate haiguste uurimisel.

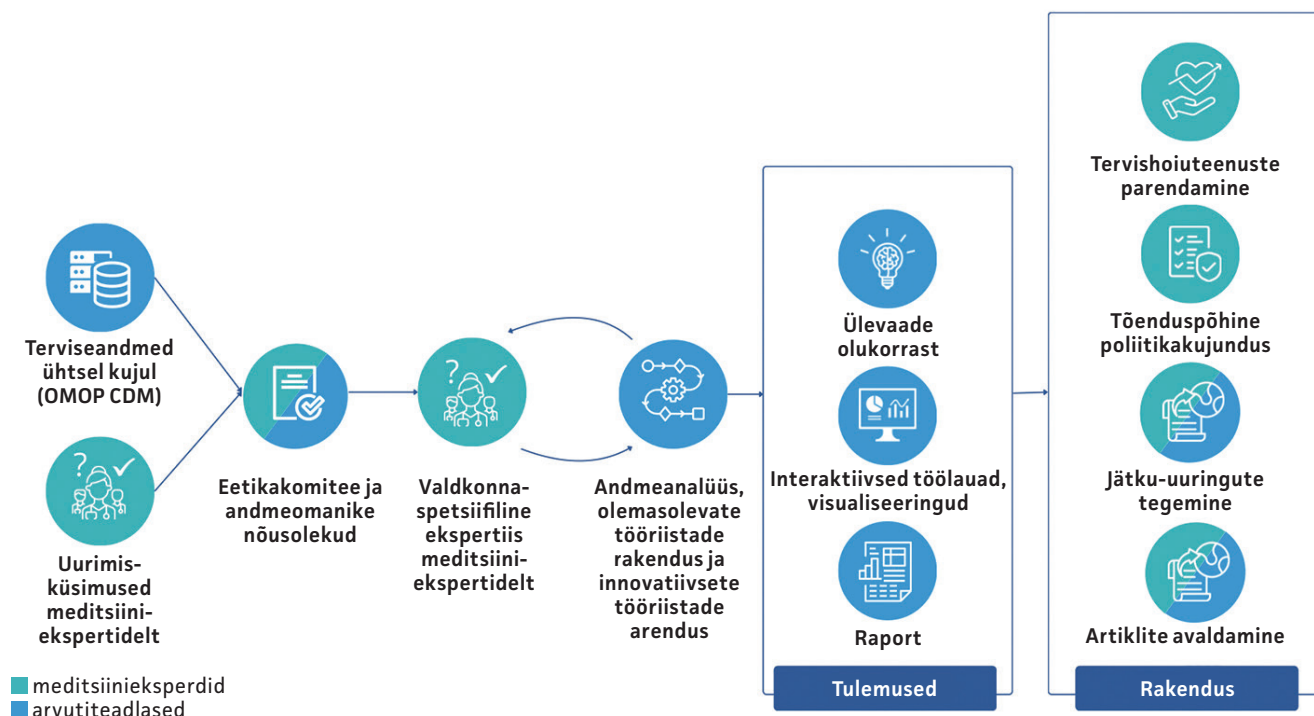
KOOSTÖÖPROTSESS JA -PÕHIMÕTTED

Est-Health-30 andmestiku kasutamine täiendavatele uurimisküsimustele vastami-seks saab parima lõpptulemuse saamiseks toimuda ainult tihedas koostöös erialaspet-sialistidega. Kõik meie täiendavad uurimis-küsimused pärinevad meditsiinivaldkonna ekspertidelt ning koostöös taotletakse uuringu läbiviimiseks vajalikud load ja nõusolekud (vt joonis 2). Andmeanalüüsi etapis rakendatakse sageli *study-a-thon*'i meetodit (3, 10), kus arstid, teadlased, epide-mioloogid, andmeanalüütikud ja andmekogu spetsialistid määratlevad ühiselt töötoas uuringusse kaasatud patsientide kohordid, huvipakkuvad teenused ja sündmused. Töötoa käigus saadakse kohe ülevaatlik kirjeldus loodud kohortidest ning sellest tulenevalt tehakse vajaduse korral kohen-dused. Samuti toimub ühiselt tulemuste tõlgendamine ning analüüsi täpsustamine.

