

TARTU ÜLIKOOL
LOODUS- JA TÄPPISTEADUSTE VALDKOND
MATEMAATIKA JA STATISTIKA INSTITUUT

Kaur-Erik Suurorg
Eksposoomi mõju insuldi tekkele

Matemaatiline statistika
Bakalaureusetöö (9 EAP)

Juhendajad: PhD Jaanika Kronberg, prof. Krista Fischer

TARTU 2026

EKSPOSOOMI MÕJU INSULDI TEKKELE

Bakalaureusetöö

Kaur-Erik Suurorg

Lühikokkuvõte

Eksposoomiks nimetatakse inimese tervist mõjutavate mittegeneetiliste tegurite kogumit. Käesoleva bakalaureusetöö eesmärk oli uurida Tartu Ülikooli Eesti geenivaramu andmete põhjal eksposoomitunnuste ja insuldiriski seoseid ning hinnata, kas keskkonnategurite põhjal on võimalik konstrueerida insuldiriski kirjeldav riskiskoor. Analüüsis kasutati Eesti geenivaramu tervise- ja küsimustikuandmeid ning EXPANSE projekti raames harmoniseeritud keskkonnaandmeid. Seoste hindamiseks kasutati Coxi võrdeliste riskide mudelit, kuhu kaasati lisaks keskkonnatunnustele ka vanus, kehamassiindeks, haridustase, suitsetamine ja sugu. Eksposoomitunnuste koosmõju uurimiseks rakendati peakomponentanalüüsi. Lisaks võrreldi teiste kohortide avaldatud andmetel põhinevat, üldtunnustel põhinevat ja eksposoomitunnuseid sisaldavaid riskiskoore. Tulemused näitasid, et klassikalised riskitegurid olid insuldiriskiga selgelt seotud, kuid kasutatud eksposoomitunnuste lisamine ei parandanud riskiskoori diskrimineerimisvõimet.

CERCS teaduseriala: P160 Statistika, operatsioonianalüüs, programmeerimine, finants- ja kindlustusmatemaatika.

Märksõnad: Eksposoom, elulemusanalüüs, Coxi mudel, riskiskoor.

THE IMPACT OF EXPOSOME ON STROKE DEVELOPMENT

Bachelor thesis

Kaur-Erik Suurorg

Abstract

The exposome refers to the set of non-genetic factors that affect human health. The aim of this bachelor's thesis was to study associations between exposome variables and stroke risk using data from the Estonian Biobank, and to evaluate whether environmental factors can be used to construct a risk score describing stroke risk. The analysis used health and questionnaire data from the Estonian Biobank together with external exposome data from the EXPANSE project. Associations were assessed using the Cox proportional hazards model, including environmental variables as well as age, body mass index, education, smoking and sex. Principal component analysis was applied to study the combined structure of exposome variables. In addition, a risk score based on published results from other European cohorts, a risk score based on classical risk factors, and risk scores including exposome variables were compared. The results showed that classical risk factors were clearly associated with stroke risk, whereas adding the available exposome variables did not improve the discriminative ability of the risk score.

CERCS research specialisation: P160 Statistics, operations research, programming, financial and actuarial mathematics.

Key Words: Exposome, survival analysis, Cox model, risk score.

Sisukord

Sissejuhatus	5
1 Taust	6
1.1 Eksposoom	6
1.2 Insult	7
2 Metoodika	9
2.1 Elulemusanalüüs	9
2.1.1 Coxi võrdeliste riskide mudel	10
2.2 Peakomponentanalüüs	12
2.3 Riskiskoor	14
3 Analüüs	16
3.1 Kirjeldav analüüs	16
3.1.1 Klassikalised riskitegurid	17
3.1.2 Eksposoomitunnused	18
3.2 Teiste Euroopa kohortide avaldatud hinnangutel põhinev riskiskoor	20
3.3 Coxi võrdeliste riskide mudelid	23
3.4 Peakomponentanalüüs	27
3.5 Riskiskoor	29
4 Tulemuste arutelu	32

Kokkuvõte	34
Kasutatud allikad	36
Lisad	38
Lisa 1	38
Lisa 2	39
Lisa 3	40
Lisa 4	41
Lisa 5	42
Lisa 6	43

Sissejuhatus

Insult on üks peamisi südame-veresoonkonnahaigustega seotud surma ja puude põhjustajaid maailmas [1]. Lisaks klassikalistele riskiteguritele, nagu vanus, suitsetamine ja kõrgvererõhutõbi, on järjest rohkem uuritud ka keskkonnategurite mõju insuldi tekkele, mida käsitletakse eksposoomi mõiste all [2]. Varasemates uuringutes on õhusaastet, linnakeskkonda ja teisi eksposoomitunnuseid seostatud kõrgema insuldiriskiga [3, 4]. Üheks suuremaks Euroopa eksposoomiprojektiks on EXPANSE, mille eesmärk on luua andmestik ja tööriistad elukeskkonna mõju uurimiseks tervisele [5].

Bakalaureusetöö eesmärk on uurida eksposoomitunnuste ja insuldiriski seoseid Eesti geenivaramu andmete põhjal ning hinnata, kas keskkonnategurite põhjal on võimalik konstrueerida insuldiriski kirjeldav riskiskoor. Töös kasutatakse Tartu Ülikooli Eesti geenivaramu, Eesti Haigekassa Ravikindlustuse andmekogu ning EXPANSE projekti käigus koostatud keskkonnaandmeid.

Töö tausta osas antakse ülevaade insuldist ning eksposoomi mõistest. Meetodika peatükis kirjeldatakse elulemusanalüüsis kasutatud Coxi võrdeliste riskide mudelit, peakomponentanalüüsi ning riskiskoori konstrueerimise põhimõtteid. Analüüsi osas kirjeldatakse uuritava valimi moodustamist, hinnatakse teiste kohortide avaldatud andmetel põhineva riskiskoori sobivust Eesti geenivaramu andmetel ning uuritakse eksposoomi tunnuste mõju Coxi mudelite ja peakomponentanalüüsi abil. Lisaks konstrueeritakse keskkonnategurite põhjal insuldiriski kirjeldavad riskiskoorid. Analüüs viidi läbi TÜ SAPU keskkonnas rakendustarkvara R abil ning töö kirjutamisel kasutati Overleafis programmi L^AT_EX.

Töö autor tänab juhendajaid Jaanika Kronbergi ja Krista Fischerit ning TÜ Eesti geenivaramut andmete kasutamise loa eest.

1 Taust

Selles peatükis kirjeldatakse eksposoomi olemust ning peamisi insuldi tekkega seotud riskitegureid.

1.1 Eksposoom

Eksposoomiks nimetatakse inimese tervist mõjutavate mittegeneetiliste tegurite kogumit kogu inimese elukaare jooksul. Eksposoom hõlmab nii väliskeskkonnast tulenevaid mõjureid, nagu õhusaaste, temperatuur ja ehitatud keskkond, kui ka individuaalseid ja käitumuslikke tegureid. [2] Lisaks üksikutele kokkupuudetele käsitletakse eksposoomi raames ka erinevate tegurite koosmõju ning kokkupuudetega seotud bioloogilisi reaktsioone [6].

Eksposoomi käsitlus erineb varasematest keskkonnatervise uuringutest selle poolest, et keskendutakse üksiktunnuste asemel mitme samaaegse kokkupuute koosmõjule. Kuna keskkonnategurid esinevad sageli koos ning on omavahel korreleeritud, ei pruugi nende eraldi analüüsimine anda täielikku pilti nende mõjust tervisele. Seetõttu on eksposoomi uurimisel oluline käsitleda erinevaid tegureid ühtse süsteemina ning kasutada meetodeid, mis võimaldavad arvesse võtta nende vastastikuseid seoseid. [6, 7]

Euroopas on eksposoomi uurimiseks loodud mitmeid rahvusvahelisi projekte, millest üks on EXPANSE projekt. Projekti eesmärk on uurida linna-keskkonna eksposoomi mõju tervisele, ühendades erinevate Euroopa riikide keskkonna-, tervise- ja ruumiandmeid. Projekti raames käsitletakse samaaegselt mitmeid keskkonnategureid, sealhulgas õhusaastet, rohealasid, müra-taset, ehitatud keskkonda ja kliimategureid, et hinnata nende seoseid krooniliste haiguste ja terviseriskidega. [5] Selliste andmete põhjal on uuritud

ka eksposoomi ja insuldi vahelist seost, käsitledes keskkonnategureid mitte üksikute tunnustena, vaid nende kombinatsioonidena [3]. Eesti geenivaramu andmed võimaldavad analoogset lähenemist rakendada ka Eesti populatsioonis, hinnates eksposoomi näitajate seoseid insuldi esinemisega ning uurides võimalust konstrueerida nende põhjal insuldi tekkeriski kirjeldav riskiskoor.

Kuigi eksposoom hõlmab teoreetiliselt kõiki inimese elu jooksul esinevaid mittegeneetilisi kokkupuuteid, on praktikas võimalik analüüsida vaid piiratud hulka mõõdetavaid tegureid. Käesolevas töös keskendutakse peamiselt väliskeskkonnaga seotud eksposoomi näitajatele, nagu õhusaaste ja ehitatud keskkonna tunnused, ning valitud elustiiliga seotud riskiteguritele. Seetõttu ei käsitleta töös eksposoomi täielikku ulatust, vaid selle piiratud alamhulka, mille kohta olid Eesti geenivaramu andmetes olemas sobivad mõõtmised.

1.2 Insult

Südame-veresoonkonna haigused põhjustavad ligikaudu kolmandiku maailma surmadest. Nende hulka kuuluv insult on üks peamisi suremuse ja pikaajalise töövõimetuse põhjuseid. [1] Insult tekib aju verevarustuse häire tagajärjel, kui ajuveresoon kas ummistub või lõhkeb. Sõltuvalt tekkemehhanismist jaotatakse insult isheemiliseks (ajuinfarkt) ja hemorraagiliseks (ajusisene verejooks). [8]

Insulti käsitletakse käesolevas töös RHK-10 klassifikatsiooni alusel [9]. Vaatluse all on järgmised diagnoosikoodid:

- I60 – subarahnoidaalne hemorraagia e ämblikuvõrkkesta-alune verevalum;
- I61 – intratserebraalne hemorraagia e peaajusisene verevalum;

- I62 – muu mittetraumaatiline koljusisene verevalum;
- I63 – peajuinfarkt;
- I64 – täpsustamata hemorraagiline või isheemiline insult.

