

Tartu Ülikool
Sotsiaalteaduse valdkond
Psühholoogia instituut

Marta Mändla
INDIVIDUAALSETE JA KESKKONNA TEGURITE SEOS ÜLEKAALULISUSE
RISKIGA
Magistritöö

Juhendaja: Kairi Kreegipuu, PhD

Läbiv pealkiri: Ülekaalulisus ja keskkond

Tartu 2026

Keskkonna ja individuaalsete tegurite seos ülekaalulisuse riskiga

Kokkuvõte

Käesoleva magistritöö eesmärk on uurida, millised individuaalsed ning keskkonna konteksti tunnused on seotud ülekaalulisuse riskiga Eesti populatsioonis. Uurimus põhineb Eesti Rahvastiku Vaimse Tervise Uuringu (RVTU) 2021–2022 andmestikul. Andmeanalüüsis kasutati ülekaalulisuse klassifitseerimiseks nii RHK-10 diagnoose (E65-E68) kui ka kehamassiindeksit (edaspidi KMI). Andmeanalüüsis kasutasin logistilist regressiooni ning mudelite stabiilsuse ja usaldusväärsuse saavutamiseks kasutasin lisaks erinevaid regulatsioonimeetodeid.

Tulemused näitasid järjekindlalt, et individuaalsed tunnused selgitavad ülekaalulise gruppi kuulumist kõige tugevamalt. Kõige tugevamaks ennustajaks osutus vanus, kus risk suurenes vanuse kasvades ning selle järgnes sugu. Ainult KMI-põhises mudelis andsid olulise panuse perekondlikud tunnused ning haridus ja sissetulek olid nõrgad ja ebajärjekindlad ennustajad. Keskkonna tunnuste ennustusvõime oli enamasti nullilähedane. See tulemus ei välista keskkonnategurite rolli, vaid viitab pigem, et keskkondlikud faktorid võivad avalduda läbi mediaator tunnuste, mida valitud andmestik ei hõlmanud.

Töö põhilised piirangud on diagnoosipõhise ülekaalulisuse madal esinemissagedus, KMI võimalik ebatäpsus ning keskkonna konteksti prediktorite kõrge omavaheline korrelatsioon. Lisaks piirab järelduste üldistatavust mudelite madal seletusvõime, mis peegeldab ülekaalulisuse välja arenemise multifaktoriaalsust. Edasised uuringud võiksid suunata tähelepanu tegelikku ruumikasutust peegeldavatele keskkonnatunnustele ning medieerivatele faktoritele.

Märksõnad: ülekaalulisus, keskkonna konteksti tegurid, logistiline regressioon

Associations Between Environmental and Individual Factors and Overweight

Abstract

The aim of this master's thesis is to examine which individual and environmental contextual characteristics are associated with the risk of being overweight in the Estonian population. The study is based on data from the Estonian National Mental Health Study (2021–2022).

Overweight and obesity were classified using both ICD-10 diagnoses (E65–E68) and body mass index (BMI). Logistic regression analysis was applied, and various regularization techniques were used to ensure model stability and reliability.

The results consistently indicate that individual characteristics are the strongest predictors of belonging to the overweight group. Age emerged as the most influential predictor, with the risk increasing with age, followed by gender. Family-related characteristics contributed significantly only in the BMI-based model, whereas education and income were weak and inconsistent predictors. Environmental characteristics showed no statistically significant independent spatial effects. This does not exclude the role of environmental factors but suggests that their influence may be mediated through variables not captured in the dataset.

The main limitations of the study include the low prevalence of diagnosis-based overweight, potential inaccuracies in BMI, and high multicollinearity among spatial predictors. In addition, the relatively low explanatory power of the models limits the generalizability of the findings and reflects the multifactorial nature of excess weight. Future research should focus on environmental characteristics that better capture actual spatial behaviour, as well as potential mediating factors.

Keywords: overweight, environmental factors, logistic regression

Ülekaalulisuse definitsioon

Maailma Tervise Organisatsioon (edaspidi WHO) defineerib ülekaalulisust (teisisõnu eelrasvumine) kui kehamassiindeksit (edaspidi KMI) vahemikus 25.0-29.9 ning rasvumist 30.0-.. (WHO, 2025). Tegemist on suuresti sama probleemi erineva tõsidusastmega ning meditsiinilises praktikas neid kahte nähtust rangelt ei eristata (Ravijuhendite Nõukoda, 2019). Nii teadustöodes kui meditsiinis käsitletakse termineid tihti sünonüümidenä ja minu töös kasutan edaspidi terminit “ülekaalulisus”.

Teadus- ja tervishoiukirjanduses ning õppematerjalide raames on välja töötatud erinevaid mudelid ülekaalulisust mõjutavate tegurite selgitamiseks. Tervisearengu Instituut (edaspidi TAI) seostab ülekaalulisust järgmiste teguritega: individuaalne käitumine, psühholoogilised, bioloogilised, sotsiaal-demograafilised, keskkondlikud, sotsio-majanduslikud ning sotsiaalsed tegurid (TAI, 2025).

Ülekaalulisuse levimus ja seos keskkonna teguritega

Ülekaalulisus on maailmas aina laialdasemalt levinum probleem. TAI 2025. aasta raporti põhisel on üle poolte Eesti ja Euroopa täiskasvanutest ülekaalulised ning see määr suureneb iga aastaga. 2022. aastal avaldatud andmed näitavad, et vanuses 16-64 on Eestis ülekaalulised või rasvunud 52.8% elanikest, võrreldes aastaga 2006, kus ülekaalulisi ja rasvunud oli 49.8% (TAI, 2015). Samuti on aastatega suurenenud liigsete kilode määr. Aastal 2014 oli vanuses 16-64 ülekaalulistel naistel 1.6 ja meestel 4.4 lisakil, kuid aastal 2024 oli naistel 2.2 ja meestel 6.2. Ühtlasi on rasvunudel tõusnud lisakilod 27.4 kilolt 29.3 kilole. Lisakilod arvutatakse kehamassi indeksi (soovituslik kaal - tegelik kaal) põhjal ja ei representeerigi seega tegelikku rasvkoe hulka, vaid statistilist keskmist. (TAI, 2025).

Ülekaalulisuse välja kujunemisel on oluline geneetiline komponent, kuid keskkonna faktorite rolli peetakse ülekaalulisuse välja arenemisel mõjukaimaks (Grundy, 1998). Kõrge kalorisisaldusega toidu tarbimine ja energia säästmine on evolutsioonilisest vaatest adekvaatne ellujäämisstrateegia. Lääne ühiskondades on kõrge kalorisisaldusega toidu kättesaadavus väga kõrge ning on palju võimalusi vältida füüsiliselt koormavaid tegevusi. Eelneva taustal on inimene suuresti säilitanud evolutsiooniliselt adekvaatsed strateegiad, kuigi keskkonnas ellu jäämine seda enam ei vaja. Seetõttu pole strateegia tarbida kõrge kalorisisaldusega toitu ja olla võimalikult energiat kokku hoidev enam kohastumuslik, vaid tekitab probleeme nagu ülekaalulisus. Ülekaalulisuse ära hoidmiseks on vaja adekvaatset

energia tarbimise ja kulutamise režiimi. See kui palju energiat tarbitakse ja kui aktiivset eluviisi elatakse saab kokku võtta terminiga käitumuslikud faktorid. Käitumuslikke faktorid omakorda mõjutavad paljud erinevad tegurid, mis jagunevad individuaalseteks, sotsiaalseteks ja keskkonna teguriteks. (Papas jt., 2007; Nicolaidis, 2019)

Nagu mainitud mängivad olulist rolli ülekaalulise välja kujunemisel käitumuslikud tegurid, mis omakorda on mõjutatud individuaalsetest, sotsiaalsetest ja keskkonna faktoritest. Individuaalsed tegurid hõlmavad endas geneetilisi, sotsiaaldemograafilisi ja sotsiaalmajanduslikke tegureid. Sotsiaalsed tegurid on seotud indiviidi sotsiaalse keskkonnaga ja neist tulenevate mõjutustega. Näiteks perekonna kontekst ja eakaaslaste arvamus; ühiskondlikud väärtused ja hinnangud ning sotsiaalset tähendust omavad keskkonna omadused (tajutud turvalisus, välimus jms) (Papas jt., 2007). Keskkonna faktorid on nii looduslikud kui ka inimese poolt loodud keskkonda tunnused. “Ülekaalulisuse epideemia” uuringud keskenduvad põhiliselt inimtegevuse tagajärjel tekkinud keskkonnale ehk täpsemalt “ehitatud keskkonnale”, kuna seda peetakse oluliseks energia tarbimise ja kulutamise tasakaalu mõjutajaks (Papas jt., 2007; Nicolaidis, 2019). Ehitatud keskkonna all mõeldakse nii otseseid tervisele kahjulikke faktoreid nagu kemikaalid ja muud patogeensed agendid kui kaudselt tervisega seotud tegurid nagu transport, elamurajooni tüüp, tööstus jms (Papas jt., 2007; Nicolaidis, 2019).

Järgnevates lõikudes annan ülevaate erinevatest teguritest, mis mängivad rolli ülekaalulise välja arenemisel. Seejärel selgitan käesoleva magistritöö eesmärki.

Individuaalsed omadused

Geneetilised

Arengud geeniuuringute valdkonnas on aidanud leida seoseid geenide ning ülekaalulisuse vahel. Ülekaalulisuse geneetilist tausta on kirjeldatud nii monogeensena kui polügeensena. Monogeenne ülekaalulisus on haruldane nähtus, kus ühe geeni mutatsiooni tulemusel areneb indiviidil välja ülekaalulisus. Vähem haruldane on polügeenne ülekaalulisus, mille korral mitmeid geneetilised faktorid koosmõjus mängivad rolli ülekaalulisuse välja arenemisel. Kuigi ülekaalulisusega on seotud sadu genee, ei käsitle käesolevas magistritöö seda teemat lähemalt. (Mahmoud jt., 2022)

Sotsiaal-demograafilised

Soolised erinevused

Eesti rahvastiku põhjal loodud statistika näitab, et keskmiselt on suurem osakaal mehi ülekaalulised (mehed 62% ja naised 45%) ja neil on rohkem lisakilosid (mehed 6.2 ja naised 2.2 lisakilo) kui naistel (TAI, 2025). See on kooskõlas varasemate leidudega, et kõrgema sisemajandusliku kogutoodanguga riikides on ülekaalulisus meeste hulgas suurem kui naiste, võrreldes madalama SKT-ga riikides, kus ülekaalulisuse seos sooga on vastupidine (Wells, 2018). Selle nähtuse selgitamiseks on välja pakutud nii füsioloogilisi, sotsiaalseid kui käitumuslikke aspekte/tegureid. Näiteks on välja pakutud, et tõenäosus ülekaalulisusele naiste hulgas on madalam, kuna hinnatakse oma kaalu täpsemalt; ollakse kaalu osas rahulolematud ning proovitakse kaalu alandada tihemini (Miscogiuro, 2024).

Sugu käsitletakse sageli mediaatorina teatud keskkonna või käitumuslike faktorite ning ülekaalulisuse vahel. Näiteks on üliõpilastel, kes elavad üksinda (võrreldes nendega, kes elavad perega) suurem risk kaalutõusule. See risk on märkimisväärselt kõrgem naistel (Yamamoto jt., 2023). Sarnased soopõhised erinevusi ülekaalulisuse riskis ilmneb veel magistritöö järgnevates peatükkides.

Vanus

Vanuse ja ülekaalulisuse seosed on keerukad ja muutuvad elu vältel. Üldiselt suureneb ülekaalulisuse esinemissagedus vanuse tõusuga keskeani, selle juures on lapsepõlve ülekaalulisus tugevalt seotud ülekaalulisusega täiskasvanueas (Rundle jt., 2020). Eestis eksisteerib üldine trend, kus ülekaalulisus suureneb vanuse tõusuga (TAI, 2025). Samuti kasvab keskmiste lisakilode määr. Näiteks aastal 2024 leiti, et vanusegrupis 35-39 on keskmiselt kaks lisakilo, kuid vanusegrupis 60-64 9.9 lisakilo. See statistika näitab läbilõiget Eesti rahvastikust ning mitte indiviidide kehakaalu muutumist elu vältel (TAI, 2025). Seega võivad esineda generatsioonide vahelisi erinevusi, mida statistika ei väljenda.

Perekondlik staatus

Maailmas on rohkelt uuritud perekondliku staatuse ja ülekaalulisuse riski vahelised seoseid. Üldine seos on, et partneriga koos elamine ja partneri kõrge KMI suurendab ülekaalulisuse riski (The & Gordon-Larsen, 2009). Need seosed on väga tugevalt seotud kultuurilise konteksti ja sooliste ning sotsiaalmajanduslike erinevustega. Näiteks naised

Kanadas, kellel on madal sotsiaalne osavõtlikkus, on märgatavalt suuremal määral ülekaalulised kuigi sama seos oli meeste puhul vastupidine (Hosseini jt., 2020; Lee jt. 2022). Mehed Jaapanis, kes elavad ja söövad üksinda, on suurema tõenäosusega ülekaalulised kui naised (Tani jt., 2015).

Sotsiaalmajanduslikud tegurid

Kõrgema sissetulekuga riikides on madalama sotsio-majanduslik staatuse (edaspidi SMS) ja kõrgema ülekaalulisuse riski vahel seos (Seifu jt., 2021). Samuti on kõrgem SMS elu jooksul seotud naistel madalama KMI-ga, kuid seos on vähem järjepidev meestel (Newton jt., 2017). Põhjuseid nende seostele on pakutud mitmeid: näiteks madalama SMS-ga perede suurem ligipääs kõrge kalorisaldusega toidule. Teiseks on pakutud, et sotsiaalne ja majanduslik ebavõrdsus vähendab psühholoogilisi ressursse toime tulla raskustega, mis omakorda mõjutab tervisekäitumist ning kõrgem stressitase on seotud kõrgema kalorisaldusega toidu tarbimisega (Claassen jt. 2019). Lisaks on välja pakutud, et stigmad kehakaalu kohta mõjutavad naisi kaalu kaotama ja toitumist piirama rohkem, eriti just kõrgema SMS-ga indiviide (Claassen at al. 2019).

Haridus

Mitmed uuringud näitavad, et kõrgem haridustase on negatiivselt seotud ülekaalulisusega (Cai jt., 2013; Furuki jt., 2023; Böckerman, 2017). Soomes läbiviidud uuring leidis, et haridusaastad on seotud kehamassiindeksiga negatiivselt. Täpsemalt iga haridusaasta üle valimi keskmise oli seotud kehamassi indeksiga $(b)=-0.22$ (Böckerman, 2017). Oluline on välja tuua, et need seosed kehtivad kõrgema sissetulekuga riikides ning sarnaste uurimuste tulemused madalama sissetulekuga riikides on pigem nõrgad või vastupidises suunas (Cohen jt., 2015). Kõrge haridustase on potentsiaalne kaitsefaktor ülekaalulisuse eest, sest see aitab teha tervislikke valikuid toitumisharjumuste ja liikumisaktiivsuse osas (Cai jt., 2013).

