

TARTU ÜLIKOOL
LOODUS- JA TÄPPISTEADUSTE VALDKOND
MATEMAATIKA JA STATISTIKA INSTITUUT

Laura Birgit Luitva

**Mitme olekuga mudelid depressiooni ja
kardiometaboolsete haiguste vaheliste seoste
hindamiseks TÜ Eesti Geenivaramu andmete
põhjal**

Matemaatika ja statistika õppekava

Matemaatilise statistika eriala

Magistritöö (30 EAP)

Põhjuhendaja: PhD Silva Kasela

Kaasjuhendaja: prof. Lili Milani

TARTU 2022

MITME OLEKUGA MUDELID DEPRESSIOONI JA
KARDIOMETABOOLSETE HAIGUSTE VAHELISTE SEOSTE
HINDAMISEKS TÕ EESTI GEENIVARAMU ANDMETE PÕHJAL

Magistritöö

Laura Birgit Luitva

Lühikokkuvõte

Magistritöö eesmärk on uurida, kas depressioon on riskiteguriks kardiometaboolsetele haigustele. Uuritakse ka depressiooni mõju surmale nii kardiometaboolse haiguse diagnoosiga kui ka diagnoosita indiviididel. Lisaks hinnatakse hüpertensiooni, hüperkolesteroleemia ja erinevat tüüpi antidepressantide kasutuse mõju kardiometaboolsetele haigustele. Töös rakendatakse elukestusanalüüsi meetodeid ning mitme olekuga mudeleid nagu konkureerivate riskide mudel ja haigus-surm mudel. Riske hinnatakse Coxi võrdeliste riskide mudeli abil. Analüüsides on ajaskaalaks vanus ning arvestatakse paremalt tsenseeritud ning vasakult tõkestatud andmetega.

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

Märksõnad: elukestusanalüüs, mitme olekuga mudelid, depressioon, kardiometaboolsed haigused

RELATIONSHIP BETWEEN DEPRESSION AND
CARDIOMETABOLIC DISEASES USING MULTI-STATE MODELS
BASED ON ESTONIAN GENOME CENTRE'S DATA

Master's thesis

Laura Birgit Luitva

Abstract

The aim of the master's thesis is to study if depression is a risk factor for

cardiometabolic diseases. Also, the effect of depression on death for individuals both with and without cardiometabolic disease is studied. In addition, the risk of hypertension, hypercholesterolaemia and the use of different types of antidepressants for cardiometabolic diseases is assessed. Survival analysis methods and multi-state models such as the competing risk model and the illness-death model are used. The Cox proportional hazards model is used for assessing risks. In the analyses, age is used as a time-scale, taking into account that the data is right censored and left truncated.

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

Key Words: survival analysis, multi-state models, depression, cardiometabolic diseases

Sisukord

Sissejuhatus	5
1 Depressioon ja kardiometaaboolsed haigused	7
2 Analüüsimetoodika	9
2.1 Elukestusanalüüs	9
2.1.1 Tsenseeritud andmed	9
2.1.2 Ajaskaala valik	10
2.1.3 Tähtsamad definitsioonid	11
2.1.4 Kaplan-Meieri hinnang üleelamisfunktsioonile	13
2.1.5 Coxi võrdeliste riskide mudel	14
2.1.6 Coxi regressioon võrdsete sündmuste toimumise aegade korral	15
2.1.7 Võrdeliste riskide eelduse kontrollimine	16
2.1.8 Kihistatud Coxi võrdeliste riskide mudel	18
2.2 Mitme olekuga mudelid elukestusanalüüsis	19
2.2.1 Tähistused ja definitsioonid	19
2.2.2 Mittehomogeenne Markovi protsess	21
2.2.3 Aalen-Johanseni hinnang konkureerivate riskide mudeli korral	22
2.2.4 Aalen-Johanseni hinnang haigus-surm mudeli korral	23
2.2.5 Coxi regressioon mitme olekuga mudelite korral	24
3 Andmete analüüs	26
3.1 Andmete ülevaade	26
3.2 Kirjeldav analüüs	29