Insuldi kujunemist mõjutavad mitmed riskitegurid, mida võib jagada elustiiliga seotud ja keskkondlikeks teguriteks. Elustiiliga seotud tegurite hulka kuuluvad näiteks suitsetamine, ebatervislik toitumine, vähene füüsiline aktiivsus ning kaasuvad haigused, nagu kõrgvererõhutõbi. [10] Viimastel aastatel on üha enam tähelepanu pööratud ka keskkonnategurite rollile haiguste kujunemisel. Uuringud on näidanud, et õhusaaste suurendab insuldi tekkeriski, samas kui ümbritseva keskkonna rohelus võib olla kaitseva toimega. Samuti on uuritud seoseid ka temperatuuri ning insuldi vahel, tuues esile madalaid talve- ning kõrgeid suvetemperatuure. [3]

2 Metoodika

Käesolevas peatükis antakse ülevaade töö metoodilisest raamistikust. Esmalt tutvustatakse elulemusanalüüsi ning Coxi võrdeliste riskide mudelit. Seejärel käsitletakse eksposoomi kirjeldamiseks kasutatavaid statistilisi meetodeid, keskendudes peakomponentanalüüsile. Peatüki lõpus käsitletakse lühidalt riskiskoori konstrueerimise ja mudelite hindamise üldpõhimõtteid. Peatükk põhineb G. James *et al.* õpiku "*An Introduction to Statistical Learning*" [11] lehekülgedel 253–256, 265, 461–465, 469–471 ning 481, kui ei ole viidatud teisiti.

2.1 Elulemusanalüüs

Elulemusanalüüs on statistiliste meetodite kogum, mida kasutatakse aja jooksul toimuva sündmuse modelleerimiseks. Selle keskmes on juhuslik suurus T , mis kirjeldab aega kindla sündmuseni. Käesoleva töö kontekstis on vaatluse all aeg uuringu algushetkest kuni insuldi esinemiseni. Elulemusanalüüsi eripära võrreldes tavapärase regressioonimeetoditega seisneb selles, et lisaks sündmuse toimumisele võetakse arvesse ka selle toimumise aega.

Elulemusandmete analüüsimisel esineb sageli tsenseerimist, mis tähendab, et kõigi vaatlusaluste kohta ei ole sündmuse täpset toimumisaega teada. Kõige sagedamini esineb parempoolset tsenseerimist, kus sündmus ei ole vaatlusperioodi lõpuks veel toimunud või isik lahkub uuringust enne sündmuse esinemist. Sellisel juhul on teada vaid see, et tegelik sündmuse aeg ületab vaadeldud ajahetke. Tsenseerimise arvestamine on elulemusanalüüsis oluline, sest selle eiramine võib viia tugevalt kallutatud hinnanguteni. Samas sisaldavad ka tsenseeritud vaatlused olulist informatsiooni, sest selle põhjal on teada, et vaadeldud perioodil isikul sündmust ei toimunud.

Olgu T mittenegatiivne juhuslik suurus, mis kirjeldab aega sündmuseni. Üleelamisfunktsioon

$$S(t) = P(T > t)$$

annab tõenäosuse, et sündmus ei ole ajaks t veel toimunud. Kui T -l leidub tihedusfunktsioon $f(t)$, siis kehtib seos

$$f(t) = -\frac{d}{dt}S(t) = -S'(t).$$

Riskifunktsioon ehk *hazard function* kirjeldab sündmuse toimumise hetkekii-
rust tingimusel, et sündmus ei ole enne aega t aset leidnud, ning selle saab
defineerida kujul

$$h(t) = \lim_{\Delta t \rightarrow 0} \frac{P(t \leq T < t + \Delta t \mid T \geq t)}{\Delta t}.$$

Kui tihedusfunktsioon eksisteerib, saab riskifunktsiooni esitada ka kujul

$$h(t) = \frac{f(t)}{S(t)}.$$

Käesolevas töös võimaldab elulemusanalüüs käsitleda aega uuringu algushet-
kest kuni insuldi esinemiseni, arvestades seejuures ka tsenseeritud vaatlusi.
Praktikas modelleeritakse sageli just riskifunktsiooni sõltuvust seletavatest
tunnustest, milleks kasutatakse näiteks Coxi võrdeliste riskide mudelit, mida
käsitletakse järgmises alapeatükis [3].

2.1.1 Coxi võrdeliste riskide mudel

Elulemusanalüüsis on sageli huvi kirjeldada, kuidas sündmuse esinemisrisk
sõltub seletavatest tunnustest. Selleks modelleeritakse tinglikku riskifunkt-

siooni kuju

$$h(t \mid \mathbf{X}),$$

kus $\mathbf{X} = (X_1, \dots, X_p)^T$ on seletavate tunnuste vektor.

Coxi võrdeliste riskide mudel eeldab, et riskifunktsiooni saab esitada kujul

$$h(t \mid \mathbf{X}) = h_0(t) \exp(\boldsymbol{\beta}^T \mathbf{X}),$$

kus $h_0(t)$ on baasriskifunktsioon ning $\boldsymbol{\beta} = (\beta_1, \dots, \beta_p)^T$ on tundmatute parameetrite vektor. Sellisel kujul jaotub riskifunktsioon kaheks komponendiks: ajast sõltuv baasrisk $h_0(t)$, mis on kõigile vaatlusalustele ühine, ning eksponentsiaalne osa $\exp(\boldsymbol{\beta}^T \mathbf{X})$, mis kirjeldab seletavate tunnuste mõju riskile.

Mudeli oluline omadus on võrdeliste riskide eeldus. Kahe indiviidi i ja j korral, kelle tunnuste vektorid on vastavalt \mathbf{X}_i ja \mathbf{X}_j , kehtib

$$\frac{h(t \mid \mathbf{X}_i)}{h(t \mid \mathbf{X}_j)} = \exp(\boldsymbol{\beta}^T (\mathbf{X}_i - \mathbf{X}_j)),$$

mis ei sõltu ajast t . See tähendab, et individidevaheline suhteline risk püsib ajas konstantne.

Parameetrite $\boldsymbol{\beta}$ hinnangud leitakse osalise tõepära meetodil. See meetod lähtub suurima tõepära meetodi ideest, kuid erinevalt tavalisest tõepärafunktsioonist ei nõua baasriskifunktsiooni $h_0(t)$ kuju täpsustamist. Olgu $t_1 < t_2 < \dots < t_D$ järjestatud sündmuse toimumise ajad ning $R(t_k)$ riskihulk hetkel t_k , s.t nende individide hulk, kellel enne aega t_k ei ole sündmus veel toimunud. Siis avaldub osaline tõepära kujul

$$L(\boldsymbol{\beta}) = \prod_{k=1}^D \frac{\exp(\boldsymbol{\beta}^T \mathbf{X}_{(k)})}{\sum_{l \in R(t_k)} \exp(\boldsymbol{\beta}^T \mathbf{X}_l)},$$

kus $\mathbf{X}_{(k)}$ tähistab selle indiviidi tunnuste vektorit, kellel sündmus toimus ajal t_k .

Mudeli parameetrite tõlgendamine põhineb *hazard ratio*'l (HR) ehk riskisuhtel. Kui X_j suureneb ühe ühiku võrra ja teised tunnused ei muutu, siis muutub risk teguri $\exp(\beta_j)$ võrra. Seega kirjeldab $\exp(\beta_j)$ suhtelist muutust sündmuse tekkeriskis.

Coxi võrdeliste riskide mudel on laialdaselt kasutatav, kuna see võimaldab hinnata seletavate tunnuste mõju sündmuse tekkeriskile ajas ilma baasriskifunktsiooni kuju täpsustamata. Seetõttu sobib see hästi olukordadesse, kus huvi keskmes on erinevate riskitegurite suhteline mõju, nagu ka käesolevas töös insuldi tekkeriski modelleerimisel.

2.2 Peakomponentanalüüs

Eksposoomi käsitlemisel on tegemist mitmemõõtmelise tunnuste kogumiga, kus erinevad keskkonnategurid, nagu õhusaaste näitajad, rohelus ja temperatuur, võivad olla omavahel tugevalt korreleeritud. Selline korrelatsioonistruktuur võib põhjustada probleeme regressioonimudelite rakendamisel, kuna multikollineaarsus muudab parameetrite hinnangud ebastabiilseks ning raskendab üksiktunnuste mõju tõlgendamist.

Sarnane probleem ilmneb ka eksposoomi ja insuldi seoseid käsitlevates empiirilistes uuringutes, kus on ka välja toodud, et keskkonnategurite omavaheline tugev seotus raskendab nende eraldiseisvat kaasamist mudelisse ja võib vähendada hinnangute usaldusväärsust [3, 6]. Seetõttu kasutatakse eksposoomi analüüsimisel sageli mõõtmete vähendamise meetodeid, mis võimaldavad kirjeldada tunnuste ühisstruktuuri kompaktsemal kujul.

Üheks selliseks meetodiks on peakomponentanalüüs (PCA ehk *Principal Com-*

ponent Analysis), mille eesmärk on teisendada algsed omavahel korreleeritud tunnused uuteks omavahel ortogonaalseteks komponentideks, säilitades võimalikult suure osa andmete dispersioonist. See võimaldab vähendada andmete dimensiooni ning kasutada saadud komponente edasises modelleerimises.

Olgu $\mathbf{X} = (x_{ij})$ mõõtmetega $n \times p$ andmemaatriks, kus n tähistab vaatluste arvu ja p tunnuste arvu. Eeldatakse, et tunnused on eelnevalt tsentreeritud, st igast tunnusest on lahutatud selle keskmine. Sageli tunnused ka standardiseeritakse, et vältida erinevatest mõõtühikutest tulenevat mõju peakomponentide kujunemisele.