Keskkonna tegurid

Elukoha tüüp

Üldpildis on välja toodud, et kortermajade elanikud on väiksema tõenäosusega ülekaalulised ning elamu suurus on negatiivses korrelatsioonis KMI-ga, kuid tulemustel on olulised soolised erinevused. Iisraeli statistikaameti poolt kahe aasta jooksul läbi viidud

uuring leidis, et peredes, kes kolivad mitme pere elamutesest (korterimajad) ühe pere elamutesse (maja) suureneb eeldatav kehamassiindeks meestel (1.04-1.16%), kuid väheneb naistel (1.30-1.46%). Seost selgitati mitme asjaoluga – näiteks traditsioonilise tööjaotusega, kus suuremasse elamusse kolimine suurendab naiste töökoormust. Lisaks on suurem kodu seotud jõukusega, mis omakorda seotud rohkemate vabaaja veetmise võimalustega. Naised kasutavad erinevaid aktiivseid vabaaja võimalusi rohkem kui mehed. Huvitavalt leiti uuringus veel, et kui ühe pere elamusesse kolival pere sünnib laps, keerab see seose vastupidi, mis viitab asjaolule, et isad mängivad lastega rohkem kui elatakse suuremas kodus. (Arbel jt., 2018). Siinkohal on oluline välja tuua, et Iisraeli kultuurikontekst erineb Eesti omast suurel määral. Täpsemalt järgib suur osa Iisraeli rahvastikust konservatiivseid usutavasid, selle juures traditsioonilisi soorolle (Scroope, 2020). Siiski viitavad väljatoodud seosed asjaolule, et elukeskkond on oluline kehalise aktiivsuse mõjutaja, kuid mõjude suund on väga sõltuv paljudest erinevatest teguritest.

Spordiasutuste kaugus

Spordikeskuste kaugus kodust ja ülekaalulisuse riski vahel on positiivne seos, kuid leitud seosed on pigem nõrgad ja statistiliselt ebaolulised. Näiteks Rootsis läbi viidud uuring tõi välja, et kodu ja töö aadressi omavaheline suurem kaugus nii tasuta kui tasulistest spordikeskustest on seotud väiksema treenimissagedusega. Sarnane seos on ka ülekaalulisusega, kuid see seos on statistiliselt ebaoluline. (Raza jt. 2022). Statistiliselt oluline seos on ülekaalulisuse ja sportimisvõimaluste ning kõnniteede kauguse vahel. Täpsemalt, mida kaugemal on kõnniteed ja spordikeskused seda suurem on ülekaalulisuse määr nii linna kui maa keskkonnas (Boehmer, 2006).

Erinevad keskkonna saastatuse tunnused

Tegemist on pigem uue uurimisvaldkonnaga, mistõttu ei saa põhjapanevalt väita, et saastatus on omaette mõjutegur ning mitte piirkonda kirjeldav tegur, kus elavad põhiliselt madala sissetulekuga pered. Siiski on proovitud nimetatud seoseid leida. Näiteks seos õhusaaste ning lapsepõlve ülekaalulisuse ja seeläbi täiskasvanute ülekaalulisuse vahel (Huang jt., 2022). Sarnane seos on ka müra- ja õhusaaste ning ülekaalulisuse riski vahel (Persson jt., 2024). Lisaks on pakutud välja, et valguse saaste häirib ööpäeva rütmi ning melatoniini tootmist ning seeläbi tõstab ülekaalulisuse riski (Han, 2021). Lõhna saaste seost ülekaalulisusega ei ole lähemalt uuritud, kuid pakun välja, et see võib mõjutada ülekaalulisuse riski sarnaselt õhusaastega kuna on sellega suuresti seotud.

Sarnaselt keskkonna saastatuse tunnustega on võimalus, et rohealade puudumine ja tööstuse lähedus pole omaette mõjutegurid, vaid üks osa madalama sissetulekuga perede elukeskkonda kirjeldavatest tunnustest. Siiski on läbi viidud erinevaid uuringuid, et leida, kuidas rohealade kaugus ja rohkus ning tööstuse lähedus on seotud ülekaalulisuse riski iseseisvate muutujatena. Tulemusi on erinevaid, kuid aastal 2020 läbiviidud meta-uuring väidab, et kuigi tegemist on viimase 20 aasta jooksul populaarust kogunud uurimisvaldkonnaga, on kõrgekvaliteediliste uuringute järele puudus (Luo jt., 2020). Tööstuse olemasolu elukeskkonna ning selle seoseid ülekaalulisega uuritakse läbi medieerivate tunnuste nagu näiteks õhusaaste, madalam kõnnitavuse indeks ja madalama populatsiooni, kuid kõrgema hoonete pindala määraga piirkonnad. Tulemused näitavad, et rohkem tööstuslikes piirkondades elavatel indiviididel (nimetatud tunnuste alusel) on kõrgem risk ülekaalulisusele (Luo jt, 2024; de Bont jt., 2021; Dixon jt., 2021).

Rohealad

Olemasolevate andmete põhjal saab väita, et leidub tugev seos kõrgema roheluse¹ indeksi ja madalama ülekaalulisuse riski vahel (Luo jt., 2020). Soome teadusuuringu tulemustel kasvab ülekaalu risk märkimiväärselt, kui elatakse rohkem kui 750 meetri kaugusel parkidest või muudest rohealadest ning rohealadest kaugemale kolimine suurendas ülekaalulisuse riski (Halonen jt., 2014). Lisaks on välja pakutud, et täpset seost rohealade ja ülekaalulisuse vahel ei ole olemasoleva teaduskirjanduse põhjal teada kuid, et rohealad motiveerivad inimesi olema rohkem aktiivsed ning vähendavad stressi (stress on samuti seotud suurema ülekaalulisuse määraga) (Han ja Lee, 2021).

Toidukohtade kaugus

Üldiselt on toidupoodide kaugus kodust seotud ülekaalulisusega negatiivselt, kuna toidupoed pakuvad rohkem tervislikke toitumisvõimalusi. Kiirtoidu restoranide lähedus oli aga positiivselt ja oluliselt seotud ülekaalulisuse määraga. Teistsorti restoranide ja ülekaalulisuse vahel olulist seost ei leitud. (Pineda, 2024)

¹ *NVDI* ehk *Normalized Difference Vegetation Index*- väärtused vahemikus -1 kuni 1 ja näitavad roheluse määra teatud piirkonnas. Madal *NVDI* väärtus on seotud kivise, mullase, lumise pinnaga ning kõrge *NVDI* väärtus tiheda metsa, raba jms (Earth Science Data Systems, 2024).

Sotsiaalsed faktorid

Sotsiaalseid faktoreid on erinevaid. Mõned neist võivad korruga kuuluda nii individuaalsete kui sotsiaalsete tegurite hulka. Näiteks eelnevalt mainitud sotsio-majanduslik ja -demograafiline info; elanike arv kodus; haridus jms. Nende tegurite lahterdamine ei ole praeguse töö raames esmatahtis ning neid eraldi välja ei tooda.

Sotsiaalne faktor, mida pole teiste kategooriate all varasemalt mainitud on ühiskonna, pere ja indiviidi väärtused, uskumused ja normid. Varasemad uuringud on välja toonud, et need mängivad rolli käitumise mõjutamisel (Crielaad jt., 2020; Gasper jt., 2022), kuid magistritöös neid ei uurita.

Eesmärk ja uurimisküsimused

Magistritöö eesmärk on uurida, millised individuaalsed ning keskkonna konteksti tunnused on seotud ülekaalulisuse ja rasvumise riskiga. Tegemist on eksploratiivse tööga, kuna Eesti valimi põhjal pole varasemalt sarnaseid uuringuid tehtud. Siiski on varasemalt sama andmestiku põhjal läbi viidud sarnaseid töid vaimse tervise ning depressiooni ja ärevuse riski seoste leidmiseks (Aganits, 2024; Pikk, 2024) Selleks kasutan Eesti Rahvastiku Vaimse Tervise Uuringu (edaspidi RVTU) raames kogutud subjektiivseid keskkonna ja demograafilisi andmeid ning geokodeeritud keskkonnaandmeid. Lisaks on kokku viidud terviseandmed, millest kasutan RHK-10 diagnoosidest E65-E68 (ülekaalulisus, rasvumine ja muud liigtoitumise vormid). Lisaks diagnoosipõhisele rühmitusele (E65–E68: ülekaalulisus, rasvumine ja teised liigtoitumise vormid) viiakse läbi analüüs kehamassiindeksi (KMI) põhise klassifikatsiooni alusel, eristades normaalkaalu ($KMI < 25$) ja ülekaalu ($KMI \geq 25$). Kahe klassifitseerimismeetodi kasutamine tuleneb andmestiku eripäradest, mis ei võimalda kõiki huvipakkuvaid tegureid mõlema lähenemise puhul ühtselt kaasata. Samas on mõlemal meetodil nii eelised kui ka piirangud. Kuna tegemist on pigem eksploratiivse tööga ning Eesti andmestikul põhinevad varasemad analoogsed uuringud puuduvad, võimaldab kahe lähenemise võrdlus paremini mõista uuritavat valdkonda. Seoste analüüsimisel kasutatakse peamise meetodina regressioonanalüüsi.

Meetod

Uuringu kirjeldus

Käesoleva magistritöö raames kasutan RVTU käigus kogutuid andmeid. Kasutuses on H2020 projekti Emotional Cities sotsiaal-füüsilised andmed, mis põhinevad indiviidide registreeritud elukohal. Lisaks kasutati neile kahele andmekogule kokku viidud indiviidide terviseandmed ehk erinevad diagnoositud häired. Uuring viidi läbi aastatel 2021-2022 kolmes laines (I ja II laine aastal 2021 ning III laine aastal 2022) Tartu Ülikooli ja Tervise Arengu Instituudi poolt. (Eesti Rahvastiku Vaimse Tervise Uuring, 2022.) Kõigi lainete raames täideti küsimustik (millesse viidi sisse muudatusi toimumiskordade vahel) ning I lainel viidi andmestikuga kokku indiviidide Rahvastikuregistris registreeritud koduaadressi ümbruskonna geokodeeritud keskkonnatunnused ning terviseandmed. Terviseandmeid koguti lisaks veel III laine raames. Uuringu põhiline eesmärk oli hinnata Eesti rahvastiku vaimse tervise seisut ja selle muutusi ühe aasta jooksul. Käesolev uuring ei kasuta tervet andmestikku, vaid valisin muutujad tuginedes uuringu eesmärkidele ja varasemale kirjandusele. Magistritöö raames uuritakse ülekaalukaalulisust kasutades kahte erinevat klassifitseerimismeetodit: esiteks RHK-10 diagnoosid E65-E68 (ülekaalulisus, rasvumine ning teised liigtoimumise vormid) ning teiseks KMI alusel.

RVTU uuringust kutsuti osa võtma 20 000 elanikku, kellest ligikaudu 700 indiviidi keeldus koheselt. Eesti jagati 17-ks mittekattuvaks piirkonnaks (maakonnad + kaks suurimat linna) eesmärgiga koguda andmeid, mis oleks võrreldavad piirkondade alusel (ehk vastajaid oleks piirkonniti võrdselt). Andmestikust eemaldasid kõik isikud, kes vastasid eitavalt küsimusele “kas elatakse samal aadressil Rahvastikuregistri järgi”. Lõppandmestikku jõudsid 18 695 indiviidi andmed. Tegemist on esindusliku kihilise juhuvalimiga. Küsimustele vastamine oli vabatahtlik ning võttis aega umbes 30 minutit (Laidra jt. 2023).

Valim

Andmeanalüüsis kasutatud valimi suurus on 18 695 inimest, kellest 9413 (50.4%) on mehed ja 9282 (49.6%) on naised. Vastajate vanused jäävad vahemikku 15-100 aastat, keskmine vanuse on 48 aastat ($SD = 21$ aastat). Analüüsis kasutasin vanusegrupi muutujat, mis koguti RVTU küsimustiku käigus. Valimis on enim esindatud vanusegrupp 70+ aastat (3510) ning kõige vähem vastajaid oli vanusegrupis 15-19 aastat (2093). Igast piirkonnast on vahemikus 1075-1124 vastajat.

Andmete kogumine toimus kolmes laines, mille käigus lisati, eemaldati ja muudeti mõningaid küsimusi. Selleks, et andmeanalüüsi kaasata võimalikult paljusid teemakohaseid kriteeriume, kasutan andmete agregeerimisel kõige liberaalsemat meetodid ehk kui tunnus esineb ühes laines, laiendatakse andmeid kõigi kolme laine ulatuses. Samuti agregeerisin ülekaalulisuse diagnoosi. Täpsemalt terviseandmeid koguti kahel korral– esimene kord katseisikutel diagnoositud häired (käesolevas magistritöö andmestikus vaid RHK-10 diagnoosid ülekaalulisus, rasvumine ja muud liigtoitumise vormid) viimase viie aasta jooksul ning teine kord uuringu jooksul diagnoositud häired (Laidra jt., 2023). Magistritöö andmeanalüüsi eesmärkidel tegin kahest muutujast ning selle kolmest tasemest (ülekaalulisus, rasvumine ning teised liigtoitumise vormid) ühe dihhotoomse muutuja nimega “ülekaalulisus”. See muutuja hõlmas endas 909 indiviidi (5.2% üldvalimist). Ülekaalulisust oli diagnoositud 272 juhul, rasvumist 175 ning teist liigtoitumise vormi 503 juhul (kokku rohkem kui muutujas “ülekaaluline”, sest ühel indiviidil võis esineda kaks erinevat diagnoosi).

Selleks, et arvutada KMI, on vaja teada indiviidi pikkust ja kaalu ning neid muutujaid küsiti RVTU I ja III laines. Katseisikuid, kes võtsid osa vähemalt ühest küsitluse voorust on kokku 6594 (u pooled neist võtsid osa mõlemast lainest). Neist 6277 olid vastanud vähemalt ühe laine jooksul pikkuse ning kaalu küsimusele, mille põhjal sai välja arvutada KMI. Neist 4008 (63.9%) olid WHO klassifikatsiooni järgi (≥ 25) ülekaalulised ning 2269 oli KMI < 25 ning arvestati normaalkaalu gruppi.

Järgnevad kolm alapeatükki kirjeldavad erinevaid andmeanalüüsi kaasatavaid muutujaid lühidalt. Andmestiku loomisprotsessist põhjaliku ülevaate saamiseks soovitan lugeda Pikk (2024) magistritööd.

Sotsiaal- demograafilised andmed

Töös kasutan osalejate Rahvastikuregistri poolt raporteeritud sugu (Laidra jt., 2023) ning enese-raporteeritud vanusegruppi kuulumist. Ühtlasi kasutan RVTU küsimustikus kogutud osalejate eneseraporteeritud sissetulekut ja haridustaset. Agregeerisin muutujad kasutades kolme laine vastustest kõige kõrgemat sissetulekut ning omandatud haridustaset. Osalejate sissetulek on jagatud nelja kategooriasse (kuni 450, 451-850, 851-1400, üle 1401). Perekondlikku staatust kirjeldavaid muutujaid koguti RVTU küsimustikus, vastates küsimustele, kas katseisikul on lapsi; elab üksi või paarilisega.

Lisaks varasemale agregeerimisele, vähendasin muutuja hariduse kaheksa taset (algharidus, põhiharidus, põhiharidus koos kutseharidusega, kesk-või gümnaasiumiharidus, keskeri- või kutseharidus, rakenduslik- või kutse kõrgharidus, kõrghariduse põhiõpe (bakalaureus), kõrghariduse kraadiõpe, (magister, doktor/kandidaat)) kolmeks tasemeks (kuni põhiharidus, gümnaasiumi ja/või kutseharidus, kõrgharidus) lihtsuse ja tõlgendatavuse eesmärkidel.