Varem on Eestis sarnast lähenemist kasutatud Est-Health-30 andmestikuga analoogselt, kuid väiksemal andmestikul ning on analüüsitud COVID-19 pikaajalist mõju tervisele (11), emakakaelavähi enne-tusele suunatud ravijuhendi rakendumist



Joonis 1. Uuringu läbiviimise ajakulu klassikalise uuringu puhul ning valmis uuringuandmebaasi taas-kasutamisel, tuginedes senisele TÜ arvutiteaduse instituudi terviseinformaatika töörühma kogemusele.



Joonis 2. Est-Health-30 koostöö töövoog ja võimalikud väljundid.

(12), kõrge riskiga HPV levimust (13), südamepuudulikkuse diagnoosimise ja ravikäsitluse praktikat (14).

TÜ arvutiteaduse instituudi terviseinformaatika töörühm on Est-Health-30 andmestikuga avatud koostööle projektides, mida iseloomustavad järgmised tunnused:

1. **Oluline tervise teema ja ühiskondlik nõudlus** – käsitletakse probleemi, millel on lai mõju rahvastiku tervisele.
2. **Tugev partnerlus ja meditsiiniline ekspertsus** – koostööpartneril on idee ja valmisolek panustada oma teadmistega uurimisküsimuste sõnastamisse ja tulemuste analüüsi.
3. **Oluline mõju** – analüüsi eeldatavad tulemused loovad lisandväärtust, mis aitab haigusi ennetada, parandada ravi kvaliteeti, tõhusust või tervishoiusüsteemi toimivust.
4. **Rakenduslik väärtus** – tellijal või koostööpartneril on selge plaan, kuidas saadud tulemusi rakendada praktikasse, jätku-uuringutesse või poliitikakujundamisse.

Selline koostöömudel tagab, et iga uus analüüs panustab realselt inimkeskse tervishoiu arengusse ning aitab muuta andmepõhise teadustöö ja tervisepoliitika loomulikuks osaks Eesti tervishoiusüsteemis.

KOKKUVÕTE

Terviseandmete väärindamine ei ole vaid tulevikuvisioon – see on juba tänane võimalus ja lausa kohustus Eesti inimeste ees. Est-Health-30 esinduslik terviseandmestik pakub suurepärase koostöövõimalust viia teadusuuringuid läbi ressursisäästlikult, kiiresti ja turvaliselt ning panustada tõendus põhisesse tervishoiusüsteemi. Uute uurimisprojektide läbiviimisega saame täiendavaid kogemusi ja oskusi, et arendada edasi metoodikaid ning kasvatada ühiskonnale saadavat kasu.

RAHASTUS

Est-Health-30 andmestiku loomiseks vajalike metoodikate arendust on toetanud Eesti Teadusagentuur (projektid RITA1/02-96, PRG1844) ja Euroopa Liit ning kaasrahanud Haridus- ja Teadusministeerium (projekt TEM-TA72). Lisaks on tööd kaasrahanud Euroopa Regionaalarengu Fond (projekt nr 2021-2027.1.01.24-0444), Euroopa Liidu teadusuuringute ja innovatsiooni programm „Euroopa horisont“ (toetusleping 101060011) ning Haridus- ja Teadusministeerium seoses „Euroopa horisondi“ osaluse laiendamise meetmega „Teaming for Excellence“ koos Euroopa Regionaalarengu Fondi ja Ühtekuuluvusfondiga (projekti nr 2021-2027.1.01.24-0444).

KIRJANDUS

1. European Health Data Space, <https://www.european-health-data-space.com>.
2. TÜ terviseinformaatika töörühm, <https://health-informatics.cs.ut.ee>.
3. Reisberg S, Mooses K, Kolde R, Kõrgvee LT, Vilo J. Uudne lähenemine – OMOP-andmemudelil põhinevad terviseuuringud. Eesti Arst 2024;103:420–9.
4. Xie J, Du M, Guo Y, Barboza C, et al. Trends in prescription opioid use in Europe: A DARWIN EU® multinational cohort