Peakomponentanalüüsi eesmärk on konstrueerida uued tunnused, mida nimetatakse peakomponentideks ja mis on algsete tunnuste lineaarkombinatsioonid. Kui ϕ_k tähistab k -nda peakomponendi kaalukordajate vektorit, siis selle komponendi väärtused kõigi vaatluste jaoks on antud vektoriga

$$\mathbf{z}_k = \mathbf{X}\phi_k, \quad k = 1, \dots, p.$$

Esimene peakomponent leitakse nii, et selle dispersioon oleks maksimaalne kõigi selliste lineaarkombinatsioonide seas, mille kaalukordajate vektori norm on 1. See tähendab, et lahendatakse optimeerimisülesanne

$$\max_{\phi_1} \text{Var}(\mathbf{X}\phi_1), \quad \text{tingimusel } \|\phi_1\| = 1.$$

Järgmised peakomponendid leitakse analoogselt, lisatingimusega, et need oleksid ortogonaalsed eelnevate komponentidega. Seega kehtib $\phi_k^T \phi_l = 0$ iga $k \neq l$ korral, mis tähendab, et saadud komponendid ei ole omavahel korreleeritud.

Matemaatiliselt vastab see ülesanne andmete kovariatsioonimaatriksi oma-

väärtuste ja -vektorite leidmisele, kusjuures iga peakomponent vastab ühele omavektorile ning selle dispersioon on võrdne vastava omaväärtusega. Sellest tulenevalt on võimalik hinnata, kui suure osa andmete dispersioonist iga komponent selgitab, ning valida edasiseks analüüsiks vaid esimesed komponendid, mis katavad ära suure osa hajuvusest.

Kuna peakomponendid on omavahel ortogonaalsed, välditakse nende kasutamisel regressioonimudelites multikollineaarsuse probleemi. Samas tuleb arvestada, et PCA on juhendamata meetod, mis ei arvesta sõltuvat muutujat, mistõttu valitud komponendid maksimeerivad küll seletavate tunnuste dispersiooni, kuid ei pruugi olla optimaalsed sündmuse tekkeriski selgitamiseks.

2.3 Riskiskoor

Elulemusanalüüsi kontekstis on loomulik viis individuaalse riski kirjeldamiseks Coxi võrdeliste riskide mudeli lineaarne prediktor. Kui riskifunktsioon on esitatud kujul

$$h(t | \mathbf{X}) = h_0(t) \exp(\boldsymbol{\beta}^T \mathbf{X}),$$

siis suurus

$$\eta = \boldsymbol{\beta}^T \mathbf{X}$$

kirjeldab indiviidi suhtelist riski võrreldes baasriskiga ning seda võib käsitleda riskiskoorina. Selline riskiskoor ei väljenda absoluutset tõenäosust, vaid võimaldab võrrelda indiviididevahelisi riske. Suurem skoor vastab kõrgemale sündmuse tekkeriskile, kusjuures riskisuhe kahe indiviidi vahel on antud nende skooride eksponentsiaalse vahe kaudu.

Eksposoomi kontekstis ei ole tunnused tavaliselt sõltumatud, mistõttu kasutatakse sageli eelnevalt dimensiooni vähendamise meetodeid. Näiteks de Bont

et al. käsitlevad keskkonnategurite koosmõju, kasutades peakomponentanalüüsi, mille abil saadud komponendid kaasatakse Coxi mudelisse [3]. Sellisel juhul põhineb riskiskoor mitte algsetel tunnustel \mathbf{X} , vaid nende lineaarsetel kombinatsioonidel Z , ning lineaarne prediktor omandab kuju

$$\eta = \boldsymbol{\gamma}^T \mathbf{Z},$$

kus \mathbf{Z} tähistab peakomponente.

Üldisemalt on võimalik riskiskoori konstrueerimisel ühendada erinevat tüüpi tunnuseid. Lisaks eksposoomi kirjeldavatele komponentidele võib mudelisse kaasata ka klassikalisi riskitegureid, nagu elustiili- ja tervisenäitajad. Sellisel juhul on riskiskoor esitatav kujul

$$\eta = \boldsymbol{\gamma}^T \mathbf{Z} + \boldsymbol{\alpha}^T \mathbf{W},$$

kus \mathbf{Z} tähistab eksposoomi komponente ning \mathbf{W} teisi seletavaid tunnuseid. Selline lähenemine võimaldab hinnata nii keskkonnategurite koondmõju kui ka nende mõju võrreldes teiste teadaolevate riskiteguritega.

Riskiskoori kvaliteeti hinnatakse eelkõige mudeli diskrimineerimisvõime ja üldistusvõime alusel. Diskrimineerimisvõimet saab hinnata näiteks kooskõlaindeksi (C-indeks) abil, mis mõõdab, kui hästi mudel järjestab individid vastavalt nende riskile. Üldistusvõime hindamiseks kasutatakse sõltumatuid andmestikke, et hinnata mudeli rakendatavust uutele andmetele.

3 Analüüs

Bakalaureusetöös kasutatakse Tartu Ülikooli Eesti geenivaramu pseudonüümitud isikuandmeid ning EXPANSE projekti raames mudeldatud keskkonnaandmeid, mis on seotud geenidoonorite elukohaandmetega. Uuring toimus vastava Eesti bioetika ja inimuuringute nõukogu (EBIN) loa alusel (1.1-12/1468), andmed väljastati Eesti geenivaramu poolt väljastustaotluse V17 alusel ning analüüs viidi läbi Tartu Ülikooli turvalises SAPU serveris.

3.1 Kirjeldav analüüs

Analüüsis kasutatakse kahte Eesti Geenivaramu kohorti, mis erinevad nii värbamisperioodi kui ka kujunemisloogika poolest. Esimese kohordi moodustavad aastatel 2002–2011 geenivaramuga liitunud geenidoonorid, teise kohordi aastatel 2018–2019 värvatud geenidoonorid. Andmestikust eemaldati isikud, kellel puudusid eksposoomi või teiste analüüsis kasutatavate tunnuste väärtused. Elulemusanalüüsis käsitletakse ainult aega esimese insuldijuhuni. Puuduvate väärtuste ülevaade tunnuste kaupa on esitatud lisas 1.

Puhastatud andmestik on kokku 180 948 isikut, kellest 36 855 kuuluvad esimesse kohorti ja 144 093 teise kohorti. Jälgimisperioodi lõpuks loetakse 31.12.2023, isiku surma aeg või insuldi diagnoosimise aeg juhul, kui insult esines jälgimisperioodi jooksul. Analüüsis käsitletakse ainult pärast geenivaramuga liitumist tekkinud insuldijuhte — isikud, kellel oli insuldidiagnoos juba enne liitumist, olid juba varasemalt andmestikust kaotatud. Kontrollideks loetakse isikud, kellel ei olnud enne ega pärast geenivaramuga liitumist ühtegi insuldidiagnoosi. Puhastatud andmestik on 'I60', 'I61' ja 'I63' juhte vastavalt 160, 225 ning 1922. Edasises analüüsis käsitletakse neid diagnoose koondatult ühe sündmusena, mida tähistatakse insuldi esinemisena. Eesti

geenivaramu andmestikus esines ka diagnoosikoode 'I62' ja 'I64', kuid diagnoosikuupäeva puudumise tõttu ei olnud neid võimalik elulemusanalüüsis kasutada.

3.1.1 Klassikalised riskitegurid

Insuldi tekkerisk sõltub tugevasti klassikalistest riskiteguritest. Käesolevas töös vaadeldakse tunnuseid *vanus*, *suitsetamine*, *kehamassiindeks* ja *haridus*. Tunnus *suitsetamine* on binaarne, kus väärtuseks 1 loetakse endised ja hetkel suitsetavad isikud ning väärtuseks 0 inimesed, kes ei ole kunagi suitsetanud. Tunnus *haridus* on samuti binaarne, kus väärtus 1 määrati vähemalt bakalau-reusekraadiga ning väärtus 0 kesk- või põhiharidusega isikutele. Haridustaset käsitletakse ühe võimaliku kaudse näitajana inimese sotsiaalmajanduslikust taustast ja sellega seotud elustiilierinevustest. Tunnuseid *vanus* ja *kehamassiindeks* kasutatakse edasises analüüsis standardiseeritud kujul, mistõttu nende mõju Coxi mudelites tõlgendatakse ühe standardhälbe suuruse muutuse kohta.

Kuna kasutatud Eesti geenivaramu kohordid erinevad teineteisest eelkõige jälgimisperioodi pikkuse poolest, arvutati mudelid esimese kohordi andmetel ning teist kohorti kasutati tulemuste valideerimiseks. Lisaks arvestatakse mudelites sugu, kuid seda ei kaasata tavapärase seletava tunnuseks, vaid stratifitseeriva ehk kihitava tunnuseks. Selline lähenemine võimaldab meeste ja naiste baasriskifunktsioonidel erineda, kuna nende insuldi baasriskid ei pruugi olla võrdelised [8].

3.1.2 Eksposoomitunnused

Analüüsis kasutatakse eksposoomi tunnuseid, mis kirjeldavad geenidoonori elukohaga seotud väliskeskkonda. Eksposoomiandmed on seotud geenidoonorite registreeritud elukoha aadressiga. Analüüsis eeldatakse, et kasutatud aadress kirjeldab piisavalt hästi inimese pikaajalist elukeskkonda ning sellega seotud eksposoominäitajaid. Vaadeldavateks tunnusteks on õhusaaste näitajad $PM_{2,5}$, NO_2 ja O_3 , rohelist kirjeldavad indeksid $NDVI$ ja $MSAVI$, vett mitteläbilaskvate pindade osakaal (*impervious surfaces*, edaspidi *asfalt*) ning *Urban index* ehk linnasuse tase (edaspidi *linnasus*). Edasises analüüsis käsitletakse neid nii eraldi kui ka omavahel seotud eksposoomitunnuste kogumina.

$PM_{2,5}$ kirjeldab õhus leiduvate peenosakeste kontsentratsiooni, mille osakeste läbimõõt on väiksem kui 2,5 mikromeetrit. NO_2 ehk lämmastikdioksiid on peamiselt liiklusest pärinev õhusaaste komponent ning O_3 ehk osoon kirjeldab maapinnalähedast osooni kontsentratsiooni. $NDVI$ (*Normalized Difference Vegetation Index*) ja $MSAVI$ (*Modified Soil Adjusted Vegetation Index*) on satelliidipiltidel põhinevad keskmise roheluse indeksid, mille suuremad väärtused viitavad suuremale taimkatte osakaalule. Tunnused $NDVI_SD$ ja $MSAVI_SD$ kirjeldavad roheluse standardhälvet. Tunnus *asfalt* kirjeldab vett mitteläbilaskvate pindade osakaalu elukoha ümbruses ning *linnasus* piirkonna urbaniseerituse taset. Tunnused $NDVI$, $MSAVI$, *asfalt*, $NDVI_SD$ ja $MSAVI_SD$ on arvutatud andmestikus inimese elukohana märgitud aadressi ümbruses 1000 meetri raadiuses. Kõik pidevad eksposoomitunnused standardiseeriti enne Coxi mudelitesse ja peakomponentanalüüsi kaasamist, mistõttu on mudelites hinnatud mõjud tõlgendatavad ühe standardhälbe suuruse muutuse kohta.