Subjektiivsed keskkonna näitajad

RVTU uuringu raames paluti osalejatel hinnata oma elukeskkonna kohta mitmeid aspekte. Valisin nendest välja muutujad, mis on kooskõlas käesoleva uuringu eesmärkidega ning omavad varasemate uuringute põhjal relevantsust.

Subjektiivset keskkonda kirjeldavad muutujad koguti, paludes katseisikutel vastata jah/ei küsimusele: “milline on lähiümbruskond, kus veedate märkimisväärselt aega?”. Vastusevariandid olid järgmised: korter, eramu, kesklinn, tööstus, rohekeskkond. Lisaks vastati küsimusele asula populatsiooni kohta (kas asula suurus väiksem kui 1000 inimest; väiksem kui 10 000 inimest või suurem kui 10 000 inimest).

Saastatuse kirjeldajad agrigeeriti kolme tasemelisteks muutujateks: “üldse mitte” ja “veidi”, “mõningal määral”; “palju”, “väga palju”. Gruppide võrdsustamiseks ning andmete dimensionaalsuse vähendamiseks agregeerisin muutujad veelkord kahe-tasemeliseks: “ei ole” ja “on”. Subjektiivsed saastatuse muutujad kirjeldavad elukeskkonna lärmi (inimtegevus) ja müra (tööstus ja transport) taset; tugeva ja vähese valguse esinemist; õhusaaste ja lõhnade määra ning rohealade puudumist.

Geokodeeritud (objektiivsed) keskkonna näitajad

Geokodeeritud keskkonna näitajad pärinevad Maa-ameti Eesti Topograafilisest Andmebaasist ning kirjeldavad osaleja registreeritud koduaadressi ümbruse (kolmes puhvertsoonis: 100 m; 500 m ja 1000 m) kompositsiooni.

Esiteks erinevate asutuste esinemissagedus kolmes puhvertsoonis. Varasemates uuringutes on kasutatud kommertsasutuste (ja ka teiste objektide) esinemissagedust erinevates puhvertsoonides ehitatud keskkonna iseloomustamiseks (James jt., 2014). Nende hulgas on analüüsi kaasatud kauplused, söögikohad, kaubanduskeskused, müügikohad, tanklakauplused, turud, kioskid, büroo- ja ärihooned ning spordirajad. Lisaks veel kaugus

erinevates huviobjektidest, nende hulgas, koolist, apteegist, perearstikeskusest, alkomüügikohas, spordirajast ja supluskohast.

Järgmine kogum geokodeeritud tunnuseid kirjeldab elukoha ümbruse hoonete kompositsiooni. Puhvertsoonides on kirjeldatud eri tüüpi (elu- ja või ühiskondlik hoone; kõrval- või tootmishoone; vundament; vare; ehitatav hoone) majade hulka ja pindala.

Veel on kasutusel andmed elukeskkonna ümbruses oleva teedevõrgustiku kohta. Täpsemalt erinevat tüüpi teede pikkus kolmes puhvertsoonis (tänavad, kõrvalmaanteed, maanteed ning muud teed, kergliiklustee, rada, ramp või ühendustee). Varasemalt on kasutatud teedevõrgustiku pikkust ülekaalulisuse riski hindamisel edukalt (Aquilera jt., 2021).

Looduskeskkonda kirjeldavad tegurid on samuti jagatud kolme puhvertsooni ning kirjeldavad nii neis asuvaid veekogusid (näiteks: järve pindala või jõe pikkus puhvertsoonis) ning lage ja puittaimestiku alasid (ala pindala tsoonis ning vegetatsiooni kõrgust).

Viimane kategooria geokodeeritud keskkonna näitajaid kirjeldavad ümbruskonna majanduslikku olukorda ning need andmed koguti Äriregistrist. Need muutujad on samuti jagatud kolme puhvertsooni järgi ning kirjeldavad maksulaekumisi, töötajate arvu ning ettevõtete arvu elukoha ümbruses.

Andmeanalüüs

Andmeanalüüsi viisin läbi kasutades R tarkvara versiooni 4.5.2 R-Studio rakenduses. Andmeanalüüs tarbes viisin läbi kirjeldava statistika, logistilise regressiooni ja regressiooni analüüsid *Minimax Convecs Penalty* (R pakett *ncvreg*) ja *Elastic Net* (R pakett *glmnet*) reguleerimismeetoditega. Magistritöös esitatud tabelid on loodud kasutades R-Studio tarkvara.

Esimeses andmeanalüüsi voorus võrdlesin indiviidide, kellel on diagnoositud ülekaalulisus, rasvumine või muu liigtoitumise vorm kõigi teistega, kellel nimetatud diagnoosi ei ole. Andmestikust eemaldasid kõik muutujad, millel puudusid vastused rohkem kui 30% juhtudest ning ülejäänud puuduvad andmed imputeeriti, kasutades funktsiooni *na.roughfix()* R-paketist *randomForest*. Maksimaalselt imputeeriti 6% andmetest, mis on kooskõlas parima praktikaga (Eekhout jt., 2014). Veel filtreeriti andmestikust Pearsoni korrelatsioonikordaja abil kõrgelt (piirmäär 0.9) korreleerunud muutujad. Antud filtreerimise

tulemusena eemaldati analüüsist kõik subjektiivsed keskkonna ning mõned sotsiaal-demograafilised ning objektiivsed keskkonna muutujad (alles jäänud muutujad märgitud lisas A). Esmalt viidi läbi logistiline regressioon. Seejärel selleks, et eemaldada mudelist müra ja valida välja vaid olulised muutujad, viidi läbi *Minimax Concave Penalty (MCP)* regressioonanalüüs, mis on sobiv meetod nii muutujate valiku täpsuse kui ka ennustusvõime poolest just mõõduka ja kõrge dimensiooniga andmete puhul (Bag jt., 2022). Kuna sõltuva muutuja jagunemine pole tasakaalus (95% vs 5%) otsustasin eelistada *MCP* reguleerimismeetodit teistega võrreldes, mis on agresiivsem madala signaaliga koefitsientide nulli viimises ja seega eemaldab mudelist müra agresiivsemalt, vähendades üle-sobitumist.

Teises voorus kaasasin analüüsi katseisikud, kes vastasid RVTU küsimustiku I ja/või III lainel kaalu ning pikkuse küsimustele. Juhul kui katseisikud olid mõlemal korral vastanud nii kaalu kui pikkuse küsimusele, arvutati välja kahe väärtuse aritmeetiline keskmine. Põhinedes nendele tulemustele, arvutasin välja kehamassiindeksi kasutades valemit kaal/pikkus^2 . Selle tulemusena saadud KMI muutujas esines äärmuslikke väärtusi, mis eemaldati kasutades IQR-meetodit (Mowbray jt., 2018). Andmeanalüüsi käigus võrreldakse WHO klassifikatsiooni järgi ülekaalulisi ($\text{KMI} \geq 25$) kõigi teistega (muutuja nimega “normaalkaal”). Andmeanalüüsis kasutatakse regulatsioonimeetodit *Elastic Net*. KMI-põhises jaotuses on grupid omavahel võrdsemalt jaotunud (63.1% ülekaaluliste grupp vs 36.9% normaalkaalu grupp), seega otsustasin kasutada *Elastic Net* reguleerimist, sest meetod jätab mudelisse väiksemate koefitsientidega muutujaid ning seega annab võimaluse säilitada suurema arvu muutujaid. Enne analüüsi viidi läbi samad andmete filtreerimissammud, kui esimese vooru juures – eemaldati muutujad, kellel esines rohkem kui 30% puuduvaid andmeid (alles jäänud muutujad märgitud lisas A) ning ülejäänud muutujate puuduvad andmed imputeeriti, kasutades funktsiooni *na.roughfix()* R-paketist *randomForest*. Suurimaks imputeerimise määraks osutus 12.2%, mis võib lisada mudelisse kallutatust kuid on aktsepteeritav imputeerimise määr lähtudes parimast praktikast (Eekhout jt., 2014). Veel filtreeriti andmestikust Pearsoni korrelatsioonikordaja abil kõrgelt (piirmäär 0.9) korreleerunud muutujad.

Analüüsid viisin läbi kõigi puhvertssoonide ja kombinatsiooniga 100 m + 1000 m vastavalt *Land Use Regression (LUR)* meetodile (Hoek jt. 2008). Samuti ainult puhvertssooniga 1000 m ja 500 m. Erinevaid puhvertsoone ja nende kombinatsioone kasutati, et leida gruppide vahelist variatsiooni kõige suuremal määral selgitav mudel. Järgnevas peatükis esitatakse diagnoosipõhise grupi logistilise regressioonanalüüsi tulemused kõigi

puhvertsoone hõlmava andmestikuga ning *MCP* ja *Elastic Net* regressioonanalüüs ainult 500 m andmestiku kohta. Tegin seda kuna mudelite tulemused sarnanesid suurel määral ning käesolevas töös mitte esitatud mudelite tulemused osutusid statistiliselt nõrgemaks ning neis esines rohkem müra. Lisaks on varasemas teadustöodes 500 m puhvertsoon enim kasutatud raadius kodu ümbruskonna kirjeldamisel (Hasanzadeh jt., 2017) ning sama andmestiku põhjal läbiviidud magistritöös kasutati samuti läbivalt 500 m puhvertsooni andmeid (Aganits, 2024).

Tulemused

Tunnuste kirjeldus

Analüüsidessse kaasatud tunnuseid, mis on kategoorilised, kirjeldab tabel Lisas A.1. Lisa A.2 kirjeldab analüüsidessse kaasatud muutujaid, mis on vähemalt järjestusskaalal.

Logistilise regressioonanalüüsi tulemused diagnoosid/mitte-diagnoositud jaotuse põhjal

Mudeli tulemused (tabel Lisas B) näitavad, et ülekaalulisuse tõenäosust mõjutavad peamiselt individuaalsed omadused (vanus ja sugu). Naised on meestega võrreldes suurema riskiga ($OR = 1.26$, $p = 0.001$) ning vanuse kasvuga suureneb ülekaalulisuse diagnoosi tõenäosus märkimisväärselt, saavutades maksimumi vanuserühmas 60–69 ($OR = 2.98$, $p < 0.001$). Asutusüksuste tüübid (v.a. alevik) olid samuti seotud oluliselt kõrgema ülekaalulisuse riskiga ($OR = 2.0–3.0$). Teiste muutujate tulemused ei osutunud statistiliselt olulisteks või oli šansside suhe 1.00. Kõikide muutujate tulemused osutusid ebastabiilseteks (usaldusvahemik $0–\infty$), mis viitab mudeli üleparametriseeritusele ja võimalikele konvergenstprobleemidele. Need probleemid ilmnevad, sest on kogum muutujaid, mis kombineeritult selgitavad peaaegu täielikult variatiivsuse sõltuva muutuja tasemete vahel. Seetõttu ei saa mudeli põhjal teha järeldusi ning see on põhjuseks, miks kasutan järgmistes analüüsidess regulatsioonimeetodeid, et eemaldada muutujad, mis ei ole mudelis olulised.

Pseudo- R^2 väärtus (*Tjur* $R^2 = 0.026$) viitab sellele, et mudel selgitab ligikaudu 2.6% ülekaalulisuse varieeruvusest. Selline suurusjärg on tüüpiline mudelites, millesse on kaasatud nii keskkonna, tervise kui ka sotsio-demograafilised muutujad (Pikk (2024), depressiooni-diagnoosi, keskkonna ja sotsio-demograafiliste mudeli Nagelkerke $R^2=0.01-0.07$).

MCP regressioonanalüüsi tulemused

MCP-regressioon 500 m puhvertsooni andmetega valis välja 28 prediktorit, mis andsid mudelile kõige tugevamad tulemused (tabel Lisas C). Tulemused näitavad, et ülekaalulisuse tõenäosust mõjutavad peamiselt individuaalsed tegurid. Naised on meestega võrreldes suurema riskiga ($OR = 1.25$) ning vanuse kasvuga suureneb ülekaalulisuse tõenäosus järsult, saavutades maksimumi vanuserühmas 60–69 ($OR = 2.18$). Asulatüüpidest osutus ainukeseks oluliseks muutujaks “linn” ($OR = 1.21$), mis viitab võimalikule ruumilisele sotsiaalmajanduslikule mõjule. Keskkonnamuutujad andsid mudelis vaid väga nõrka signaali ($OR \approx 1.00$), mis viitab sellele, et detailne ruumiline kontekst ei pruugi selgita ülekaalulise diagnoosi sama hästi kui individuaalsed tegurid.

MCP-regressiooni hindamisel kasutati mitut pseudo- R^2 mõõdikut, mis annavad ülevaate mudeli selgitusvõimest. McFadden R^2 (0.046), Cox–Snell R^2 (0.018) ja Nagelkerke R^2 (0.054) viitavad sellele, et mudel selgitab ligikaudu 4–5% varieeruvusest ülekaalulisuse ja normaalkaalu grupi vahel. Kuigi need väärtused on suhteliselt väikesed, on selline suurusjärg tüüpiline tervisekäitumise ja individuaalsete ning keskkonnategurite mudelitele (Pikk, 2024), kus tulemuse kujunemist mõjutavad paljud tegurid, mis ei ole esindatud mudelis.

Väga väike alpha väärtus ($\alpha = 0.01$) näitab, et MCP mudel käitus peaaegu täielikult *ridge*-regressioonina, sisaldades siiski väikesel määral MCP'le omast muutujate eemaldamist. Mudel ei surunud koefitsiente nulli ega teinud agressiivset tunnuste valikut. Selle asemel stabiliseeris mudel koefitsiendid, vähendades nende suurust ja hoidis korreleeruvad prediktorid mudelis. Väga madal alpha on tüüpiline olukordades, kus prediktorid on omavahel tugevalt seotud ja andmestikus puudub tugev signaal. Lambda ($\lambda = 0.09947$) määrab regulatsiooni tugevuse MCP mudelis. See väärtus viitab mõõdukalt tugevale karistusele, mis vähendab koefitsientide suurust ja muudab mudeli konservatiivsemaks. Selline kombinatsioon tähendab, et mudel keskendus pigem koefitsientide silumisele kui tunnuste valikule.

Kehamassiindeksi-põhine analüüs

Järgnevad tulemused on 500 m puhvertsooni andmestiku põhjal tehtud analüüside kohta ning on saadud kasutades *Elastic Net* reguleerimismeetodit. Mudel valis välja 23 prediktorit, mis andsid parima signaali kehakaalukategooriate eristamisel. Valitud prediktorite kogum püsis stabiilne ristvalideerimise andmejaotuste lõikes. Sarnaselt diagnoosi-põhisele

analüüsile ilmnes, et ülekaalulisuse kujunemist selgitavad peamiselt individuaalsed tegurid ning keskkonnamuutujate mõju jäi pigem nõrgaks.