- study including seven European countries. *Front Pharmacol* 2025;16:1608051.
- Gandaglia G, Pellegrino F, Golozar A, et al. Clinical characterization of patients diagnosed with prostate cancer and undergoing conservative management: a PIONEER analysis based on big data. *Eur Urol* 2024;85:457–65.
 - Burkard T, López-Güell K, Gorbachev A, et al. Calculating daily dose in the Observational Medical Outcomes Partnership Common Data Model. *Pharmacoepidemiol Drug* 2024;33:e5809.
 - Pineda-Moncusí M, Rekkas A, Martínez Pérez A, et al. Changes in use and utilisation patterns of drugs with reported shortages between 2010 and 2024 in Europe and North America: a network cohort study. *The Lancet Public Health* 2025;10:e835–47.
 - Teadus- ja arendustegevuse ning innovatsiooni korralduse seadus (17.06.2025). Riigi Teataja, <https://www.riigiteataja.ee/akt/112072025001>.
 - Inimgeeniuringute seadus, 749 SE eelnõu. Riigikogu, <https://www.riigikogu.ee/tegevus/eelnoud/eelnou/18ffea88-8a90-464f-b220-905cb3e1c4f0/Inimgeeniuringute%20seadus/>.
 - Hughes N, Rijnbeek PR, van Bochove K, et al. Evaluating a novel approach to stimulate open science collaborations: a case series of “study-a-thon” events within the OHDSI and European IMI communities. *JAMIA Open* 2022;5:ooac100.
 - Mooses K, Vesilind K, Oja M, et al. The use of prescription drugs and health care services during the 6-month post-COVID-19 period. *Sci Rep* 2023;13:11638.
 - Mooses K, Šavrova A, Pajusalu M, et al. Using electronic health records to evaluate the adherence to cervical cancer prevention guidelines: A cross-sectional study. *Preventive Medicine*. 2024;183:107982.
 - Uusküla A, Oja M, Tamm S, et al. Prevalence of type-specific human papillomavirus infection by grade of cervical cytology in Estonia. *JAMA Netw Open* 2023;6:e2254075.
 - Blöndal M, Oja M, Elmet M, Irs A, Kasepalu T, Reisberg S jt. Terviseandmete teise kasutamine südamepuudulikkuse diagnoosimise ja ravikäsitluse hindamisel. *Eesti Arst* 2026;105:11–9.

Liigne rasva osakaal kehas võib põhjustada metaboolseid häireid ka normaalse kehamassiindeksiga isikutel

Traditsiooniliselt hinnatakse liigse rasvahulga ladestumist kehas kehamassiindeksi (KMI) väärtuse alusel: normaalne KMI on 18,5–24,9 kg/m², ülekaalu KMI 25–29,9 kg/m², rasvumise KMI 30 ja enam kg/m². Samas on leitud, et rasvumisele iseloomulikke metaboolseid häireid võib esineda ka normaalse KMI-ga isikutel ning seda nimetatakse normaalkaaluga rasvumiseks.

USA mitme ülikooli koostöös tehtud uuringus testiti hüpoteesi,

mille kohaselt röntgen-absorptsiomeetria (DEXA) meetodil mõõdetud liigne rasva osakaal kehas võib seonduda metaboolsete häiretega ka normaalse KMI korral. Uuringusse kaasatud 20 613 isikul mõõdeti DEXA-meetodil rasva osakaalu protsent keha koosseisust ja hinnati selle seost insuliiniresistentsuse, metaboolse sündroomi, düslipideemia, 2. tüüpi diabeedi, arteriaalse hüpertensiooni ja üldsuremusega. DEXA-uuringud tehti kahel korral: aastatel 1999–2006 ja 2011–2018, jälgimisperiood oli 2019. aasta lõpuni.

Uuritustest ligi kolmandikul, 6424 isikul, oli KMI normaalväärtuses. Andmete analüüsil, kus

arvestati sugu, vanust, sotsiaalmajanduslikke tegureid, suitsetamist, alkoholi tarvitamist ja perekonnaseisu, ilmnes, et selle grupi uuritutel suurendas rasva osakaalu kasv iga 10 protsendipunkti võrra oluliselt üldise suremuse riski, insuliiniresistentsuse riski, metaboolse sündroomi riski ja düslipideemia riski.

Autorite hinnangul vallanduvad rasva osakaalu suurenemise korral organismis metaboolsetele häiretele viivad mõjud ka normaalse KMI-ga isikutel.

REFEREERITUD

Chen Z, Li J, Medina-Inojosa BJ, et al. Body fat and cardiometabolic disease burden and all-cause mortality in adults with normal body mass index. *Mayo Clin Proc* 2026;101:249–58.

LÜHIDALT