Tabelis 1 on esitatud klassikaliste riskitegurite ning eksposoomitunnuste kirjeldus kohordi ja soo lõikes. Nähtub, et teises kohordis on insuldijuhtude suhteline osakaal väiksem kui esimeses kohordis, mis on tõenäoliselt seotud lühema jälgimisperioodiga. Teises kohordis on suurem kõrgharidusega isikute osakaal ning naiste seas esineb rohkem suitsetamist kui esimeses kohordis. Eksposoomitunnuste põhjal on kohortide keskmised näitajad üldiselt sarnased. Mõnevõrra erinevad õhusaaste näitajad, kus teises kohordis on $PM_{2,5}$ ja NO_2 keskmised väärtused madalamad ning O_3 väärtused kõrgemad kui esimeses kohordis. Pidevate tunnuste histogrammid on toodud lisas 2.

Tabel 1: Tunnuste kirjeldus grupiti. K1 viitab esimesele ning K2 teisele kohordile.

	K1 mehed	K1 naised	K2 mehed	K2 naised
inimeste arv	11 938	24 914	50 044	94 034
insultide arv	527	814	439	526
<i>Vanus</i> , keskmine \pm sd	40,4 \pm 16,5	43,9 \pm 16,0	43,7 \pm 15,3	45,1 \pm 15,9
<i>KMI</i> , keskmine \pm sd	26,38 \pm 4,55	26,25 \pm 5,65	26,88 \pm 4,43	25,57 \pm 5,34
<i>Suitsetajad</i> , n (%)	7 150 (59,9)	8 605 (34,5)	30 156 (60,3)	38 769 (41,2)
<i>Kõrgharidus</i> , n (%)	4 052 (33,9)	10 624 (42,6)	21 849 (43,7)	50 783 (54,0)
<i>Eksposoomitunnused</i> (keskmine \pm sd)				
$PM_{2,5}$	9,06 \pm 1,60	9,14 \pm 1,65	7,46 \pm 1,32	7,52 \pm 1,31
NO_2	11,07 \pm 5,87	11,53 \pm 5,73	9,63 \pm 4,94	10,10 \pm 4,88
O_3	63,00 \pm 4,77	62,90 \pm 4,82	65,27 \pm 3,47	65,20 \pm 3,41
<i>NDVI</i>	0,55 \pm 0,09	0,54 \pm 0,09	0,54 \pm 0,10	0,53 \pm 0,09
<i>NDVI_SD</i>	0,14 \pm 0,03	0,14 \pm 0,03	0,14 \pm 0,02	0,14 \pm 0,02
<i>MSAVI</i>	0,70 \pm 0,08	0,69 \pm 0,08	0,68 \pm 0,08	0,68 \pm 0,08
<i>MSAVI_SD</i>	0,13 \pm 0,02	0,13 \pm 0,02	0,13 \pm 0,02	0,13 \pm 0,02
<i>Asfalt</i>	12,35 \pm 16,08	13,56 \pm 16,01	16,81 \pm 17,55	18,41 \pm 17,54
<i>Linnasus</i>	18,00 \pm 22,39	19,99 \pm 22,93	21,37 \pm 22,39	23,51 \pm 23,20

3.2 Teiste Euroopa kohortide avaldatud hinnangutel põhinev riskiskoor

Eksposoomi mõju hindamiseks konstrueeriti insuldi riskiskoor, mis põhineb Jeroen de Bont *et al.* artiklis [3] esitatud hinnangutel. Riskiskoori koostamisel kasutati üksiktunnuste põhjal hinnatud riskisuhteid (HR), mis on esitatud protsentuaalse muutusena insuldi riskis kindla eksposoomitunnuse ühiku suurenemisel. Skoori konstrueerimisel lähtuti viiest eksposoomitegurist, mis olid välja toodud ka uuringus: $PM_{2,5}$, NO_2 , O_3 , $NDVI$ ning *asfalt*. Välja jäid *BC* (*black carbon*) ehk tahm, mida ei ole geenivaramu andmestikus, *linnasus* ning *MSAVI*, mis puudusid uuringust, ning temperatuuriga seotud näitajad, mille mõju oli varasemas uuringus nõrgem ja ebaühtlasem. Riskiskoori toimivust hinnati Eesti geenivaramu I kohordi põhjal, kuna selle geenidonorite jälgimisaeg on pikem. Pikem jälgimisperiood võimaldab paremini hinnata, kas eksposoomi põhjal konstrueeritud riskiskoor on seotud insuldi tekkeriskiga.

Esmalt kasutati registripõhiste kohortide hinnanguid riskiskoori konstrueerimiseks. Registripõhised kohortid on moodustatud rahvastikuregistri põhjal, kus andmestikus leiduvad vaid diagnoosi- ning keskkonnaandmed. Kuna temperatuuriga seotud tunnused ei olnud kõigi vaatluste jaoks kättesaadavad ning nende mõju oli uuringus nõrgem ja ebaühtlasem, jäetakse need riskiskoori konstrueerimisel välja. Seega põhineb skoor õhusaaste, roheluse ja vett mitteläbilaskvate pindade näitajatel.

Euroopa registripõhiste kohortide hinnangutel põhinevad riskisuhted teisen-dati logaritmilisele kujule, kasutades seost

$$\beta = \log(HR),$$

kus $HR = 1 + \Delta\%/100$. Saadud koefitsiente kasutati lineaarse kombinatsioo-
nina, kus iga tunnus skaleeriti vastavalt uuringus kasutatud ühikule.

Isiku i registripõhiste kohortide hinnangutel põhinev eksposoomi riskiskoor
on defineeritud kujul

$$S_i^{(R)} = \beta_1 \cdot \frac{PM_{2,5i}}{5} + \beta_2 \cdot \frac{NO_{2i}}{10} + \beta_3 \cdot \frac{O3_i}{10} + \beta_4 \cdot \frac{NDVI_i}{0,1} + \beta_5 \cdot \frac{asfalt_i}{10},$$

kus koefitsiendid β_j on arvatud uuringus leitud registripõhiste kohortide
riskisuhete hinnangute põhjal:

$$\beta_1 = \log(1,031), \beta_2 = \log(1,021), \beta_3 = \log(0,973), \beta_4 = \log(0,990), \beta_5 = \log(1,007).$$

Saadud skoor $S_i^{(R)}$ on suhteline eksposoomi riskiskoor, mis võimaldab järjes-
tada isikuid eksposoomi põhjal vastavalt nende hinnangulisele insuldiriskile.
Edasises analüüsis kasutatakse standardiseeritud skoori, et hinnata selle seost
insuldi esinemisega Coxi võrdeliste riskide mudelis.

Sarnaselt konstrueeriti ka traditsiooniliste kohortide avaldatud tulemustel
põhinev suhteline eksposoomi riskiskoor. Selle eesmärk oli hinnata, kas ris-
kiskoori toimivus Eesti andmetel sõltub sellest, kumma kohorditüübi põhjal
saadud hinnanguid kasutatakse.

Isiku i traditsiooniliste kohortide hinnangutel põhinev eksposoomi riskiskoor
on defineeritud kujul

$$S_i^{(T)} = \beta_1 \cdot \frac{PM_{2,5i}}{5} + \beta_2 \cdot \frac{NO_{2i}}{10} + \beta_3 \cdot \frac{O3_i}{10} + \beta_4 \cdot \frac{NDVI_i}{0,1} + \beta_5 \cdot \frac{asfalt_i}{10},$$

kus koefitsiendid β_j on arvatud traditsiooniliste kohortide hinnangute põh-

jal:

$$\beta_1 = \log(0,820), \beta_2 = \log(1,089), \beta_3 = \log(0,911), \beta_4 = \log(0,941), \beta_5 = \log(0,998).$$

Mõlema skoori toimivust hinnati Eesti andmetel Coxi võrdeliste riskide mudelis, kuhu lisati lisaks eksposoomi riskiskoorile ka klassikalised riskitegurid: vanus, kehamassiindeks, suitsetamine ja haridus. Tabelis 2 on esitatud ilma riskiskoorita ning riskiskoore sisaldavate mudelite riskisuhted, nende riskisuhte 95% usaldusintervallid ning C-indeksid.

Tabel 2: Avaldatud Euroopa kohortide hinnangutel põhinevate eksposoomi riskiskooride seos insuldi riskiga Eesti andmetel.

Mudel	Riskiskoori HR	95% intervall	p -väärtus	C
Ilma riskiskoorita	–	–	–	0,815
Registripõhiste kohortide riskiskoor	0,981	(0,927; 1,038)	0,500	0,814
Traditsiooniliste kohortide riskiskoor	1,010	(0,955; 1,067)	0,735	0,814

Ilmneb, et kumbki avaldatud Euroopa kohortide hinnangutel põhinev eksposoomi riskiskoor ei osutunud Eesti andmetel statistiliselt oluliseks. Registripõhiste kohortide riskiskoori korral oli riskisuhe 0,981 ja p -väärtus 0,500 ning traditsiooniliste kohortide riskiskoori korral vastavalt 1,010 ja 0,735. Samuti ei parandanud riskiskooride lisamine mudeli diskrimineerimisvõimet C-indeksi põhjal. Ilma eksposoomi riskiskoorita mudeli C-indeks oli 0,815, kuid mõlema riskiskoori lisamisel 0,814. Klassikalised riskitegurid, sealhulgas vanus, kehamassiindeks, suitsetamine ja haridus, jäid mudelites statistiliselt oluliseks. Seega ei kandunud Euroopa kohortide põhjal saadud eksposoomi riskiskoorid otseselt üle Eesti Geenivaramu andmetele.

Sellest tulenevalt tekib vajadus Eesti andmetele sobivama riskiskoori konstrueerimiseks, võttes arvesse ka Eesti andmestikus olemasolevaid täiendavaid eksposoominäitajaid.