Kõige tugevamaks ülekaalulisuse ennustajaks osutus vanus. Vanusegrupid alates grupist 30-39 näitasid suurenenud tõenäosust kuuluda ülekaalu kategooriasse võrreldes noorimate vastajatega (vanusegrupp 20-29 $RRR=0.932$ ehk väiksem risk kuuluda ülekaaluliste gruppi võrreldes noorima vanusegrupiga ehk 15-19 aastased). Kõige kõrgem suhteline riskisuhe ilmnes vanuserühmas 60–69 aastat ($RRR = 4.293$), millele järgnesid 70+ ($RRR = 4.290$), 50–59 ($RRR = 3.276$) ning 40-49 ($RRR=1.936$). Need tulemused kattuvad nii *MCP*-regressiooni analüüsi tulemustega kui varasema teadustööga (TAI, 2025).

Järgmine oluliseks teguriks osutus sugu– naistel oli väiksem tõenäosus kuuluda ülekaaluliste hulka ($RRR = 0.681$). See on vastupidine tulemus diagnoosi-põhise analüüsiga, kuid kooskõlas TAI statistikaga (2025). Erinevus diagnoosipõhise mudeliga viitab sellele, et tervishoiusüsteemi diagnoosipraktikad ja enese-raporteeritud kaalunäitajad võivad mõjutada tulemusi erinevalt. Haridustase ennustas kehakaalukategooriatesse kuulumist mõõdukalt: kõige kõrgemasse haridustaseme gruppi kuulunud vastajatel oli väiksem tõenäosus kuuluda ülekaaluliste hulka ($RRR = 0.756$). See on kooskõlas rahvusvaheliste uuringutega, mis näitavad, et kõrgem haridus toetab tervislikumaid toitumis- ja liikumisharjumusi (Cai jt., 2013).

Nõrka kuni mõõdukat mõju omasid kaks perekondlikku tegurit — laste olemasolu ($RRR = 1.298$) ja partneriga koos elamine ($RRR = 1.69$). See tulemus toetab varasemaid leide (The & Gordon-Larsen, 2009). Neljast kolm sissetuleku kategooriat näitasid vaid nõrkasid ja ebajärjekindlaid seoseid (kuni 850 eurot $RRR 1.074$; kuni 1400 eurot $RRR = 1.177$ ja üle 1400 eurot $RRR = 1.152$).

Keskkonnamuutujad andsid mudelis väga nõrkasid signaale. Asulaüksuse tüübid linn ($RRR = 1.109$), linnaosa ($RRR = 1.040$) ja küla ($RRR = 1.037$) ning elukoha tüüp populatsiooniga üle 10 000 ($RRR = 1.034$) osutusid mudelis olulisteks teguriteks. Ehitatud keskkonna teguritest omasid nõrka mõju kolm tunnust: kioskid ($RRR = 1.061$) ja büroo- ja ärihooned ($RRR = 0.874$) ning hoonete hulk tüüp: ehitatavad hooned ($RRR = 1.024$). Looduslikku keskkonda kirjeldavatest teguritest andis signaali vegetatsiooni kõrgus ($RRR = 0.989$) ning supluskohta veekvaliteet “väga hea” ($RRR = 1.029$).

Mudel suutis eristada normaalkaalu ja ülekaalu mõõduka täpsusega ($Accuracy = 0.688$). Mudelis sensitiivsus oli madal (0.35), mis tähendab, et mudelil oli raskusi normaalkaalus indiviidide tuvastamisega, kuid spetsiifilisus kõrge (0.89), mis näitab, et mudel oli edukas ülekaalulise grupi tuvastamisel. See viitab asjaolule, et ülekaalulisuse eristamine normaalkaalu grupist on antud andmestiku põhjal keerukas.

Tulemuste interpreteerimine

MCP analüüsis luuakse palju mudelid erinevate lambda väärtuste juures ning regulatsiooni protsessi jooksul arvutatakse välja šansside suhted kõigile muutujatele nende lambda väärtuste juures. Analüüs valib lõpptulemuse ristvalideerimise abil, otsides lambda väärtust, mis annab kõige väiksema veamäära ($CV\ error$). Selle tulemusena jäävad mõned prediktorid mudelisse, kuid nende šansside suhte väärtus on üks, mis tähendab, et koefitsient oli optimaalsel lambdaväärtusel null (efekti ei ole). Need muutujad ei panusta lõppmudeli eristusvõimesse, kuid *MCP* ei eemalda neid täielikult, kuna nende koefitsiendid ei olnud null kõigis reguleerimisprotsessis loodud mudelites. Seega eristab *MCP* kahte tüüpi ebaolulisi prediktoreid: täielikult eemaldatud muutujad ja mudelisse alles jäänud, kuid nullmõjuga prediktorid. Lisaks oli muutujate eemaldamine vähem agressiivne kuna alpha väärtus oli lähedal nullile.

MCP analüüsi tulemusel saadud šansside suhted ei ole võrdselt tõlgetatavad klassikalise logistilise regressiooni sama nime kandvate väärtustega. Täpsemalt peegeldavad *MCP* šansside suhted (OR) prediktorite mudelisisest tähtsust, mitte täpset riski suurenemist populatsioonis. Ehk šansside suhe 2.18 muutujal vanusegrupp 60-69 aastat näitab, et muutujal on tugev mudelisisene mõju eristada gruppe üksteisest.

(Zhang, 2010; Miller & Breheny, 2023)

Elastic Net regressioonianalüüsi tulemusel saadud RRR ehk suhtelised riskisuhted sarnanevad üldiselt klassikalise logistilise regressiooni tulemusel saadavate šansside suhetega (OR). RRR näitab, mitu korda on šansid kuuluda uuritavasse kategooriasse suuremad või väiksemad võrreldes referentskategoriaga (magistritöös on referents “normaalkaal”), hoides kõik muud mudelis olevad tegurid konstantsena. Sarnaselt *MCP* tulemustega $RRR > 1$ viitab suurenenud riskile, $RRR < 1$ viitab vähenenud riskile ning $RRR = 1$ tähendab, et seost ei esine. (Friedman jt., 2010; Zou & Hastie, 2005, Tray jt., 2023)

Arutelu

Magistritöö eesmärk oli leida seoseid individuaalsete ja keskkonnamuutujate ning ülekaalulisuse vahel. Käesoleva töö raames valiti andmestikust järgmised tunnused, millest koosneb individuaalsete tunnuste rühm: sugu, vanus, haridus, sissetulek ja perekondliku staatuse näitajad. Keskkonda kirjeldavad muutujad jagunesid kahte kategooriasse vastavalt andmete allikale – subjektiivsed, kogutud RVTU küsimustikus, ning objektiivsed, geokodeeritud keskkonnaandmed erinevatest allikatest, ning andmed Rahvastiku- ja Äriregistrist. Töös on esitatud kaks mudelit, mis põhinevad 500 m puhvertsooni andmestikul. See lähtub varasemast teadustööst, kus kasutatakse puhvertsooni vahetu lähikeskkonna määramiseks (Hasanzadeh jt., 2017, Aganits, 2024) ning asjaolust, et teiste puhvertsooni kombinatsioonide mudelid olid statistiliselt nõrgemad.

Tulemused näitasid, et vanus oli järjepidevalt kõige tugevam ülekaalulisuse ennustaja. Nii diagnoosipõhises kui ka KMI-põhises mudelis suurenes ülekaalulisuse tõenäosus vanuse kasvades. See saavutas suurima väärtuse vanusegrupis 60-69 aastat, kus hoides teised muutujad samana, oli tõenäosus kuuluda ülekaalulisuse gruppi KMI-põhises mudelis ligi neli korda suurem võrreldes vanusegrupiga 15-19 aastat. Tõenäosus püsis kõrge vanuses 70+. Vanusegruppi 20-29 kuulumine oli aga kaitsetegur, mis tähendas umbes 12% väiksemat tõenäosust kuuluda ülekaaluliste gruppi võrreldes vanusegrupiga 15-19 aastat. See on kooskõlas varasemate uuringutega, mille järgi suureneb ülekaalulisuse levimus keskeas ning püsib kõrgena vanemaealiste seas (Rundle jt., 2020; TAI, 2025). Seda seost võib selgitada vanusega seotud kehalise aktiivsuse languse ja ainevahetuse aeglustumisega (Colleluori & Villareal, 2021).

Teine oluline tegur oli sugu. Diagnoosi-põhises mudelis olid suurema ülekaalulisuse riskiga naised. See ei lähe kokku TAI varasemate leidudega, et mehed on suuremal määral ülekaalulised (TAI, 2025). See erinevus tulemustes võib tuleneda asjaolust, et mehed kalduvad alahindama oma kuuluvust kaalukategooriasse (Haakstad jt., 2021) ning naised otsivad suurema tõenäosusega terviseteenuseid (Golinelli jt., 2025). KMI-põhises mudelis ilmnisid varasemate leidudega kooskõlas tulemused, kus naistel oli 30% väiksem tõenäosus kuuluda ülekaaluliste hulka. (TAI, 2025).

Sotsiaalmajanduslikud tegurid (haridus, sissetulek) ei osutunud tugevateks ennustajateks. Kuigi diagnoosipõhises mudelis ei olnud haridus statistiliselt oluline ennustaja,

ilmnes KMI-põhises mudelis, et haridustase eristas kehakaalukategooriaid vähesel määral: kõrgharidusega isikutel oli ligikaudu 24% väiksem tõenäosus kuuluda ülekaalulisuse rühma võrreldes inimestega kellel on põhiharidus. Need muustrid on kooskõlas rahvusvaheliste uuringutega, et kõrgema sissetulekuga riikides on madalam haridustase üks olulisemaid ülekaalulisuse riskitegureid (Cai jt., 2013; Furuki jt., 2023). Sissetuleku seos ülekaalulisusega esines vaid KMI-põhises mudelis (~7% kuni 850 eurot; ~18% kuni 1400 eurot ning ~15% rohkem kui 1400 eurot võrreldes grupiga vähem kui 450 eurot). Need seosed olid heterogeensed, mis võib peegeldada Eesti konteksti eripära, kus sissetulek ei pruugi olla seotud toitumis- ja liikumisvõimaluste/harjumustega kui pole tegemist väga väikese sissetulekuga, kus toidu kättesaadavus on märkimisväärselt piiratud.

KMI-põhine regressioon näitas, et kaalukategooriatesse kuulumisel mängivad rolli sotsiaal-demograafilised tegurid, sealhulgas perega seotud tunnused nagu laste olemasolu ja partneriga koos elamine. Indiviididel, kes vastasid küsimustikus, et elavad koos lastega, oli ligikaudu 28% suurem tõenäosus kuuluda ülekaaluliste gruppi ning partneriga koos elamine suurendas tõenäosust kuuluda ülekaaluliste gruppi 17% võrreldes indiviididega, kes ei ela koos lastega ja partneriga. See on kooskõlas varasemate leidudega, mille järgi pereeluga kaasnev ajapuudus, stress ja muutunud toitumisharjumused võivad soodustada kaalutõusu (Normal jt., 2015).

Kuigi varasemad uuringud on leidnud seoseid ehitatud keskkonna ja ülekaalulisuse vahel (Papas jt., 2007; Luo jt., 2020), jäi töös geokodeeritud keskkonnamuutujate mõju nõrgaks. Enamike objektiivsete keskkonna tunnuste šansside suhted olid *MCP*-mudelis ligikaudu üks, mis viitab praktiliselt olematule mõjule. Siiski näitab asjaolu, et muutujaid ei eemaldatud mudelist täielikult, et keskkonna tegurid on seotud ülekaalulisuse riskiga. Need muutujaid, võivad potentsiaalselt kirjeldada pigem linnalist elukeskkonda, mis läheks kokku KMI-põhise analüüsi tulemustega, kus nii populatsioon üle 10 000 kui ka asutusüksuse tüüp “linn” ja “linnaosa”, ennustasid ülekaalulisuse gruppi kuulumist väikesel määral. See tulemus on vastupidine mujal maailmas läbiviidud uuringutega, mis on leidnud, et maalises (*rural*) piirkonnas elamine on seotud suurema ülekaalulisuse riskiga (Cohen jt., 2018). Seoses suuna erinevus Eestis võib olla seotud kultuurilise kontekstiga, kus maalistes piirkondades elatakse pigem eramajades ja tehakse tihemini füüsilist tööd (aiatööd, puude lõhkumine jms) ning kasvatatakse puu-ja juurvilju. Ameerika valimi põhjal koostatud uuringus toodi välja, et maalises piirkondades on ülekaalulise määr kõrgem ning puu-ja juurviljade tarbimine

väiksem. Seega kõrgem ülekaalulise risk ja väiksem puu- ja juurviljade tarbimine on omavahel tugevalt seotud ja ülekaalulisus maalistes piirkondades rohkem levinud (Cohen jt., 2018). Pakun, et Eestis võib eksisteerida vastupidine seos värske toidu kättesaadavuse ja maal elamise vahel, kuna maalistes piirkondades pole haruldus viljapuuaed või kasvuhoone.

Siiski ilmnes KMI-põhises analüüsis vastupidiseid tulemusi, kus büroo- ja ärihoonete hulk puhvertsoonis oli nõrk kaitsetegur (u 13% väiksem tõenäosus) ning asulaüksuse tüüp “küla” väga nõrk riskitegur (~4%) ülekaalulisuse gruppi kuulumisel. Need seosed viitavad pigem linnalise piirkonna kaitsvale efektile, kuid pakun välja, et need leiud esindavad hoopis spetsiifilisemaid seoseid, mitte ei kirjelda linnalist vs maalist keskkonda. Tulemustes esines veel kaks sarnast spetsiifilist seost – nii kioskite (~6%) kui ehitavate hoonete (~2%) nõrka riskiteguri näol. Büroo- ja ärihooneid ning ehitusjärgus hooneid esineb suuremal hulgal äärelinnades ning uuselamu rajoonides. Nendes piirkondades elamine on seotud suurema autost sõltumisega, et jõuda kooli ja tööle (Prieto-Curiel & Barroso, 2025), mis omakorda viitab väiksemale füüsilise aktiivsuse tasemele igapäeva tegevuses. Statistikaameti andmete järgi on maapiirkondades (eeldatavasti ka külades) vanemate inimeste osakaal suurem (Statistikaamet, 2025) seega “küla” riskitegurina võib kirjeldada seost kehakaalu ja vanuse vahel. Geokodeeritud andmestikus ei eraldatud kiirtoidurestorane teistest restoranidest, mis võib olla põhjuseks, miks toidukohtade muutuja ei osutunud mudelites oluliseks ennustajaks (Elbel jt., (2019) leidis seose ainult kiirtoidu restoranide ja ülekaalulisuse vahel laste hulgas, teised toidukohatüübid ei ostunud statistiliselt olulisteks). Julgen aga väita, et Eesti kontekstis on kioskid sarnased kiirtoidu restoranidega, kuna tihti pakutakse neis kõrgelt töödeldud ning suure kalorisaldusega toite, seega on ka mudelis ilmnenu riskitegur nõrk.