3.3 Coxi võrdeliste riskide mudelid

Insuldiriski analüüsimisel on huvipakkuvaks suuruseks aeg vaatluse algusest kuni insuldi esmadiagnoosini. Kuna kõigil isikutel vaatluse lõpuks insulti ei diagnoositud, esineb andmestikus paremalt tsenseeritud vaatlusi. Sellise andmestruktuuri korral sobib analüüsiks Coxi võrdeliste riskide mudel, mis võimaldab hinnata, kuidas erinevad tunnused on seotud insuldi tekkeriskiga ajas. Analüüsis on sündmus defineeritud binaarse tunnusena *insult*, kus väärtus 1 tähistab insuldi esinemist pärast geenivaramuga liitumist ning väärtus 0 tsenseeritud vaatlust. Aeg sündmuseni on esitatud tunnusena *aeg*, mis mõõdab ajavahemikku geenidoonoriks saamisest kuni esimese insuldi diagnoosi või tsenseerimiseni päevades.

Coxi mudelid hinnatakse I kohordi põhjal, kuna selles kohordis on jälgimisaeg pikem ning seetõttu on mudeli hindamiseks rohkem informatsiooni. II kohorti kasutatakse hiljem hinnatud mudelite võimaliku ülekantavuse ja enustusvõime hindamiseks. Mudelis kasutatav olulisuse nivoo on $\alpha = 0,05$. Pidevad tunnused on enne modelleerimist standardiseeritud, mistõttu hinnatud kordajad on tõlgendatavad kui mõju ühe standardhälbe suuruse muutuse kohta.

Esmalt modelleeritakse insuldiriski üldtunnuste *haridus*, *kehamassiindeks*, *suitsetamine*, *vanus* ja *sugu* abil. Tunnuseid lisatakse mudelisse ükshaaval ning hinnatakse nende statistilist olulisust. Selgus, et kõik nimetatud tunnused on olulised tegurid. Tabelis 3 on esitatud lõpliku üldtunnuste mudeli

hinnatud riskisuhted, nende 95% usaldusintervallid ja p-väärtused, kusjuures kõik tunnused on mudelis statistiliselt olulised.

Tabel 3: Üldtunnuste põhjal hinnatud Coxi mudel.

Tunnus	HR (riskisuhe)	95% usaldusintervall	p-väärtus
$scale(vanus)$	3,37	(3,16; 3,60)	< 0,0001
$scale(KMI)$	1,17	(1,11; 1,23)	< 0,0001
$I(kõrgharidus)$	0,76	(0,67; 0,86)	< 0,0001
$I(suitsetamine)$	1,20	(1,07; 1,35)	0,003

Vanuse suurenemine ühe standardhälbe ühiku võrra on seotud ligikaudu 3,37-kordse insuldiriski kasvuga, mis viitab tugevale seosele vanuse ja insuldi esinemise vahel. Ka kehamassiindeksil on positiivne seos insuldiriskiga: ühe standardhälbe suurune kehamassiindeksi suurenemine tõstab riski ligikaudu 17%. Suitsetamine on samuti statistiliselt oluline riskitegur, suurendades insuldi tekkeriski ligikaudu 20%. Seevastu kõrgharidus on seotud väiksema riskiga — kõrgharidusega isikutel on hinnanguliselt umbes 24% madalam insuldirisk võrreldes nendega, kellel kõrgharidus puudub.

Saadud mudeli puhul kontrolliti ka võrdeliste riskide eeldust, kasutades Schoenfeldi jääke (R-i funktsioon `cox.zph`). Üldtunnuste korral ei leitud võrdeliste riskide eelduse rikkumist, mistõttu on Coxi mudeli kasutamine antud juhul põhjendatud.

Edaspidi lisatakse üldtunnuste mudelisse ükshaaval üheksa standardiseeritud keskkonnategurit, et hinnata iga tunnuse eraldi seost insuldiriskiga. Standardiseerimise tõttu on hinnatud kordajad võrreldavad ning tõlgendatavad ühe standardhälbe suuruse muutuse kaudu.

Saadud tulemused näitavad, et ükski keskkonnategur ei ole statistiliselt oluline insuldiriski selgitaja valitud olulisuse nivoo $\alpha = 0,05$ juures. Analüüsi

käigus selgus ka asjaolu, et ükshaaval keskkonnategurite lisamine mudelisse ei muuda oluliselt üldtunnuste hinnatud riskisuhteid ega nende statistilist olulisust ning seda kõigi üheksa keskkonnateguri puhul. Seetõttu on allolevas tabelis 4 välja toodud vaid keskkonnategurite ülevaade.

Tabel 4: Keskkonnategurite mõju insuldiriskile Coxi mudelis.

Keskkonnategur	HR (riskisuhe)	95% intervall	p-väärtus
$PM_{2,5}$	0,98	(0,93; 1,04)	0,54
NO_2	1,01	(0,96; 1,07)	0,65
O_3	0,96	(0,91; 1,01)	0,13
$NDVI$	1,01	(0,95; 1,06)	0,79
$NDVI_SD$	0,99	(0,94; 1,05)	0,74
<i>Asfalt</i>	0,98	(0,93; 1,04)	0,55
<i>Linnasus</i>	1,03	(0,98; 1,09)	0,28
$MSAVI$	1,01	(0,96; 1,07)	0,70
$MSAVI_SD$	0,97	(0,92; 1,03)	0,27

Kuigi ükski keskkonnategur ei ole statistiliselt oluline, võib märkida, et osooni (O_3) hinnanguline mõju erineb teistest mõnevõrra rohkem. Hinnatud riskisuhe on 0,96, mis viitab võimalikule negatiivsele seosele insuldiriskiga, kuid p-väärtus ei võimalda teha usaldusväärseid järeldusi. Selline tulemus ei pruugi tähendada, et keskkonnategurid ei ole insuldiriskiga seotud, vaid võib viidata sellele, et nende mõju avaldub mitme omavahel korreleeritud tunnuse koosmõjuna. Eksposoomi tunnused on sageli tugevalt korreleeritud, mis võib viia multikollineaarsuseni ning vähendada üksiktunnuste statistilist olulisust Coxi mudelis [3].

Lisaks üksiktunnuste mudelitele hinnati ka mudeleid, kuhu kaasati korraga kaks keskkonnategurit. Selle eesmärk oli uurida, kas keskkonnategurite mõju võib avalduda alles siis, kui arvestada samaaegselt teise keskkonnatunnuse taset. Kõigi võimalike paaride läbivaatamisel ilmnes, et mõnede tunnuste kombinatsioonide korral osutuvad keskkonnategurid statistiliselt oluliseks,

kuigi üksiktunnustena nad seda ei olnud.

Näiteks osutusid vett mitteläbilaskvad pinnad ja lämmastikdioksiidi tase samaaegsel kaasamisel Coxi mudelisse mõlemad statistiliselt oluliseks. Sarnane tulemus ilmnes ka linnastumise ning vett mitteläbilaskvate pindade kombinatsiooni puhul ning mõlema tunnuse p-väärtus oli eriti väike. Paare läbi vaadates on tegu suurehulgalise hüpoteeside testimisega ning seetõttu tuleks arvestada ka Bonferroni korrigeerimisega [12]. Arvestades, et meil on 9 keskkonnategurit, siis paarikaupa neid mudelitesse paigutades, peame testimise 36 mudelit. Meetodit arvesse võttes peab mõlema keskkonnateguri p-väärtus olema väiksem kui $\frac{0,05}{36} \approx 0,0014$.

Ainuke oluliseks osutuv Coxi mudel kasutab *linnasuse* ning *asfaldi* paari. Lisas 3 on toodud ülevaade kahe tunnuse hajuvusdiagrammidest mõlemas kohordis. Kuigi käesolevas uurimistöös ei ole võimalik lähemalt seost uurida, siis andmed tundub olevat võimalik jagada klastritesse. Seetõttu võib olla ka Coxi mudeli olulisuse taga täpsem põhjus. Tabelis 5 on toodud Coxi mudeli tunnuste olulisusetõenäosused, riskisuhted ning usaldusintervallid riskisuhtele.

Tabel 5: Üldtunnuste ning kahe keskkonnateguriga (asfalt + linnasus) Coxi mudeli mõju insuldiriskile.

Tunnus	HR (riskisuhe)	95% usaldusintervall	p-väärtus
<i>scale(vanus)</i>	3,37	(3,16; 3,60)	< 0,0001
<i>scale(KMI)</i>	1,17	(1,11; 1,23)	< 0,0001
<i>I(kõrgharidus)</i>	0,76	(0,67; 0,87)	< 0,0001
<i>I(suitsetamine)</i>	1,20	(1,07; 1,36)	0,003
<i>Asfalt</i>	0,81	(0,71; 0,92)	0,0008
<i>Linnasus</i>	1,23	(1,10; 1,37)	0,0004

Tulemus viitab sellele, et keskkonnategurite mõju ei avaldu üksiktunnuste tasemel, vaid pigem nende koosmõjus. See muudab üksikute tunnuste ots-

ese tõlgendamise keeruliseks ning viitab vajadusele käsitleda eksposoomi kui omavahel seotud tunnuste süsteemi. Seetõttu analüüsitakse edaspidi keskkonnategurite koosmõju, kasutades peakomponentanalüüsi.

3.4 Peakomponentanalüüs

Keskkonnategurite eraldi modelleerimine ei andnud statistiliselt olulisi tulemusi ning tunnuste paarikaupa kaasamisel saadi vaid üks statistiliselt oluline, kuid raskesti tõlgendatav mudel. Seetõttu uuritakse keskkonnategurite mõju edasi peakomponentanalüüsi abil, käsitledes eksposoomi kui omavahel seotud tunnuste kogumit ning püüdes leida sisukamaid ja paremini tõlgendatavaid tunnuste kombinatsioone.

Analüüs viidi läbi I kohordi andmetel, kasutades üheksat standardiseeritud keskkonnategurit. Saadud peakomponente kasutati Coxi võrdeliste riskide mudelis, kus neid lisati üldtunnuste mudelile. Peakomponendid lisati Coxi mudelisse analoogselt alg tunnustega, kusjuures hinnatakse nende seost insuldiriskiga.

Ükshaaval peakomponentide Coxi mudelisse lisamisel osutus statistiliselt oluliseks üksnes seitsmes peakomponent. Täiendavalt hinnati mudelit, kuhu kaasati samaaegselt kõik peakomponendid. Ka sellisel juhul jäi statistiliselt oluliseks üksnes seitsmes peakomponent, mis viitab sellele, et antud komponent sisaldab insuldiriskiga seotud informatsiooni, mida teised komponendid ei kirjelda. Tabelis 6 on esitatud Coxi mudeli hinnangud, kus on lisatud seitsmes peakomponent. PC_7 mõju Coxi mudelis avaldub läbi selle kordaja $\beta_{PC_7} = \log(0,71) < 0$, mistõttu suurem PC_7 väärtus vähendab lineaarset riskiskoori ning seeläbi ka suhtelist insuldiriski.

Tabel 6: Coxi mudel koos seitsmenda peakomponendiga.

Tunnus	HR (riskisuhe)	95% usaldusintervall	p-väärtus
$scale(vanus)$	3,38	(3,17; 3,61)	< 0,0001
$scale(KMI)$	1,17	(1,11; 1,23)	< 0,0001
$I(suitsetamine)$	1,20	(1,06; 1,35)	0,0029
$I(kõrgharidus)$	0,76	(0,67; 0,87)	< 0,0001
PC_7	0,71	(0,59; 0,86)	0,0006

Seitsmenda peakomponendi tunnuste laadungid on esitatud tabelis 7. Laadungite põhjal iseloomustab komponent eelkõige *asfaldi* positiivset ning *linnasuse* ja NO_2 negatiivset panust. Tulemus sarnaneb varasemale analüüsile, kus Coxi mudelisse *linnasuse* ja *asfaldi* samaaegsel kaasamisel osutusid mõlemad tunnused statistiliselt oluliseks.

Tabel 7: Seitsmenda peakomponendi tunnuste laadungid.

Tunnus	Laadung
<i>Asfalt</i>	0,866
<i>Linnasus</i>	-0,322
NO_2	-0,321
<i>NDVI</i>	0,167
<i>MSAVI</i>	0,097
<i>MSAVI_SD</i>	0,055
$PM_{2,5}$	-0,041
<i>NDVI_SD</i>	-0,036
O_3	0,024

Kuigi seitsmes peakomponent seletab vaid väikese osa keskkonnategurite koguvarieeruvusest (lisa 4), võib see viidata sellele, et seos insuldiriskiga ei paikne peamistes tunnuste hajuvussuundades, vaid täpsemas keskkonnategurite kombinatsioonis. Siiski tuleb arvestada, et kõrgema järgu peakomponendid on tundlikumad juhuslikule mürale, mistõttu käsitletakse saadud tulemust ettevaatlikult [11]. Lisas 5 on ülevaade kõigi peakomponentide laadungitest.

Peakomponentanalüüsi tulemused viitavad sellele, et insuldiriskiga ei ole seotud üksikud keskkonnategurid eraldi, vaid eelkõige nende kombinatsioonid, mis avalduvad selgemalt kui algtunnuste eraldi modelleerimisel. Selline tulemus on kooskõlas eksposoomi käsitleusega kui omavahel korreleeritud tegurite süsteemiga, kuid tuleb arvestada, et tuvastatud seos põhineb andmepõhisel komponendil ning võib sõltuda konkreetse kohordi struktuurist. Seetõttu ei käsitleta saadud komponenti üksnes kirjeldava tulemuse, vaid potentsiaalse ennustava tunnuse, mille praktilist rakendatavust on võimalik hinnata riskiskoori kontekstis.

3.5 Riskiskoor

Käesolevas peatükis konstrueeritakse I kohordi andmetel hinnatud Coxi mudelite põhjal riskiskoorid ning hinnatakse nende ennustusvõimet II kohordi andmetel. Vaatluse all on kolm mudelit: ainult üldtunnuseid sisaldav mudel, kahe keskkonnateguriga mudel (*asfalt* ja *linnasus*) ning seitsmendal peakomponendil põhinev mudel.

Kõigi mudelite korral defineeritakse riskiskoor Coxi mudeli lineaarse prediktori alusel, kasutades I kohordi andmetel hinnatud kordajaid. Selline lähenemine võimaldab hinnata, kas leitud seosed on rakendatavad ka sõltumatul andmestikul ning vältida testandmete kasutamist mudeli parameetrite hindamisel. Riskiskooride diskrimineerimisvõimet hinnati II kohordi andmetel kooskõlaindeksi (C-indeks) abil.

Kahe keskkonnateguriga mudeli (*asfalt* ja *linnasus*) põhjal saadud riskiskoori C-indeks oli II kohordis 0,8246 ning peakomponendil põhineva mudeli korral 0,8251. Võrdluseks hinnati ka ainult üldtunnustel põhineva mudeli diskrimineerimisvõimet, mille C-indeks oli 0,8258. Seega ei parandanud keskkonna-

tegurite lisamine riskiskoori diskrimineerimisvõimet võrreldes ainult üldtunnuseid sisaldava mudeliga.

Tulemus viitab sellele, et kuigi mõned keskkonnategurid osutusid Coxi mudelis statistiliselt oluliseks, sisaldavad üldtegurid käesolevas andmestikus juba suure osa insuldiriski kirjeldavast informatsioonist. Samuti võib see viidata sellele, et kasutatud eksposoomi tunnused ei kirjelda piisavalt täpselt keskkonnamõju, näiteks on puudu ülipeened osakesed, millel on suurem potentsiaal erinevatesse keha kudedesse jõuda, või nende mõju on võrreldes üldtunnustega suhteliselt väike [13]. Täpsem ülevaade peenosakeste võimalikust mõjust inimese tervisele on toodud lisas 6.

Täiendavalt võrreldi riskiskooride jaotust insuldi saanud indiviidide ja kogu kohordi lõikes II kohordis (vt tabel 8). Tabelist nähtub, et insuldi saanud indiviidide riskiskoori jaotus on kõigi mudelite korral nihkunud suuremate väärtuste suunas võrreldes kogu kohordiga. Näiteks on paarismudeli korral insuldijuhtude mediaanne riskiskoor ligikaudu 1,52, samas kui kogu kohordis on see ligikaudu $-0,15$. Sarnane muster ilmneb ka peakomponendil põhineva ning ainult üldtunnuseid sisaldava mudeli korral.

Tabel 8: Coxi mudelite põhjal arvatatud riskiskooride jaotus teises kohordis vastavalt mudeli tüübile (paaris vastab asfaldi ning linnasusega mudelile).

	Min	1. kv	Mediaan	Keskmine	3. kv	Max
Paaris (kogu)	-2,8896	-1,1206	-0,1450	-0,0511	0,9356	4,4775
PC_7 (kogu)	-2,8145	-1,1237	-0,1496	-0,0511	0,9368	4,5095
Üldtunnused (kogu)	-2,6582	-1,1207	-0,1470	-0,0511	0,9336	4,5547
Paaris (insult)	-2,2553	0,7988	1,5181	1,3838	2,1432	3,6213
PC_7 (insult)	-2,1705	0,7835	1,5128	1,3870	2,1545	3,5627
Üldtunnused (insult)	-2,4401	0,9811	1,7153	1,5517	2,3290	3,8445

Kuigi saadud tulemused näitavad, et riskiskoorid suudavad indiviide insuldiriski alusel järjestada, tuleb arvestada, et nende ennustusvõime võib olla mõjutatud kasutatud andmestiku eripäradest, sealhulgas haigusjuhtude arvust ning valitud keskkonnateguritest. Seetõttu ei saa saadud tulemust käsitleda lõpliku hinnanguna mudelite üldistatavusele, vaid pigem indikatsioonina, et üldistel riskiteguritel põhinevad mudelid sisaldavad informatsiooni insuldiriski kohta.

Oluline on rõhutada, et riskiskoor ei võimalda täpselt ennustada individuaalse insuldi tekkimist. Kõrge riskiskoor ei tähenda, et indiviidil tekib kindlasti insult, vaid pigem seda, et ta kuulub suurema riskiga rühma. Käesoleva analüüsi põhjal ilmneb eelkõige seos selle vahel, et insuldi saanud indiviididel on keskmiselt kõrgem riskiskoor kui kogu populatsioonis.

4 Tulemuste arutelu

Euroopa kohortide avaldatud hinnangutel [3] põhinevate riskiskooride kasutamine ei andnud Eesti andmetel häid tulemusi. Nii administratiivsete kui ka traditsiooniliste kohortide põhjal konstrueeritud suhtelised riskiskoorid ei osutunud Coxi mudelites statistiliselt olulisteks ega parandanud mudelite diskrimineerimisvõimet. See võib viidata sellele, et Euroopa kohortide põhjal hinnatud eksposoomi mõjud ei pruugi Eesti populatsioonile otse üle kanduda ning tekkis vajadus uurida keskkonnategurite mõju Eesti andmetel eraldi.

Tulemused näitasid, et klassikalised insuldi riskitegurid kirjeldavad insuldiriski Eesti geenivaramu andmetel hästi. Vanus, kehamassiindeks ja suitsetamine olid seotud kõrgema insuldiriskiga ning kõrgharidus väiksema riskiga. Sellised tulemused on kooskõlas üldiste teadmistega insuldi riskifaktoritest, mille järgi vanust, ülekaalu ja suitsetamist peetakse olulisteks insuldiriski mõjutavateks teguriteks [8, 10]. Geenivaramu andmetes ei olnud otsest sotsiaalmajanduslikku seisundit kirjeldavat tunnust. Käesolevas töös käsitleti haridustaset ühe võimaliku kaudse näitajana inimese sotsiaalmajanduslikust taustast ja elustiilist. Kuna kõrgharidus oli seotud väiksema insuldiriskiga, võib see viidata sellele, et kõrgema haridustasemega indiviididel võivad olla keskmiselt tervislikumad eluviisid.

Keskkonnategurite eraldi modelleerimisel Coxi mudelites ei osutunud ükski kasutatud eksposoomitunnus statistiliselt oluliseks insuldiriski kirjeldajaks, kuigi varasemates uuringutes on õhusaastet ja linnakeskkonda seostatud kõrgema insuldiriskiga [3, 4]. Samas on eksposoomitunnused sageli omavahel tugevalt korreleeritud ning nende mõju ei pruugi avalduda üksiktunnuste tasemel, mistõttu uuriti edaspidi keskkonnategurite koosmõjusid.