Põhjuseid, miks käesoleva magistr töö raames läbiviidud analüüsid leidsid nõrkasid seoseid ülekaalulisuse ja keskkonna muutujate vahel on mitmeid. Üheks põhjuseks võib olla Eesti ruumiline homogeensus. Võrreldes riikidega, kus on varasemalt läbi viidud sarnaseid uuringuid, on Eestis linna ja maapiirkondade erinevused väiksemad, seda eriti just rohealade kättesaadavuse osas. Kuna varieeruvus piirkondade vahel on väiksem, siis on ka statistilised signaalid nõrgemad. Samuti ei pruugi olla koduümbrus seotud kõrgema kehalise aktiivsuse või toitumisharjumustega. Aina enam levinud on erinevad treenimisvõimalused (grupitrennid, kodused treeningud jms), mis ei ole seotud lühiajalise jalutuskäigu kaugusel olevate võimalustega. Ülekaalulisust paremini ennustava mudeli loomiseks on seega tarvis liikumisharjumusi ja toitumist otseselt kirjeldavaid tegureid. Viimaseks on varasemad

uuringud näidanud, et keskkonna mõju võib olla vahendatud stressi või tervisekäitumise kaudu (Han & Lee, 2021). Kuna käesolevas töös neid mediaatoreid ei vaadeldud, võib seos jääda varjatuks. Seetõttu ei saa järeldada käesolevas töös saadud tulemuste põhjal, et keskkond ei mängi olulist rolli ülekaalulisuse kujunemisel. On võimalik, et magistritöös kasutatud andmestik ja meetodid ei võimaldanud seda mõju selgelt tuvastada.

Rohekeskkonna seoseid ülekaalulisega ilmnas ainult KMI-põhises mudelis, kus vegetatsiooni kõrgus oli väga nõrk kaitsetegur (~1%) ning supluskohta väga hea veekvaliteet nõrk riskitegur (~3%). Vegetatsiooni nõrk, kuid statistiliselt oluline seos ülekaalulisuse riskiga käib kokku varasemate järeldustega, kus maalises piirkonnas elamine on seotud väiksema ülekaalulisuse riskiga. Veekvaliteedi seos võib tuleneda muutuja ebaühtlasest jaotusest (93% väga hea vs 7% halb) gruppide vahel, mistõttu ei tee tulemuse põhjal järeldusi.

Andmeanalüüsist tulenevaid põhjuseid, miks ei olnud võimalik leida seoseid ülekaalulisuse ja keskkonnamuutujate vahel, on mitmeid. Diagnoosipõhise ülekaalulisuse muutuja esinemissagedus oli valimis väga väike (5.1%), mis vähendas mudelite statistilist jõudu ja suurendas ebastabiilsust. See esinemissagedus on palju madalam kui TAI (2025) poolt mõõdetud ülekaalulisuse ja rasvumise esindatus populatsioonis (52.8%). See erinevus sõltub mitmest tegurist. Esiteks on võimalus, et määratakse mõni teine seotud diagnoos (näiteks hüpertensioon). Samuti ei pruugi nii kehamassiindeksi kui rasvkoe osakaalu testi järgi ülekaalulised indiviidid vajada meditsiinilisi sekkumisi, mis on diagnoosi saamise eelduseks. Samuti võib esineda mitmeid faktoreid, mis peatavad indiviidi tervishoiuteenuseid otsimast, seal hulgas sotsiaalmajanduslik staatus, uskumused, aversioon meditsiini vastu jms (Keller jt., 2024).

KMI-põhine muutuja oli jaotunud ühtlasemalt (63.1% ülekaalus ja rasvunud KMI grupist), kuid vastupidiselt diagnoosipõhise grupiga, oli ülekaalulisuse osakaal suurem kui TAI poolt raporteeritud määr (52.8%). Grupp loodi RVTU küsimustikus kogutud andmete põhjal seega tegemist on enese-raporteeritud määratlusega. Ülekaalulise ja rasvumise ülesindatus valimis võib olla tingitud indiviidide suuremast teadlikkusest (pikkuse ja kaalu osas), kui on varasemaid kokkupuuteid kehakaalu probleemidega.

Töö piirangud ja võimalikud uurimissuunad

Tegemist on kõrge dimensioonilise andmestikuga, kus esines multikollineaarsust ja üleparameetrisust. Selle tulemusena ei töötanud klassikalised logistilise regressiooni meetod (CI 0-∞) ning kõrge regulatsiooniga meetodite (MCP) tulemusena saadi väga väikesed koefitsendid. Siiski ruumiliste andmete suurem filtreerimine (näiteks hoonete varemete muutuja eemaldamise kaudu) ei olnud piisavalt põhjendatav tuginedes varasema teadustöö vähesusest Eesti valimi põhjal. Masinõppe meetodid võiksid potentsiaalselt edukalt luua kõrge ennustusvõimega mudeleid tuginedes kõrge dimensioonilisel andmestikul (Scheinker jt., 2019).

Teiseks ei osutunud diagnooside olemasolu esinduslikuks hindamaks tegelikku ülekaalulisuse esindatust. Põhinedes TAI poolt raporteeritud ülekaalulisuse määrale võib järeldada, et suur osa mitte-ülekaalulisuse gruppi langenud indiviididest on ülekaalulised teiste mõõtmisvahendite järgi. Seega on tulemuste tõlgendatavus suurel määral piiratud ning sobib pigem selgitamaks, mis tegurid mängivad rolli diagnoosi saamisel. Lisaks diagnoosi-põhilisele klassifitseerimisele on KMI-põhine klassifitseerimine problemaatiline. Seda esiteks, kuna tegemist on enese-raporteringuga, mis on tihti kallutatud. Lisaks on kritiseeritud KMI-t ülekaalulisuse hindamisel, kuna see ei võta arvesse indiviidi lihasmassi; rasvkoe ladestumist erinevatesse kehaosadesse; rassilist ja etnilist kuulumust ning soolisi erinevusi ülekaalulisuse välja arenemisel (Willimas, 2024). Töö puudujääke saaks vähendada, luues andmestikke mille algeesmärk on võimalikult täpselt ning multifaktoriaalselt selgitada indiviidi tervisekäitumist. Kasulik oleks näiteks mõõta inimese tegelikku liikumistrajektoori ning selle ümbruskonda. Lisaks tuleks kasutada gruppe paremini eristavaid ja objektiivsemaid klassifitseerimismeetodeid nagu näiteks DXA-uuring (luutiheduse ja kehakoostise määramise uuring) (Tu, 2023).

Mõlema mudeli võime kirjeldada gruppide vahelist varieeruvust oli väike. See kinnitab asjaolu, et ülekaalulisus on multifaktoriaalne nähtus, mille edukaks ennustamiseks on vaja palju erinevaid muutujaid. Selleks, et hinnata ülekaalulisuse ja keskkonna vahelisi seoseid täpsemalt, tuleks mudelisse lisada erinevaid mediaator tunnused nagu stress, tervishoiuteenuste kättesaadavus jms.

Viimaseks, kuigi 500 m puhvertsooni andmestiku valik on põhjendatav eelnevalt väljatoodud põhjustel, ei eksisteeri üksmeelt parima kirjeldatavuse saavutamiseks.

Töö tugevused

Töö üheks olulisimaks tugevuseks on valimi maht ning esindatavus. Täpsemalt on valim nii maakondade kui ka vanusegruppide osas ühtlaselt jaotunud, mis tähendab, et valim on Eesti populatsiooni esindav ning sellest võib järeldada, et tulemused on üldistatavad kogu populatsioonile. Sellist sorti andmestik, mis hõlmab endas nii subjektiivseid kui ka objektiivseid keskkonna andmed, tervise- kui ka individuaalseid andmed, niivõrd suure valimi kohta on pigem vähe levinud ning seega unikaalne võimalus avastada seoseid Eestis.

Andmeanalüüsi tugevused seisnevad modernsete regulatsioonimeetodite ning kahe erineva klassifitseerimismeetodi kasutamises. Sobivate regulatsioonimeetodite kasutamine aitab kasutada kõrge dimensioonilisi andmestikke, millega esinevad probleemid klassikalise logistilise regressioonianalüüsi käigus, sealhulgas konvergeerimisprobleemid ja üle-sobitumine. Kahe erineva klassifitseerimismeetodi tulemusel saadud sarnased leiud kinnitavad, et tegurid on olulised ülekaalulisuse ennustamisel. Üksteisest erinevad tulemused illustreerivad aga klassifitseerimismeetodite erinevusi ja kallutusi.

Kokkuvõte

Käesoleva magistritöö eesmärk oli uurida, millised individuaalsed ja keskkonna tegurid on seotud ülekaalulisuse riskiga. Analüüsis kasutati RVTU andmeid. Ülekaalulisust klassifitseeriti kahel viisil: diagnoosipõhiselt (RHK-10 koodid E65–E68) ning enese-raporteeritud KMI alusel. Andmeanalüüsis kasutati logistilist regressiooni ning selle regulatsioonimeetodeid *MCP* ja *Elastic Net*.

Tulemused näitasid, et ülekaalulisuse kõige tugevamad ennustajad olid individuaalsed tegurid. Nii diagnoosi- kui KMI-põhises mudelis oli vanus järjekindlalt kõige olulisem tegur: risk suurenes vanuse kasvades ning saavutas maksimumi vanuserühmas 60–69 aastat.

Sugu osutus oluliseks, kuid selle mõju suund sõltus mudelist: diagnoosi-põhises mudelis oli suurem riski naistel, samas kui KMI-põhises mudelis oli naistel väiksem tõenäosus kuuluda ülekaaluliste hulka. Haridustase mõjutas kehakaalukategooriatesse kuulumist mõõdukalt – kõrgharidusega vastajatel oli väiksem risk. Sissetuleku mõju oli nõrk ja ebajärjekindel.

Perekondlikel teguritel oli mõõdukas mõju vaid KMI-põhises mudelis: laste ja partneriga koos elamine suurendas tõenäosust kuuluda ülekaaluliste gruppi.

Keskkonnamuutujate mõju osutus väga nõrgaks. Nii subjektiivsed kui geokodeeritud keskkonnatunnused andsid mudelites minimaalseid signaale, kuid potentsiaalselt kirjeldasid need linnalise elukoha tüübi positiivset mõju ülekaalulisuse riskile. Lisaks esines KMI-põhises mudelis neli spetsiifilist seost kirjeldavat muutujat: riskiteguriteks osutusid küla, ehitatavad hooned ja kiosk ning kaitseteguriks büroo- ja ärihooned. Need seosed ei ühti linnalise ja maalise piirkonna järeldustega ning sellest järeldasin, et tegemist on spetsiifiliste mõjutajatega.

Kokkuvõttes saab magistritöö tulemustest järeldada, et Eesti populatsioonis selgitavad ülekaalulisuse riski peamiselt individuaalsed tegurid, samas kui keskkonnamõjud jäävad kasutatud andmestiku põhjal nõrgaks. Tulemused rõhutavad vajadust täpsemate andmete järele, et mõista ülekaalulisuse seost keskkonnaga.

Tehisaru kasutus

Magistritöös kasutati inspiratsiooniallikana, statistilistest meetodite mõistmiseks, sünonüümide leidmiseks, programmeerimis abiliseks ning tõlkimiseks Microsoft Copilot tehisintellekti rakendust. Asjakohaste allikate otsingul aitas Consensus AI tehisaru rakendus.

Tänuõnad

Soovin tänada juhendajat Kairi Kreegipuud toetuse ja nõu eest. Lisaks soovin tänada Sergio Vela Morenot toetust andmast andmeanalüüsi käigus.

Eesti Rahvastiku Vaimse Tervise Uuringut rahastas SA Eesti Teadusagentuuri Euroopa Regionaalarengu Fondist toetava RITA tegevuse 1 "Strateegilise TA tegevuse toetamine" raames. Uuringu viisid läbi Tervise Arengu Instituut ja Tartu Ülikool. Lisaks toetas uuringu läbi viimist Euroopa Liidu Horisont 2020 teadus- ja innovatsiooniprogrammi grandileping 957751 projectile "eMotional Cities: Mapping the cities through the senses of those who make them".

Kasutatud kirjandus

- Aganits, L., (2024). Geograafilise ruumi sotsiaalse konteksti andmete seos depressiooni- ja ärevuse riskiga. [Magistritöö, Tartu Ülikool]. Tartu Ülikool digiarhiiv ADA
- Arbel, Y., Fialkoff, C., & Kerner, A. (2018). Does the housing unit's type and size affect health? *Business and Economic Research*, 9(1), 45. <https://doi.org/10.5296/ber.v9i1.14125>
- Bag, S., Gupta, K., & Deb, S. (2022). A review and recommendations on variable selection methods in regression models for binary data. <https://doi.org/10.48550/arXiv.2201.06063>
- Breiman, L. Random Forests. *Machine Learning* 45, 5–32 (2001). <https://doi.org/10.1023/A:1010933404324>
- Böckerman, P., Viinikainen, J., Pulkki-Råback, L., Hakulinen, C., Pitkänen, N., Lehtimäki, T., Pehkonen, J., & Raitakari, O. T. (2017). Does Higher Education protect against obesity? evidence using Mendelian randomization. *Preventive Medicine*, 101, 195–198. <https://doi.org/10.1016/j.ypmed.2017.06.015>
- Cai, L., He, J., Song, Y., Zhao, K., & Cui, W. (2013). Association of obesity with socio-economic factors and obesity-related chronic diseases in rural Southwest China. *Public Health*, 127(3), 247–251. <https://doi.org/10.1016/j.puhe.2012.12.027>
- Claassen, M. A., Klein, O., Bratanova, B., Claes, N., & Corneille, O. (2019). A systematic review of psychosocial explanations for the relationship between socioeconomic status and body mass index. *Appetite*, 132, 208–221. <https://doi.org/10.1016/j.appet.2018.07.017>
- Cohen, A. K., Rai, M., Rehkopf, D. H., & Abrams, B. (2013). Educational attainment and obesity: A systematic review. *Obesity Reviews*, 14(12), 989–1005. <https://doi.org/10.1111/obr.12062>
- Cohen, S. A., Greaney, M. L., & Sabik, N. J. (2018). Assessment of dietary patterns, physical activity and obesity from a national survey: Rural-urban health disparities in older adults. *PLOS ONE*, 13(12). <https://doi.org/10.1371/journal.pone.0208268>

- Colleluori, G., & Villareal, D. T. (2021). Aging, obesity, sarcopenia and the effect of diet and exercise intervention. *Experimental Gerontology*, *155*, 111561. <https://doi.org/10.1016/j.exger.2021.111561>
- Consensus. (2026). Consensus (April 2026 version) [Large language model]. <https://consensus.app>
- Crielaard, L., Dutta, P., Quax, R., Nicolaou, M., Merabet, N., Stronks, K., & Sloot, P. (2020). Social norms and obesity prevalence: From cohort to system dynamics models. *Obesity Reviews*, *21*. <https://doi.org/10.1111/obr.13044>.
- de Bont, J., Márquez, S., Fernández-Barrés, S., Warembourg, C., Koch, S., Persavento, C., Fochs, S., Pey, N., de Castro, M., Fossati, S., Nieuwenhuijsen, M., Basagaña, X., Casas, M., Duarte-Salles, T., & Vrijheid, M. (2021). Urban environment and obesity and weight-related behaviours in primary school children. *Environment International*, *155*, 106700. <https://doi.org/10.1016/j.envint.2021.106700>
- Dixon, B. N., Ugwoaba, U. A., Brockmann, A. N., & Ross, K. M. (2020). Associations between the built environment and dietary intake, physical activity, and obesity: A scoping review of reviews. *Obesity Reviews*, *22*(4). <https://doi.org/10.1111/obr.13171>
- Earth Science Data Systems, N. (2024, September 30). *Normalized difference vegetation index (NDVI)*. <https://www.earthdata.nasa.gov/topics/land-surface/normalized-difference-vegetation-index-ndvi>
- Eekhout, I., de Vet, H. C. W., Twisk, J. W. R., Brand, J. P. L., de Boer, M. R., & Heymans, M. W. (2014). Missing data in a multi-item instrument were best handled by multiple imputation at the item score level. *Journal of Clinical Epidemiology*, *67*(3), 335–342. <https://doi.org/10.1016/j.jclinepi.2013.09.009>
- Elbel, B., Tamura, K., McDermott, Z. T., Wu, E., & Schwartz, A. E. (2019). Childhood obesity and the Food Environment: A population-based sample of Public School Children in New York City. *Obesity*, *28*(1), 65–72. <https://doi.org/10.1002/oby.22663>