Koosmõjusid modelleeriti nii kahe keskkonnateguri samaaegsel kaasamisel

Coxi mudelitesse kui ka peakomponentanalüüsi abil. Mõlemal juhul ilmnes sarnane muster, kus olulisimaks osutus linnasuse ja vett mitteläbilaskvate pindade kombinatsioon. Peakomponentanalüüsis osutus statistiliselt oluliseks ainult seitsmes peakomponent, mida iseloomustasid peamiselt *asfaldi* positiivne ning *linnasuse* ja NO_2 negatiivne laadung. Samas seletas peakomponent vaid väikest osa keskkonnategurite varieeruvusest ja nii selle põhjal konstrueeritud riskiskoor kui ka kahe keskkonnateguriga Coxi mudeli riskiskoor ei parandanud mudelite diskrimineerimisvõimet II kohordi põhjal.

Analüüsi üheks oluliseks piiranguks on kasutatud eksposoomitunnuste valik. Käesolevas töös oli kasutada vaid piiratud hulk väliskeskkonna näitajaid ning analüüsist jäid välja näiteks müratase, siseõhu kvaliteet, töölased kokkupuuted ja liikuvusmustrid. Samuti kasutati peenosakeste kirjeldamiseks ainult tunnust $PM_{2,5}$, kuigi varasemates uuringutes on leitud, et ka PM_{10} ja ultrapeened osakesed (UFP ehk *ultrafine particles*) võivad olla seotud südameveresoonkonna ja närvisüsteemi haigustega [13, 14].

Käesolevas töös kasutatud meetodika ei ole ainus võimalik viis eksposoomi mõju uurimiseks. Peakomponentanalüüs võimaldas vähendada tunnuste mõõdet ning käsitleda eksposoomi tegureid koosmõjus, kuid kuna PCA ei arvesta sõltuvat tunnust, ei pruugi saadud komponendid olla optimaalsed insuldiriski kirjeldamiseks. Edasistes uuringutes võiks kaaluda ka teisi lähenemisi, näiteks Elastic Net meetodit või eksposoomitunnuste temaatilist grupeerimist [7]. Samuti võiks uurida eksposoomi mõju erinevates alarühmades, näiteks kõrgema geneetilise riskiga indiviidide seas eraldi. Varasemates uuringutes on kirjeldatud ka geneetiliste ja keskkonnategurite koosmõju südame-veresoonkonna haiguste riskis [15].

Kokkuvõte

Bakalaureusetöö eesmärk oli uurida eksposoomi tunnuste ja insuldiriski seoseid Eesti geenivaramu andmete põhjal ning hinnata, kas keskkonnategurite põhjal on võimalik konstrueerida insuldiriski kirjeldav riskiskoor. Töös kasutati Tartu Ülikooli Eesti geenivaramu tervise- ja küsimustikuandmeid ning EXPANSE projekti raames koostatud keskkonnaandmeid.

Esmalt hinnati, kas de Bont *et al.* Euroopa uuringu põhjal konstrueeritud eksposoomi riskiskoor sobib Eesti andmetele [3]. Teiste kohortide avaldatud andmetel põhinev riskiskoor ei osutunud Coxi mudelites statistiliselt oluliseks ega parandanud mudelite diskrimineerimisvõimet. Seetõttu ei saanud eeldada, et varasemas Euroopa uuringus leitud eksposoomi mõjud kanduvad Eesti geenivaramu andmetele otseselt üle.

Seejärel koostati Coxi võrdeliste riskide mudelid Eesti andmete põhjal. Klassikalistest riskiteguritest olid vanus, kehamassiindeks ja suitsetamine seotud suurema insuldiriskiga ning kõrgharidus väiksema insuldiriskiga. Need tulemused on kooskõlas varasemate teadmistega insuldi riskiteguritest [8, 10]. Üksikute keskkonnategurite lisamisel mudelisse ei osutunud ükski eksposoomitunnus statistiliselt oluliseks.

Kuna eksposoomitunnused võivad olla omavahel seotud, uuriti lisaks keskkonnategurite koosmõju. Paarikaupa mudelites osutus oluliseks linnasuse ja vett mitteläbilaskvate pindade kombinatsioon. Peakomponentanalüüsis osutus oluliseks seitsmes peakomponent, mis kirjeldas samuti peamiselt linnasuse ja vett mitteläbilaskvate pindade vahelist erinevust. Kuna tegemist oli kõrgema järgu peakomponendiga, tuleb seda tulemust tõlgendada ettevaatlikult [11].

Töö viimases osas võrreldi kolme riskiskoori: ainult üldtunnustel põhinevat

riskiskoori, linnasust ja vett mitteläbilaskvaid pindu sisaldavat riskiskoori ning seitsmendal peakomponendil põhinevat riskiskoori. Teise kohordi andmetel ei parandanud keskkonnategurite lisamine riskiskoori diskrimineerimisvõimet. Seega ei andnud kasutatud eksposoomitunnused käesolevas töös klassikalistele riskiteguritele põhinevale insuldiriski mudelile praktilist lisaväärtust.

Tulemuste tõlgendamisel tuleb arvestada, et keskkonnategurite seose puudumist insuldiriskiga ei saa käesoleva töö põhjal lõplikult järeldada. Varasemates uuringutes on leitud seoseid õhusaaste, linnakeskkonna ja insuldi tekke vahel [3, 4]. Käesoleva töö tulemused viitavad sellele, et kasutatud eksposoomitunnuste, valimi ja meetoodika korral ei olnud keskkonnategurite lisapanus riskiskoori ennustusvõimes selgelt eristatav. Edasistes analüüsidest võiks kasutada täpsemaid ja mitmekesisemaid eksposoomi tunnuseid, näiteks ultrapeeneid osakesi (UFP) ja PM_{10} , uurida tunnuseid temaatiliste rühmadena, hinnata keskkonnategurite mõju eraldi erinevates riskirühmades ning käsitleda ka geneetilise riski ja keskkonnategurite võimalikku koosmõju.

Kasutatud allikad

- [1] Roth, G. A., C. Johnson, A. Abajobir, F. Abd-Allah, S. F. Abera *et al.* „Global, Regional, and National Burden of Cardiovascular Diseases for 10 Causes, 1990 to 2015“. *Journal of the American College of Cardiology* 70.1 (2017), lk. 2–25. DOI: [10.1016/j.jacc.2017.04.052](https://doi.org/10.1016/j.jacc.2017.04.052).
- [2] Wild, C. P. „Complementing the genome with an “exposome”: the outstanding challenge of environmental exposure measurement in molecular epidemiology“. *Cancer Epidemiology, Biomarkers & Prevention* 14.8 (2005), lk. 1847–1850. DOI: [10.1158/1055-9965.EPI-05-0456](https://doi.org/10.1158/1055-9965.EPI-05-0456).
- [3] Bont, J. de, R. Pickford, C. Åström, F. Coloma, K. Dimakopoulou *et al.* „Mixtures of long-term exposure to ambient air pollution, built environment and temperature and stroke incidence across Europe“. *Environment International* 179 (september 2023), lk. 1–11. DOI: [10.1016/j.envint.2023.108136](https://doi.org/10.1016/j.envint.2023.108136). (Vaadatud 17.03.2026).
- [4] Shah, A. S. V., K. K. Lee, D. A. McAllister, A. Hunter, H. Nair *et al.* „Short term exposure to air pollution and stroke: systematic review and meta-analysis“. *BMJ* 350 (2015), h1295. DOI: [10.1136/bmj.h1295](https://doi.org/10.1136/bmj.h1295).
- [5] EXPANSE Project. *EXPANSE projekti veebileht*. URL: <https://expanseproject.eu> (vaadatud 06.05.2026).
- [6] Vermeulen, R., E. L. Schymanski, A.-L. Barabási ja G. W. Miller. „The exposome and health: Where chemistry meets biology“. *Science* 367.6476 (2020), lk. 392–396. DOI: [10.1126/science.aay3164](https://doi.org/10.1126/science.aay3164).
- [7] Billionnet, C., D. Sherrill ja I. Annesi-Maesano. „A review of common statistical methods for dealing with exposome-health associations“. *Environment International* 186 (2024), lk. 108675. DOI: [10.1016/j.envint.2024.108675](https://doi.org/10.1016/j.envint.2024.108675).

- [8] Kurvits, T., G. Sõerunurk, K. Suur, L. Pääbo, D. Palumäe *et al.* *Elu pärast insulti*. Tartu Ülikooli Kliinikum, 2021. URL: https://www.kliinikum.ee/wp-content/uploads/2021/02/kliinikum_insult_web.pdf (vaadatud 11.03.2026).
- [9] Tervise ja Heaolu Infosüsteemide Keskus. *RHK-10 klassifikaator*. URL: <https://rhk.sm.ee> (vaadatud 06.05.2026).
- [10] *Avoiding Heart Attacks and Strokes: Don't Be a Victim – Protect Yourself*. World Health Organization, 2005. ISBN: 9241546727. URL: <https://world-heart-federation.org/wp-content/uploads/2005/04/publications-avoiding-english.pdf> (vaadatud 11.03.2026).
- [11] James, G., D. Witten, T. Hastie ja R. Tibshirani. *An Introduction to Statistical Learning: With Applications in R*. 2. väljaanne. Springer, 2021.
- [12] Bland, J. M. ja D. G. Altman. „Multiple significance tests: the Bonferroni method“. *BMJ* 310.6973 (1995), lk. 170. DOI: [10.1136/bmj.310.6973.170](https://doi.org/10.1136/bmj.310.6973.170).
- [13] Schraufnagel, D. E. „The health effects of ultrafine particles“. *Experimental & Molecular Medicine* 52 (2020), lk. 311–317. DOI: [10.1038/s12276-020-0403-3](https://doi.org/10.1038/s12276-020-0403-3).
- [14] Kim, K., E. Kabir ja S. D. Kabir. „A review on the human health impact of airborne particulate matter“. *Environment International* 74 (2015), lk. 136–143. DOI: [10.1016/j.envint.2014.10.005](https://doi.org/10.1016/j.envint.2014.10.005).
- [15] Hartiala, J. A., W. S. Schwartzman, Y. C. Klimentidis *et al.* „Gene-environment interactions for cardiovascular disease“. *Current Atherosclerosis Reports* 23.12 (2021), lk. 75. DOI: [10.1007/s11883-021-00974-9](https://doi.org/10.1007/s11883-021-00974-9).