- Furuki, H., Sonoda, N., & Morimoto, A. (2023). PS-P08-1: Relationship between education level and obesity among healthy Japanese people. *Journal of Hypertension*, 41(Suppl 1). <https://doi.org/10.1097/01.hjh.0000917200.72247.eb>
- Friedman, J., Hastie, T., & Tibshirani, R. (2010). Regularization paths for generalized linear models via coordinate descent. *Journal of Statistical Software*, 33(1). <https://doi.org/10.18637/jss.v033.i01>
- Golinelli, D., Sanmarchi, F., Guarducci, G., Palombarini, J., Benetti, P., Rosa, S., & Lenzi, J. (2025). Gender differences in healthcare utilization across Europe: Evidence from the European Health Interview Survey. *Health Policy*, 162, 105448. <https://doi.org/10.1016/j.healthpol.2025.105448>
- Haakstad, L. A., Stensrud, T., & Gjestvang, C. (2021). Does self-perception equal the truth when judging own body weight and height? *International Journal of Environmental Research and Public Health*, 18(16), 8502. <https://doi.org/10.3390/ijerph18168502>
- Halonen, J. I., Kivimäki, M., Pentti, J., Stenholm, S., Kawachi, I., Subramanian, S. V., & Vahtera, J. (2014). Green and blue areas as predictors of overweight and obesity in an 8-year follow-up study. *Obesity*, 22(8), 1910–1917. <https://doi.org/10.1002/oby.20772>
- Han, S.-J., & Lee, S.-H. (2021). Nontraditional risk factors for Obesity in modern society. *Journal of Obesity & Metabolic Syndrome*, 30(2), 93–103. <https://doi.org/10.7570/jomes21004>
- Hasanzadeh, K., Broberg, A., & Kyttä, M. (2017). Where is my neighborhood? A dynamic individual-based definition of home ranges and implementation of multiple evaluation criteria. *Applied Geography*, 84, 1–10. <https://doi.org/10.1016/j.apgeog.2017.04.006>
- Hoek, G., Beelen, R., de Hoogh, K., Vienneau, D., Gulliver, J., Fischer, P., & Briggs, D. (2008). A review of land-use regression models to assess spatial variation of Outdoor Air Pollution. *Atmospheric Environment*, 42(33), 7561–7578. <https://doi.org/10.1016/j.atmosenv.2008.05.057>
- Hosseini, Z., Veenstra, G., Khan, N. A., & Conklin, A. I. (2020). Associations between social connections, their interactions, and obesity differ by gender: A population-based,

- cross-sectional analysis of the Canadian Longitudinal Study on Aging. *PLOS ONE*, *15*(7). <https://doi.org/10.1371/journal.pone.0235977>
- Huang, C., Li, C., Zhao, F., Zhu, J., Wang, S., & Sun, G. (2022). The association between childhood exposure to ambient air pollution and obesity: A systematic review and meta-analysis. *International Journal of Environmental Research and Public Health*, *19*(8), 4491. <https://doi.org/10.3390/ijerph19084491>
- James, P., Berrigan, D., Hart, J. E., Aaron Hipp, J., Hoehner, C. M., Kerr, J., Major, J. M., Oka, M., & Laden, F. (2014). Effects of puhver size and shape on associations between the Built Environment and Energy Balance. *Health & Place*, *27*, 162–170. <https://doi.org/10.1016/j.healthplace.2014.02.003>
- Keller, A. S., Hurt, R. T., Schroeder, D. R., DeJesus, R. S., Ganesh, R., Fokken, S. C., Mundi, M. S., Bonnes, S. L., Lawson, D. K., Njeru, J. W., Vincent, A., Ebbert, J. O., Ghosh, K., Abu Lebdeh, H. S., Hensrud, D. D., Nanda, S., & Croghan, I. T. (2024). Perceived health concerns and barriers to care in persons with overweight and obesity: A patient survey. *Journal of Primary Care & Community Health*, *15*. <https://doi.org/10.1177/21501319241303608>
- Lee, X. Y., Yusof, N. W., Pillai, N. K., Yap, C. G., & Jahan, N. K. (2022). Review Paper: Association between Social Network and obesity among adult population. *Open Journal of Endocrine and Metabolic Diseases*, *12*(02), 20–46. <https://doi.org/10.4236/ojemd.2022.122003>
- Luo Y-N, Huang W-Z, Liu X-X, jt. Greenspace with overweight and obesity: A systematic review and meta-analysis of epidemiological studies up to 2020. *Obesity Reviews*. 2020; *21*:e13078. <https://doi.org/10.1111/obr.13078>
- Luo, C., Wei, T., Jiang, W., Yang, Y., Zhang, M.-X., Xiong, C.-L., & Tung, T.-H. (2024). The association between Air Pollution and obesity: An umbrella review of meta-analyses and systematic reviews. *BMC Public Health*, *24*(1). <https://doi.org/10.1186/s12889-024-19370-4>

- Mahmoud, R., Kimonis, V., & Butler, M. G. (2022). Genetics of obesity in humans: A clinical review. *International Journal of Molecular Sciences*, 23(19), 11005. <https://doi.org/10.3390/ijms231911005>
- Mowbray, F. I., Fox-Wasylyshyn, S. M., & El-Masri, M. M. (2018). Univariate outliers: A conceptual overview for the nurse researcher. *Canadian Journal of Nursing Research*, 51(1), 31–37. <https://doi.org/10.1177/0844562118786647>
- Microsoft. (2026). Copilot (GPT-4) [Large language model]. <https://copilot.microsoft.com>
- Miller, R., & Breheny, P. (2023). Feature-specific inference for penalized regression using local false discovery rates. *Statistics in Medicine*, 42(9), 1412–1429. <https://doi.org/10.1002/sim.9678>
- Newton, S., Braithwaite, D., & Akinyemiju, T. F. (2017). Socio-economic status over the life course and obesity: Systematic Review and meta-analysis. *PLOS ONE*, 12(5). <https://doi.org/10.1371/journal.pone.0177151>
- Nicolaidis, S. (2019). Environment and Obesity. *Metabolism*, 100, 153942. <https://doi.org/10.1016/j.metabol.2019.07.006>
- Papas, M. A., Alberg, A. J., Ewing, R., Helzlsouer, K. J., Gary, T. L., & Klassen, A. C. (2007). The built environment and Obesity. *Epidemiologic Reviews*, 29(1), 129–143. <https://doi.org/10.1093/epirev/mxm009>
- Persson, Å., Pyko, A., Stucki, L., Ögren, M., Åkesson, A., Oudin, A., Tjønneland, A., Rosengren, A., Segersson, D., Rizzuto, D., Helte, E., Andersson, E. M., Aasvang, G. M., Gudjonsdottir, H., Selander, J., Christensen, J. H., Leander, K., Mattisson, K., Eneroth, K., ... Eriksson, C. (2024). Long-term exposure to transportation noise and obesity: A pooled analysis of eleven Nordic cohorts. *Environmental Epidemiology*, 8(4). <https://doi.org/10.1097/ee9.0000000000000319>
- Pikk, J., (2024). Seosed elukohalähedase roheluse ja vaimse heaolu vahel. [magistritöö, Tartu Ülikool]. Tartu Ülikooli digiarhiiv ADA

- Pineda E, Stockton J, Scholes S, Lassale C, Mindell JS. Food environment and obesity: a systematic review and meta-analysis. *BMJ Nutrition, Prevention & Health*. 2024;7:. <https://doi.org/10.1136/bmjnph-2023-000663>
- Prieto-Curiel, R., & Barroso, F. (2025). The growing dominance of cars in suburban areas. *Transportation Research Interdisciplinary Perspectives*, 32, 101559. <https://doi.org/10.1016/j.trip.2025.101559>
- Ravijuhendite Nõukoda. (2019). Ülekaalulise või rasvunud patsiendi käsitus esma-tasandil. RJ-E/34.1-2019.
- Raza, A., Pulakka, A., Hanson, L. L., Westerlund, H., & Halonen, J. I. (2022). Distance to sports facilities and low frequency of exercise and obesity: A cross-sectional study. *BMC Public Health*, 22(1). <https://doi.org/10.1186/s12889-022-14444-7>
- Rundle, A. G., Factor-Litvak, P., Suglia, S. F., Susser, E. S., Kezios, K. L., Lovasi, G. S., Cirillo, P. M., Cohn, B. A., & Link, B. G. (2020). Tracking of obesity in childhood into adulthood: Effects on body mass index and fat mass index at age 50. *Childhood Obesity*, 16(3), 226–233. <https://doi.org/10.1089/chi.2019.0185>
- Scheinker, D., Valencia, A., & Rodriguez, F. (2019). Identification of factors associated with variation in US county-level obesity prevalence rates using epidemiologic vs Machine Learning models. *JAMA Network Open*, 2(4). <https://doi.org/10.1001/jamanetworkopen.2019.2884>
- Scroope, C. (2020) *Israeli - family*. Cultural Atlas. <https://culturalatlas.sbs.com.au/israeli-culture/israeli-culture-family>
- Seifu, C. N., Fahey, P. P., & Atlantis, E. (2021). Unhealthy diet pattern mediates the disproportionate prevalence of obesity among adults with socio-economic disadvantage: An Australian representative cross-sectional study. *Nutrients*, 13(4), 1363. <https://doi.org/10.3390/nu13041363>
- Statistikaamet. (2025, May 27). *RV0282U: Rahvastik Soo, Vanuserühma Ja Elukoha JÄRGI, I. jaanuar. Eesti Statistika*. https://andmed.stat.ee/et/stat/rahvastik__rahvastikunaitajad-ja-koosseis__rahvaarv-ja-rahvastiku-koosseis/RV0282U/table/tableViewLayout2

- Tani, Y., Kondo, N., Takagi, D., Saito, M., Hikichi, H., Ojima, T., & Kondo, K. (2015). Combined effects of eating alone and living alone on unhealthy dietary behaviors, obesity and underweight in older Japanese adults: Results of the Jages. *Appetite*, *95*, 1–8. <https://doi.org/10.1016/j.appet.2015.06.005>
- Tervisearengu Instituut. (2025) Ülemaailmne rasvumise päev: probleem on Eestis suur nii täiskasvanute kui laste seas. <https://www.tai.ee/et/uudised/ulemaailmne-rasvumise-paev-probleem-eestis-suur-nii-tai-skasvanute-kui-laste-seas>.
- Tervise Arengu Instituut. (n.d.) *Ülekaal*. Tervisliku toitumise informatsioon. <https://toitumine.ee/kehakaal/ulekaal>
- The, N. S., & Gordon-Larsen, P. (2009). Entry into romantic partnership is associated with obesity. *Obesity*, *17*(7), 1441–1447. <https://doi.org/10.1038/oby.2009.97>
- Tu, L. (2023, July 5). Better Ways Than BMI to Measure Obesity. *Scientific America*. other. Vaadatud May 7, 2026,.
- Tay, J. K., Narasimhan, B., & Hastie, T. (2023). Elastic net regularization paths for all generalized linear models. *Journal of Statistical Software*, *106*(1). <https://doi.org/10.18637/jss.v106.i01>
- Wells, J. C. K., Marphatia, A. A., Cole, T. J., & McCoy, D. (2012). Associations of economic and gender inequality with global obesity prevalence: Understanding the female excess. *Social Science & Medicine*, *75*(3), 482–490. <https://doi.org/10.1016/j.socscimed.2012.03.029>
- World Health Organization. (2025). Obesity and Overweight. <https://www.who.int/news-room/fact-sheets/detail/obesity-and-overweight>
- Williams, S. (2024, November 14). BMI, aka body mass index: what the science says. *Stanford Medicine News Center*. Retrieved April 22, 2025, from <https://med.stanford.edu/news/insights/2024/11/bmi-body-mass-index-alternatives-science.html>.

- Yamamoto, R., Shinzawa, M., Yoshimura, R., Taneike, M., Nakanishi, K., Nishida, M., ... Moriyama, T. (2021). Living alone and prediction of weight gain and overweight/obesity in university students: a retrospective cohort study. *Journal of American College Health*, 71(5), 1417–1426. <https://doi.org/10.1080/07448481.2021.1927052>
- Zhang, C.-H. (2010). Nearly unbiased variable selection under minimax concave penalty. *The Annals of Statistics*, 38(2). <https://doi.org/10.1214/09-aos729>
- Zou H., Hastie T., Regularization and Variable Selection Via the Elastic Net, *Journal of the Royal Statistical Society Series B: Statistical Methodology*, Volume 67, Issue 2, April 2005, Pages 301–320, <https://doi.org/10.1111/j.1467-9868.2005.00503.x>

Lisad

Lisa A

Tabel A.1. Kategoorilisi muutujaid kirjeldav tabel

	Gruppide %line jaotus	Asümmeetria kordaja	Järskus kordaja
Sugu***		0.01	-2.00
Mees	50.35		
Naine	49.65		
Vanusegrupp***		-0.51	-1.01
15-19	11.20		
20-29	12.31		
30-39	15.42		
40-49	14.36		
50-59	14.82		
60-69	13.13		
70+	18.78		
Haridus**		0.22	-2.33
põhiharidus	20.28		
gümnaasiumi ja kutseharidus	49.92		
kõrgharidus	29.80		
Sissetulek**		0.29	-2.07
kuni 450 eurot	13.75		
451-850 eurot	40.91		
851-1400 eurot	28.19		
üle 1401 eurot	17.16		
Lapsed**		0.00	-2.75
elab koos lastega	65.20		
ei ela koos lastega	34.80		
Üksi elamine		0.00	-2.75
elab üksi	75.56		
ei ela üksi	24.80		
Paarilisega elamine**		0.00	-2.75
elab paarilisega	65.54		
ei ela paarilisega	34.46		
Koduloomad		0.00	-2.75
ei	46.96		
jah	53.04		
Elukohatüüp**		0.35	-2.33
alla 1000 elanikuga asulas või maal	44.60		
1000–10 000 elanikuga asulas	29.15		
üle 10 000 elanikuga asulas	26.25		
Elukoha alatüüp***		0.00	-2.75
maaline	43.97		
linnaline	56.05		
Asulaüksuse tüüp***		0.41	-1.81
alev	1.93		

linn	9.39		
omavalitsuse sisene linn	30.23		
linnaosa	7.12		
alevik	14.10		
küla	37.22		
Supluskoha veekvaliteet**		0.00	-2.75
halb	7.15		
väga hea	92.85		
Müra		0.00	-2.75
“ei”	52.52		
“jah”	47.48		
Lärm		0.00	-2.75
“ei”	57.65		
“jah”	42.35		
Tugev valgus		0.00	-2.75
“ei”	78.28		
“jah”	21.72		
Vähene valgus		0.00	-2.75
“ei”	68.48		
“jah”	31.52		
Õhusaaste		0.00	-2.75
“ei”	61.01		
“jah”	38.99		
Lõhnad		0.00	-2.75
“ei”	59.88		
“jah”	40.12		
Vaade		0.00	-2.75
“ei”	72.14		
“jah”	27.86		
Roheala puudumine		0.00	-2.75
“ei”	74.13		
“jah”	25.87		
Eramu		0.00	-2.75
ei	71.78		
jah	28.22		
Korter		0.00	-2.75
ei	68.21		
jah	31.79		
Rohe		0.00	-2.75
ei	55.09		
jah	44.91		

*- muutujad, mis pärast andmete filtreerimist olid sisendiks diagnoosi-põhises mudelis

** - muutujad, mis pärast andmete filtreerimist olid sisendiks KMI-põhises mudelis

***- muutujad, mis pärast andmete filtreerimist olid sisendiks mõlemas mudelis

Tabel A.2 Vähemalt järjestusskaalal olevaid muutujaid kirjeldav tabel

Muutuja	M	SD	Min	Max	Asümmeetria- kordaja	Järskus- kordaja
Populatsioon (100m)	1563.25	2598.73	1	17173	2.74	8.87
Populatsioon (500m)***	1426.82	2324.86	1	17031	2.6	7.66
Populatsioon (1000m)	1169.75	1904.37	1	11721	2.45	6.13
Turg (100m)	3.9	8.49	1	47	4.5	19.95
Turg (500m)***	6.09	11.6	1	61	3.5	11.78
Turg (1000m)	7.53	14.15	1	61	2.97	7.73
Tanklakauplus (100m)	1	0	1	1	NA	NA
Tanklakauplus (500m)***	1.08	0.28	1	3	3.56	12.41
Tanklakauplus (1000)	1.42	0.77	1	7	2.33	6.87
Söögikoht (100m)	1.18	0.55	1	4	3.27	10.68
Söögikoht (500m)***	2.14	2.25	1	22	3.5	16.64
Söögikoht (1000m)	4.06	5.31	1	49	3.6	17.3
Müügikoht (100m)	1.42	1.15	1	8	3.06	9.41
Müügikoht (500m)***	3.31	4.48	1	30	3.32	12.72
Müügikoht (1000m)	5.97	8.21	1	42	2.49	6.18
Kiosk (100m)	1.05	0.22	1	2	4.07	14.65
Kiosk (500m)***	1.56	1.04	1	12	2.8	12.08
Kiosk (1000m)	2.34	2.25	1	20	2.81	11.68
Kauplus (100m)	1.27	0.74	1	9	4.04	22.17
Kauplus (500m)***	4.01	4.28	1	58	3.22	18.98
Kauplus (1000m)	9.48	11.5	1	113	3.01	15.33
Kaubanduskeskus (100m)	1.54	1	1	6	2.38	6.49
Kaubanduskeskus (500m)***	2.87	2.8	1	19	2.45	6.33
Kaubanduskeskus (1000m)	5	5.07	1	27	1.95	3.55

Muutuja	M	SD	Min	Max	Asümmeetria- kordaja	Järskus- kordaja
Büroo- ja ärihoone (100m)	1.07	0.29	1	3	4.65	22.92
Büroo- ja ärihoone (500m)***	1.33	0.83	1	8	4.39	24.33
Büroo- ja ärihoone (1000m)	1.95	1.8	1	12	2.98	9.35
Spordiraja kaugus***	1174.94	1786.97	0.92	25216.49	3.1	14.05
Supluskoha kaugus***	12370.23	12875.72	105.39	73495.91	1.49	1.94
Perearstikeskuse kaugus***	3501.86	4668.82	0.01	75261.57	3.05	21.15
Apteegi kaugus***	3868.26	5596.28	0	75170.64	2.58	11.83
Alkoholi müügikoha kaugus***	1890.98	2677.32	0	30192	2.2	5.78
Kaugus koolist***	2155.38	2769.97	0	35121	2.42	8.94
Kaugus merest***	60855.52	59421.32	22.57	212222.6 6	0.63	-0.91
Kaugus siseveekogust***	81843.06	58966.27	30.87	237969.9 4	0.74	-0.62
Kaitseala kaugus***	1728.37	1844.3	0	47045.77	4.97	85.01
Hoonete hulk tüüp: Elu- või ühiskondlik hoone (100m)	8.29	7.74	0	62	1.5	2.34
Hoonete pindala tüüp: Elu- või ühiskondlik hoone (100m)	2395.44	2047.78	0	20371.34	1.2	3.29
Hoonete hulk tüüp: Kõrval- või tootmishoone (100m)	8.43	8.59	0	109	2.1	9.14
Hoonete pindala tüüp: Kõrval- või tootmishoone (100m)	692.71	728.98	0	15082.89	3.21	26.99
Hoonete hulk tüüp: Vundament (100m)	0.13	0.42	0	6	4.13	22.88
Hoonete pindala tüüp: Vundament (100m)	15.19	73.32	0	2266.32	10.38	170.24
Hoonete hulk tüüp: Vare (100m)	0.15	0.46	0	11	4.99	49.72

Muutuja	M	SD	Min	Max	Asümmeetria- kordaja	Järskus- kordaja
Hoonete pindala tüüp: Vare (100m)	15.94	82.38	0	4145.19	20.91	783.96
Hoonete hulk tüüp: Ehitatav hoone (100m)	0.02	0.13	0	4	10.84	167.14
Hoonete pindala tüüp: Ehitatav hoone (100m)	3.53	55.5	0	2874.27	30.8	1230.17
Hoonete hulk kokku (100m)	17.02	15.07	1	114	1.53	2.25
hoonete pindala kokku (100m)	3122.81	2257.91	0	21853.78	1.04	2.49
Hoonete hulk tüüp: Elu- või ühiskondlik hoone (500m)	129.03	134.44	0	708	1.28	1.17
Hoonete pindala tüüp: Elu- või ühiskondlik hoone (500m)	35718.67	36806.39	0	291904.35	1.2	1.62
Hoonete hulk tüüp: Kõrval- või tootmishoone (500m)***	142.62	136.22	0	848	1.24	1.24
Hoonete pindala tüüp: Kõrval- või tootmishoone (500m)***	16373.05	15252.95	0	138742.78	1.39	2.53
Hoonete hulk tüüp: Vundament (500m)***	2.24	2.47	0	31	2.69	15.02
Hoonete pindala tüüp: Vundament (500m)***	278.37	567.61	0	12887.96	12.04	228.2
Hoonete hulk tüüp: Vare (500m)***	2.12	2.37	0	40	2.3	12.63
Hoonete pindala tüüp: Vare (500m)***	351.15	743.02	0	14091.98	7.2	87.55
Hoonete hulk tüüp: Ehitatav hoone (500m)***	0.29	0.78	0	11	4.24	26.02
Hoonete pindala tüüp: Ehitatav hoone (500m)	142.7	678.62	0	15301.66	7.3	68.88
Hoonete hulk kokku (500m)	276.3	265.8	1	1503	1.2	0.99
Hoonete pindala kokku	52863.94	47132.78	146.09	304757.8	0.85	0.16

Muutuja	M	SD	Min	Max	Asümmeetria- kordaja	Järskus- kordaja
(500m)				7		
Hoonete hulk tüüp: Elu- ja ühiskondlik hoone (1000m)	385.89	415.2	1	2217	1.25	0.99
Hoonete pindala tüüp: Elu- või ühiskondlik hoone (1000m)	105570.92	119664.58	62.43	805039.6	1.45	2.51
Hoonete hulk tüüp: Kõrval- või tootmishoone (1000m)	436.31	417.4	0	2094	1.11	0.45
Hoonete pindala tüüp: Kõrval- või tootmishoone (1000m)	56181.17	51713.96	0	342603.9	1.18	1.18
Hoonete hulk tüüp: Vundament (1000m)	7.3	5.75	0	63	1.97	7.15
Hoonete pindala tüüp: Vundament (1000m)	994.68	1318.59	0	31963.13	8.24	135.76
Hoonete hulk tüüp: Vare (1000m)	6.91	5.82	0	95	2.66	20
Hoonete pindala tüüp: Vare (1000m)	1144.35	1585.37	0	16457.91	3.63	19.09
Hoonete hulk tüüp: Ehitatav hoone (1000m)	1.01	2.01	0	19	3.44	15.23
Hoonete pindala tüüp: Ehitatav hoone (1000m)	482.78	1636.17	0	25224.09	6.23	53.49
Hoonete hulk kokku (1000m)	837.42	826.15	2	3879	1.13	0.51
Hoonete pindala kokku (1000m)	164373.9	161369.84	146.09	917041.6 4	1.05	0.59
Kergliiklustee (100m)	29.32	78.89	0	962.34	3.41	14.86
Kõrvalmaantee (100m)	25.7	63.37	0	409.07	2.33	4.31
Muu tee (100m)	105.52	167.03	0	1223.24	2	4.21
Põhimaantee (100m)	4.59	28.91	0	391	6.93	52.64
Rada (100m)	103	184.52	0	1186.23	2.47	6.55
Ramp või ühendustee	0	0	0	0	NA	NA

Muutuja	M	SD	Min	Max	Asümmeetria- kordaja	Järskus- kordaja
(100m)						
Tugimaantee (100m)	11.08	42.16	0	360.01	3.85	14.15
Tänav (100m)	338.52	321.19	0	1422.31	0.42	-0.96
Teed kokku (100m)	617.92	384.49	0	2133.38	0.76	0.02
Kergliiklustee (500m)***	786.62	1081.56	0	5954.13	1.82	3.35
Kõrvalmaantee (500m)***	439.93	581.89	0	3128.4	1.12	0.35
Muu tee (500m)***	1484.06	1579.68	0	11511.74	1.49	2.82
Põhimaantee (500m)	0	0	0	0	NA	NA
Rada (500m)***	2142.5	2752.85	0	18139.28	2.11	4.5
Ramp või ühendustee (500m)	14.23	89.32	0	2233.27	11.65	178.5
Tugimaantee (500m)***	263.19	466.15	0	2709.62	1.67	2.12
Tänav (500m)	5526.22	5222.04	0	20567.49	0.38	-1.15
Teed kokku (500m)***	10822.38	6973.24	505.94	35584.44	0.7	-0.09
Kergliiklustee (1000m)	2693.92	3318.73	0	18546.76	1.67	2.73
Kõrvalmaantee (1000m)	1387.67	1405.14	0	7705.54	0.85	0.08
Muu tee (1000m)	5259.31	4110.38	0	37572.46	1.16	3.41
Põhimaantee (1000m)	624.89	1025.12	0	7606.68	1.63	2.42
Rada (1000m)	7102.4	8353.04	0	46297.71	2.05	4.19
Ramp või ühendustee (1000m)	63.8	281.98	0	4756.84	8.59	94.89
Tugimaantee (1000m)	886.17	1220.68	0	6159.23	1.28	1.06
Tänav (1000m)	16582.14	16733.1	0	65114.57	0.59	-0.94
Teed kokku (1000m)	34600.3	23509.74	1836.1 7	119338.4 5	0.83	0.01
Puittaimestik (100m)	2433.28	4636.43	0	31374.93	2.44	6.13
Puittaimestik (500m)***	134354.93	147687.1	0	777206.1 2	1.44	1.64

Muutuja	M	SD	Min	Max	Asümmeetria- kordaja	Järskus- kordaja
Puittaimestik (1000m)	697248.2	582360.25	0	3099009.09	1.03	0.6
Jõe pikkus (100m)	3.52	25.03	0	473	8.45	83.13
Jõe pikkus (500m)***	183.03	436.08	0	3080	2.48	5.61
Jõe pikkus (1000m)	784.55	1220.32	0	7277	1.44	1.3
Kaldajoone pikkus (100m)	24.4	69.01	0	838	4.24	22.71
Kaldajoone pikkus (500m)***	809.72	1073.92	0	9351	1.81	3.52
Kaldajoone pikkus (1000m)	3424.98	3170.4	0	23244	1.22	1.6
Järve pindala (100m)	42.71	744.23	0	37215.22	25.62	814.41
Järve pindala (500m)	102660.16	765786.62	0	11854950	10.86	132.35
Järve pindala (1000m)	2353058.75	13350260.03	0	116300589	7.31	55.6
Paisjärve pindala (100m)	61.23	913.62	0	29545	19.84	458.77
Paisjärve pindala (500m)	105758.83	652727.57	0	8185439	8.18	73.4
Paisjärve pindala (1000m)	1035385.21	4606777.48	0	37949756	5.78	36.52
Tehisjärve pindala (100m)	11.7	519.96	0	41033	67.34	4923.24
Tehisjärve pindala (500m)	32603.63	324845.8	0	5166040	12.6	167.39
Tehisjärve pindala (1000m)	450982.23	2735582.42	0	25351050	7.65	60.89
Vegetatsiooni kõrgus (100m)	4.6	2.6	0	21.8	1.36	2.81
Vegetatsiooni kõrgus (500m)	5.140316	2.77585	0.074177	21.22147	1.155255	1.627537
Vegetatsiooni kõrgus (1000m)	5.47	2.64	0.06	20.62	0.92	0.98
Spordiraja pikkus (100m)	0.08	0.33	0	4	4.88	28.56

Muutuja	M	SD	Min	Max	Asümmeetria- kordaja	Järskus- kordaja
Spordiraja pikkus (500m)***	2.58	3.31	0	23	1.47	2.01
Spordiraja pikkus (1000m)	7.36	8.45	0	55	1.5	2.26
Lagedad alad (100m)	6574.95	8120.49	0	31317.88	1.03	-0.24
Lagedad alad (500m)***	237873.15	210701.91	0	772377.11	0.59	-0.88
Lagedad alad (1000m)	1041723.32	773834.38	0	3069143.19	0.4	-1.01
Asfalt (100m)	4646.28	4394.19	0	30412.97	1.08	1.14
Asfalt 500 (500m)***	83973.29	84074.71	0	583464.93	1.12	0.86
Asfalt (1000m)	261631.12	281229.68	0	1874900.55	1.32	1.48
Looduslikud alad kokku(100m)	15569.93	8393.92	0	31374.96	-0.35	-1.04
Looduslikud alad kokku(500m)***	463701.56	204172.33	16292.19	784902.63	-0.15	-1.2
Looduslikud alad kokku(1000m)	1994389.08	785781.31	72206.22	3133910.19	-0.29	-1.19
Ühistranspordi peatuse arv (100m)	0.21	0.56	0	5	2.85	8.16
Ühistranspordi väljumised keskmiselt nädalas (100m)	22.2	134.59	0	4313	12.43	211.69
Ühistranspordi peatuse arv (500m)***	3.96	4.04	0	49	1.72	6.63
Ühistranspordi väljumised keskmiselt nädalas(500m)***	557.71	1630.29	0	26191	6.18	51.86
Ühistranspordi peatuse arv (1000m)	12.08	13.05	0	112	1.73	4.75
Ühistranspordi väljumised	1884.29	5420.02	0	67833	5.47	37.79