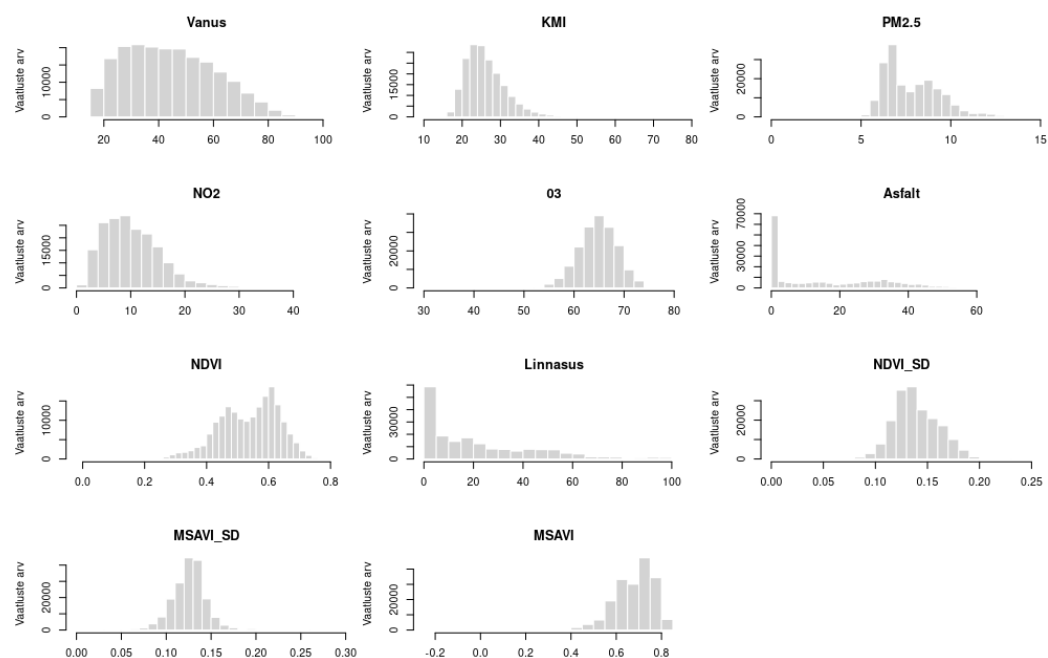
Lisad

Lisa 1

Tabel 9: Puuduvate väärtuste ülevaade tunnuste kaupa

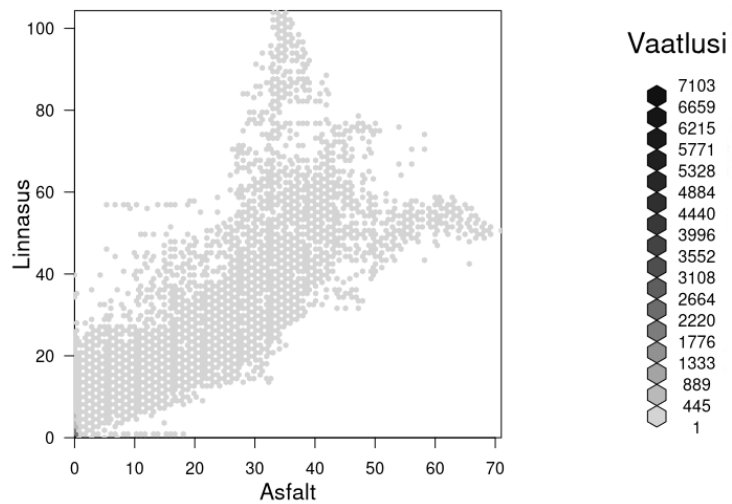
Tunnus	Kirjeldus	Puuduvad väärtused
Vanus	Vanus geenivaramuga liitumisel	0 (0,00%)
Sugu	Bioloogiline sugu	0 (0,00%)
KMI	Kehamassiindeks	530 (0,27%)
Suitsetamine	Suitsetamise staatus	1011 (0,51%)
Haridus	Haridustase (põhiharidus, keskharidus, kõrgharidus)	1292 (0,66%)
PM2,5	Peenosakeste kontsentratsioon	55 (0,03%)
NO2	Lämmastikdioksiidi kontsentratsioon	55 (0,03%)
O3	Osooni kontsentratsioon	55 (0,03%)
NDVI	Keskmine roheluse indeks 1000 m raadiuses	709 (0,36%)
NDVI sd	Roheluse indeksi standardhälve 1000 m raadiuses	730 (0,37%)
MSAVI	Keskmine modifitseeritud roheluse indeks 1000 m raadiuses	709 (0,36%)
MSAVI sd	Modifitseeritud roheluse indeksi standardhälve 1000 m raadiuses	730 (0,37%)
Asfalt	Vett mitteläbilaskvate pindade osakaal 1000 m raadiuses	37 (0,02%)
Linnasus	Piirkonna linnastumise tase	32 (0,02%)

Lisa 2

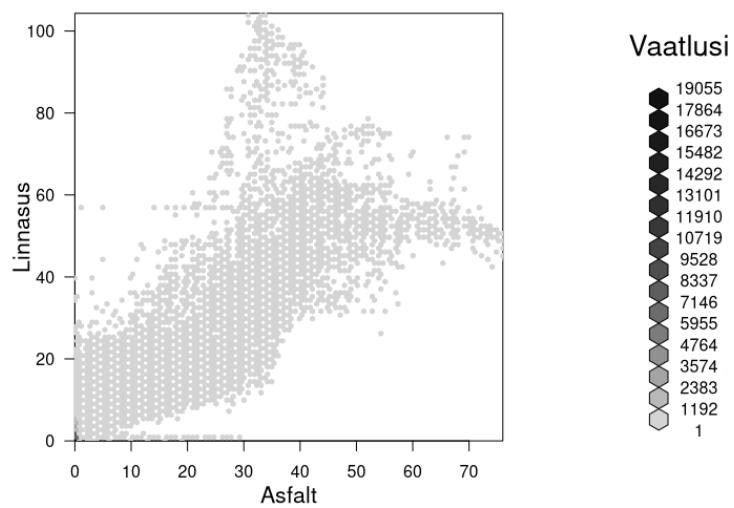


Joonis 1: Pidevate tunnuste histogrammid.

Lisa 3



Joonis 2: Linnasuse ja asfaldi hajuvusdiagramm kohordis 1.



Joonis 3: Linnasuse ja asfaldi hajuvusdiagramm kohordis 2.

Lisa 4

Tabel 10: Peakkomponentide standardhälbed ja seletatud dispersioon.

	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9
Standard- hälve	2,27	1,36	0,98	0,71	0,54	0,37	0,29	0,22	0,06
Osakaal dispersioonist	0,571	0,204	0,106	0,057	0,032	0,016	0,009	0,005	0,000
Kumulatiivne dispersioon	0,571	0,775	0,881	0,937	0,970	0,985	0,994	1,000	1,000

Lisa 5

Tabel 11: Peakomponentide tunnuste laadungite ülevaade.

Tunnus	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9
<i>Asfalt</i>	0,422	0,054	-0,021	0,036	-0,212	-0,036	0,866	0,138	0,038
<i>NDVI</i>	-0,401	-0,204	0,041	-0,271	-0,427	-0,068	0,167	-0,148	-0,697
<i>NDVI_SD</i>	-0,279	0,504	0,323	-0,069	-0,142	-0,008	-0,036	0,733	-0,017
<i>PM_{2,5}</i>	0,011	0,394	-0,806	-0,431	-0,023	0,068	-0,041	0,042	-0,012
<i>NO₂</i>	0,411	0,014	0,016	-0,075	-0,343	-0,776	-0,321	0,043	-0,011
<i>O₃</i>	-0,306	0,115	-0,382	0,804	-0,289	-0,128	0,024	-0,008	0,004
<i>MSAVI</i>	-0,397	-0,227	0,020	-0,279	-0,428	-0,054	0,097	-0,091	0,714
<i>MSAVI_SD</i>	-0,030	0,693	0,307	-0,009	-0,096	-0,012	0,055	-0,640	0,044
<i>Linnasus</i>	0,402	-0,021	0,044	0,068	-0,599	0,607	-0,322	0,028	-0,024

Lisa 6

Tabel 12: Peenosakeste võimalik mõju inimese organismile Kyung Kim *et al.* uuringu põhjal [14].

Osakeste tüüp	Suurus	Võimalik levik organismis	Võimalik mõju tervisele
PM_{10}	$< 10 \mu\text{m}$	Peamiselt ülemised hingamisteed	Hingamisteede ärritus ja põletik
$PM_{2,5}$	$< 2,5 \mu\text{m}$	Võivad jõuda kopsualveolidesse ja vereringesse	Seostatud südame-veresoonkonna ja hingamisteede haigustega
UFP	$< 0,1 \mu\text{m}$	Võivad liikuda vereringesse ning potentsiaalselt jõuda kesknärvisüsteemi	Seostatud neuroinflammatsiooni ja oksüdatiivse stressiga

Lihtlitsents lõputöö reprodutseerimiseks ja üldsusele kättesaadavaks tegemiseks

Mina, Kaur-Erik Suurorg,

1. annan Tartu Ülikoolile tasuta loa (lihtlitsentsi) minu loodud teose "Eksposoomi mõju insuldi tekkele", mille juhendajad on Jaanika Kronberg ja Krista Fischer, reprodutseerimiseks eesmärgiga seda säilitada, sealhulgas lisada digitaalarhiivi DSpace kuni autoriõiguse kehtivuse lõppemiseni.
2. Annan Tartu Ülikoolile loa teha punktis 1 nimetatud teos üldsusele kättesaadavaks Tartu Ülikooli veebikeskkonna, sealhulgas digitaalarhiivi DSpace kaudu Creative Commons'i litsentsiga CC BY NC ND 3.0, mis lubab autorile viidates teost reprodutseerida, levitada ja üldsusele suunata ning keelab luua tuletatud teost ja kasutada teost ärieesmärgil, kuni autoriõiguse kehtivuse lõppemiseni.
3. Olen teadlik, et punktides 1 ja 2 nimetatud õigused jäävad alles ka autorile.
4. Kinnitan, et lihtlitsentsi andmisega ei riku ma teiste isikute intellektuaalomandi ega isikuandmete kaitse õigusaktidest tulenevaid õigusi.

Kaur-Erik Suurorg

12.05.2025