Muutuja	M	SD	Min	Max	Asümmeetria- kordaja	Järskus- kordaja
keskmiselt nädalas(1000m)						
Lähim ühistranspordi peatus (m)***	393.33	489.29	7.62	13045.59	5.61	74.56
Ettevõtete arv (100m)	4.68	12.19	0	560	22.78	887.15
Ettevõtete arv (500m)***	66.5	161.35	0	4068	10.98	186.44
Ettevõtete arv (1000m)	211.35	534.35	0	8593	8.26	91.66
Makselaekumised (100m)	64317.93	888433.64	0	49621789 .17	36.47	1590
Maksulaekumised (500m)***	1340298.28	6854719.42	0	16572530 6.2	13.07	209.31
Maksulaekumised (1000m)	4988320.66	20751956.0 9	0	35704719 7.2	9.58	109.57
Töötajate arv (100m)	19.54	137.26	0	8772	44.93	2632.63
Töötajate arv (500m)***	409.53	1325.15	0	26034	10.69	149.12
Töötajate arv (1000m)	1438.42	4324.51	0	74740	8.95	105.08

*- muutujad, mis pärast andmete filtreerimist olid sisendiks diagnoosi-põhises mudelis

** - muutujad, mis pärast andmete filtreerimist olid sisendiks KMI-põhises mudelis

***- muutujad, mis pärast andmete filtreerimist olid sisendiks mõlemas mudelis

Lisa B

Diagnoosi-põhise grupi logistilise regressioonianalüüsi tulemused

Tabel B.1

Muutuja	Šansisuhe (odds ratios; OR)	standardviga	CI	p
Intercept	0.01	0.01	0–∞	<0.001
Sugu [naine]	1.26	0.09	0–∞	0.001
Vanusgrupp 20-29	0.63	0.12	0–∞	0.019
Vanusgrupp 30-39	0.74	0.13	0–∞	0.092
Vanusgrupp 40-49	1.67	0.26	0–∞	0.001
Vanusgrupp 50-59	2.36	0.35	0–∞	<0.001
Vanusgrupp 60-69	2.98	0.43	0–∞	<0.001
Vanusgrupp 70+	1.66	0.25	0–∞	0.001
Asutüsüksusetüüp [linn]	2.89	1.17	0–∞	0.008
Asutüsüksusetüüp [omavalitsuse sisene linn]	2.33	0.88	0–∞	0.025
Asutüsüksusetüüp [linnaosa]	2.50	1.09	0–∞	0.036
Asutüsüksusetüüp [alevik]	1.94	0.75	0–∞	0.086
Asutüsüksusetüüp [küla]	2.37	0.97	0–∞	0.034
Elukohta alatüüp [maaline]	0.82	0.11	0–∞	0.136
Populatsioon (100m)	1.00	0.00	0–∞	0.699
Hoonete hulk tüüp: Kõrval- või tootmishoone (100m)	1.00	0.01	0–∞	0.971
Hoonete pindala tüüp: Kõrval- või tootmishoone (100m)	1.00	0.00	0–∞	0.553
Hoonete hulk tüüp: Vundament (100m)	0.99	0.09	0–∞	0.950
Hoonete hulk tüüp: Vare (100m)	1.05	0.08	0–∞	0.535
Hoonete pindala tüüp: Elu- või ühiskondlik hoone (100m)	1.00	0.00	0–∞	0.298
Hoonete pindala kokku (100m)	1.00	0.00	0–∞	0.721
Hoonete hulk tüüp: Elu- või ühiskondlik hoone (500m)	1.00	0.00	0–∞	0.132
Hoonete hulk tüüp Kõrval- või tootmishoone (500m)	1.01	0.02	0–∞	0.693
Hoonete hulk tüüp: Vare (500m)	1.00	0.02	0–∞	0.844
Hoonete hulk tüüp: Ehitatav hoone (500m)	0.91	0.07	0–∞	0.235
Hoonete pindala kokku (500m)	1.00	0.00	0–∞	0.914

Muutuja	Šansisuhe (odds ratios; OR)	standardviga	CI	p
Hoonete pindala tüüp: Elu- või ühiskondlik hoone (500m)	1.00	0.00	0–∞	0.173
Hoonete pindala tüüp: Kõrval- või tootmishoone (500m)	1.00	0.00	0–∞	0.496
Hoonete pindala tüüp: Vundament (500m)	1.00	0.00	0–∞	0.087
Hoonete pindala tüüp: Vare (500m)	1.00	0.00	0–∞	0.025
Hoonete hulk tüüp: Vundament (1000m)	1.00	0.01	0–∞	0.981
Hoonete hulk tüüp: Vare (1000m)	1.00	0.01	0–∞	0.773
Hoonete hulk tüüp: Ehitatav hoone (1000m)	0.98	0.03	0–∞	0.623
Hoonete pindala kokku (1000m)	1.00	0.00	0–∞	0.467
Hoonete pindala tüüp: Elu- või ühiskondlik hoone (1000m)	1.00	0.00	0–∞	0.399
Hoonete pindala tüüp: Kõrval- või tootmishoone (1000m)	1.00	0.00	0–∞	0.578
Hoonete pindala tüüp: Vundament (1000m)	1.00	0.00	0–∞	0.646
Hoonete pindala tüüp: Vare (1000m)	1.00	0.00	0–∞	0.375
Teed kokku (100m)	1.00	0.00	0–∞	0.218
Tänav (100m)	1.00	0.00	0–∞	0.709
Kergliiklustee (100m)	1.00	0.00	0–∞	0.579
Kõrvalmaantee (100M)	1.00	0.00	0–∞	0.398
Muu tee (100m)	1.00	0.00	0–∞	0.963
Rada (100m)	1.00	0.00	0–∞	0.106
Teed kokku (500m)	1.00	0.00	0–∞	0.917
Tänav (500m)	1.00	0.00	0–∞	0.700
Kergliiklustee (500m)	1.00	0.00	0–∞	0.536
Kõrvalmaantee (500m)	1.00	0.00	0–∞	0.653
Muu tee (500m)	1.00	0.00	0–∞	0.216
Rada (500m)	1.00	0.00	0–∞	0.664
Tugimaantee (500m)	1.00	0.00	0–∞	0.329
Teed kokku (1000m)	1.00	0.00	0–∞	0.505
Tänav (1000m)	1.00	0.00	0–∞	0.003
Kergliiklustee (1000m)	1.00	0.00	0–∞	0.641
Kõrvalmaantee (1000m)	1.00	0.00	0–∞	0.336
Muu tee (1000m)	1.00	0.00	0–∞	0.848
Põhimaantee (1000m)	1.00	0.00	0–∞	0.003
Rada (1000m)	1.00	0.00	0–∞	0.689

Muutuja	Šansisuhe (odds ratios; OR)	standardviga	CI	p
Tugimaantee (1000m)	1.00	0.00	0–∞	0.750
Vegetatsiooni kõrgus (1000m)	0.98	0.03	0–∞	0.447
Vegetatsiooni kõrgus (500m)	1.01	0.03	0–∞	0.683
Vegetatsiooni kõrgus (100m)	1.02	0.02	0–∞	0.338
Puittaimestik (100m)	1.00	0.00	0–∞	0.279
Puittaimestik (500m)	1.00	0.00	0–∞	0.583
Puittaimestik (1000m)	1.00	0.00	0–∞	0.826
Lagedad alad (100m)	1.00	0.00	0–∞	0.445
Lagedad alad (500m)	1.00	0.00	0–∞	0.705
Lagedad alad (1000m)	1.00	0.00	0–∞	0.217
Looduslikud alad kokku (100m)	1.00	0.00	0–∞	0.716
Looduslikud alad kokku (500m)	1.00	0.00	0–∞	0.188
Looduslikud alad kokku (1000m)	1.00	0.00	0–∞	0.438
Asfalt (100m)	1.00	0.00	0–∞	0.153
Asfalt (500m)	1.00	0.00	0–∞	0.899
Asfalt (1000m)	1.00	0.00	0–∞	0.370
Jõgede pikkus (1000m)	1.00	0.00	0–∞	0.346
Kaldajoone pikkus (500m)	1.00	0.00	0–∞	0.242
Kaldajoone pikkus (1000m)	1.00	0.00	0–∞	0.128
Kaugus merest	1.00	0.00	0–∞	0.177
Kaugus siseveekogust	1.00	0.00	0–∞	0.160
Supluskoha veekvaliteet [väga hea]	1.15	0.18	0–∞	0.381
Supluskoha kaugus	1.00	0.00	0–∞	0.580
Kaitseala kaugus	1.00	0.00	0–∞	0.604
Maksulaekumised (100m)	1.00	0.00	0–∞	0.511
Maksulaekumised (500m)	1.00	0.00	0–∞	0.088
Maksulaekumised (1000m)	1.00	0.00	0–∞	0.737
Töötajate arv (100m)	1.00	0.00	0–∞	0.147
Töötajate arv(500m)	1.00	0.00	0–∞	0.266
Ettevõtete arv (1000m)	0.98	0.01	0–∞	0.038
Kaugus koolist	1.00	0.00	0–∞	0.710
Kaugus alkoholi müügikohast	1.00	0.00	0–∞	0.845
Lähim ühistranspordi peatuse kaugus	1.00	0.00	0–∞	0.574
Ühistranspordi peatuste arv (100m)	1.09	0.08	0–∞	0.202
Ühistranspordi peatuste arv (500m)	0.98	0.02	0–∞	0.393
Ühistranspordi peatuste arv (1000m)	1.00	0.01	0–∞	0.796
Ühistranspordi väljumised keskmiselt nädalas(500m)	1.00	0.00	0–∞	0.148
Spordiraja pikkus (100m)	1.02	0.11	0–∞	0.844

Muutuja	Šansisuhe (odds ratios; OR)	standardviga	CI	p
Spordiraja pikkus (500m)	1.03	0.02	0–∞	0.135
Spordiraja pikkus (1000m)	0.99	0.01	0–∞	0.478
Spordiraja kaugus	1.00	0.00	0–∞	0.381
Perearstikeskuse kaugus	1.00	0.00	0–∞	0.059
Apteegi kaugus	1.00	0.00	0–∞	0.262
Kauplus (100m)	0.92	0.09	0–∞	0.399
Kauplus (500m)	0.99	0.02	0–∞	0.708
Söögikoht (1000m)	1.02	0.06	0–∞	0.729
Kaubanduskeskus (500m)	1.02	0.03	0–∞	0.599
Müügikoht (500m)	1.00	0.02	0–∞	0.890
Tanklakauplus (500m)	0.99	0.11	0–∞	0.899
Kiosk (500m)	0.94	0.08	0–∞	0.477
Büroo- ja ärihoone (500m)	1.00	0.12	0–∞	0.979
Kauplus (1000m)	1.00	0.01	0–∞	0.998
Söögikoht (1000m)	0.99	0.03	0–∞	0.847
Kaubanduskeskus (1000m)	1.00	0.02	0–∞	0.842
Müügikoht (1000m)	1.02	0.01	0–∞	0.228
Tanklakauplus (1000m)	1.15	0.07	0–∞	0.034
Turg (1000m)	1.01	0.01	0–∞	0.371
Kiosk (1000m)	1.08	0.05	0–∞	0.101
Büroo- ja ärihoone (1000m)	0.95	0.07	0–∞	0.482

Lisa CKokkuvõte *Minimax Convecs Penalty* regressioonanalüüsi diagnoosipõhisele grupile

Tabel C.1

Muutuja	Koefitsient	Šansisuhe (odds ratio)
(Intercept)	-3.38	0.03
Vanusgrupp 20-29	-0.66	0.52
Vanusegrupp 30-39	-0.51	0.60
Vanusgrupp 40-49	0.23	1.26
Vanusgrupp 50-59	0.55	1.73
Vanusgrupp 60-69	0.78	2.18
Vanusgrupp 70+	0.24	1.27
Sugu [naine]	0.22	1.25
Asutusüksusetüüp [linn]	0.19	1.21
Asutusüksusetüüp [linnaosa]	0.06	1.06
Spordiradade pikkus (500m)	0.02	1.02
Müügikohtade arv (500m)	0.02	1.02
Hoonete hulk tüüp: ehitatav hoone (500m)	-0.09	0.91
Hoonete hulk tüüp: vare (500m)	0.01	1.01
Hoonete hulk tüüp: Kõrval- või tootmishoone (500m)	0.00	1.00
Hoonete pindala tüüp: Kõrval- või tootmishoone (500m)	0.00	1.00
Kergliiklusteed (500m)	0.00	1.00
Kõrvalmaantee (500m)	0.00	1.00
Rada (500m)	0.00	1.00
Asfalt (500m)	0.00	1.00
Kaugus siseveekogust	0.00	1.00
Supluskoha kaugus	0.00	1.00
Kaitseala kaugus	0.00	1.00
Maksulaekumised (500m)	0.00	1.00
Väljumised bussipeatusest keskmiselt nädalas (500m)	0.00	1.00
Spordiraja kaugus	0.00	1.00
Perearstikeskuse kaugus	0.00	1.00
Apteegi kaugus	0.00	1.00

Lisa D*Elastic Net* regressioonanalüüsi tulemused KMI-põhisele grupile

Tabel D.1 Mudeli kokkuvõte

Accuracy	95% CI	No Information Rate	Mcnemar's Test P-Value
0.688	(0.6763, 0.6996)	0.6293	< 2.2e-16

Tabel D.2 Segadusmaatriks

Tegelik klass		
Ennustus	Normaalkaal	Ülekaal
Normaalkaal	784	425
Ülekaal	1483	3424

Tabel D.3

Muutuja	Koefitsient	Suhteline riskisuhe (RRR; Relative Risk Ratio)
(Intercept)	-0.486	0.615
vanusegrupp 20-29	-0.071	0.932
vanusegrupp 30-39	0.313	1.368
vanusegrupp 40-49	0.661	1.936
vanusegrupp 50-59	1.187	3.276
vanusegrupp 60-69	1.457	4.293
vanusegrupp 70+	1.456	4.290
sugu [naine]	-0.385	0.681
Haridus [kõrgharidus]	-0.280	0.756
lapsed [elab lastega]	0.261	1.298
paariline [elab paarilisega]	0.156	1.169
sissetulek [kuni 850]	0.071	1.074
sissetulek [kuni 1400]	0.163	1.177
sissetulek [üle 1400]	0.141	1.152
asutusüksusetüüp [linn]	0.104	1.109
asutusüksusetüüp [linnaosa]	0.039	1.040
asutusüksusetüüp [küla]	0.037	1.037
elukohatüüp [üle 10 000 elanikuga asula]	0.033	1.034
büroo- ja ärihoone	-0.135	0.874

kiosk	0.059	1.061
hoone hulk tüüp: ehitatav hoone	0.024	1.024
vegetatsiooni kõrgus	-0.011	0.989
supluskoha veekvaliteet [väga hea]	0.028	1.029

Lihtlitsents lõputöö reprodutseerimiseks ja üldsusele kättesaadavaks tegemiseks

Mina, Marta Mändla

1. annan Tartu Ülikoolile tasuta loa (lihtlitsentsi) minu loodud teose “Individuaalsete ja keskkonna tunnuste seos ülekaalulisuse riskiga”, mille juhendaja on Kairi Kreegipuu, pHD, reprodutseerimiseks eesmärgiga seda säilitada, sealhulgas lisada Tartu Ülikooli digitaalarhiivi kuni autoriõiguse kehtivuse lõppemiseni;
2. annan Tartu Ülikoolile loa teha punktis 1 nimetatud teos üldsusele kättesaadavaks Tartu Ülikooli veebikeskkonna, sealhulgas digitaalarhiivi kaudu Creative Commons'i litsentsiga CC BY NC ND 4.0, 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.

Marta Mändla
11.05.2026