3.3	Naiivne elukestusanalüüs	33
3.4	Mitme olekuga mudelid	41
3.4.1	Konkureerivate riskide mudel	42
3.4.2	Haigus-surm mudel	49
	Kokkuvõte	55
	Kasutatud allikad	57
	Lisad	60

Sissejuhatus

Magistritöö eesmärk on uurida depressiooni mõju kardiometaboolsete haiguste tekkele. Täpsemalt on vaatluse all viis kardiometaboolset haigust: südame isheemiatõbi, insult, teist tüüpi diabeet, kodade virvendus ja laperdus ning rinnaangiin. Töös uuritakse depressiooni mõju ka surmale nii tervetel indiviididel kui ka kardiometaboolse haigusega indiviididel. Lisaanalüüsina hinnatakse hüpertensiooni, hüperkolesteroleemia ja erinevat tüüpi antidepressantide kasutuse mõju kardiometaboolsetele haigustele. Antidepressantide tüübid on jagatud kolmeks: selektiivsed serotoniini tagasihaarde inhibiitorid ehk SSRI (*selective serotonin reuptake inhibitors*), tritsüklilised ning ülejäänud.

Depressiooni mõju kardiometaboolsetele haigustele uuritakse kahe erineva lähenemisega. Esmalt viiakse iga haiguse korral läbi tavaline ühe sündmusega elukestusanalüüs, kus teiste kardiometaboolsete haiguste esinemist on ignoreeritud. Selline lähenemine on antud juhul naiivne, sest huvipakkuvatel kardiometaboolsetel haigustel on suuresti samad riskifaktorid ning võivad esineda koos. Sellega arvestamiseks tehakse analüüs läbi ka mitme olekuga mudeliga nagu konkureerivate riskide mudel. Sellisel juhul vaadeldakse igal indiviidil vaid esimesena saadud kardiometaboolset haigust. Lisaks võetakse konkureeriva riskina arvesse ka surma. Selleks, et hinnata depressiooni mõju surmale kardiometaboolse haiguseta ja haigusega indiviididel, kasutatakse haigus-surm mudelit, mis on samuti mitme olekuga mudel.

Töös kasutatakse Tartu Ülikooli Eesti Geenivaramu pseudonüümitud andmeid, mis on lingitud Haigekassa andmebaasiga. Analüüse tehakse kasutatakse ajaskaalana vanust ning arvestatakse, et andmed on paremalt tsenseeritud ning vasakult tõkestatud. Tunnuste mõju hinnatakse Coxi võrdeliste riskide mudelitega. Illustratsiooniks kujutatakse joonistel ühe sündmuse korral Kaplan-Meieri hinnanguid üleelamisfunktsioonile. Mitme sündmuse korral näidatakse joonised olekutõenäosuste hinnangutest, mis on leitud Aalen-Johanseni hinnangute abil. Analüüside

läbiviimiseks kasutati statistikatarkvara R.

Magistritöö esimeses peatükis antakse ülevaade depressioonist ja kardiometaboolsetest haigustest. Teises peatükis tutvustatakse analüüsimeetodikat, kus kirjeldatakse elukestusanalüüsi meetodeid ning mitme olekuga mudeleid. Kolmandas peatükis antakse ülevaade andmetest, viiakse läbi kirjeldav analüüs ning seejärel näidatakse analüüside tulemusi.

Autor tänab juhendajaid Silva Kaselat ja Lili Milanit rohkete nõuannete eest.

1 Depressioon ja kardimetaboolsed haigused

Depressiooni all kannatab 5% täiskasvanutest kogu maailmas, seejuures esineb depressiooni rohkem naistel (*Depression* 2021). Kardimetaboolsed haigused (nt kardiovaskulaarhaigused ja diabeet) on peamiseks surmapõhjuseks kogu maailmas (De Waard, A.-K. M. *et al.*, 2019). Peamised elustiiliga seotud kardimetaboolsete haiguste riskifaktorid on suitsetamine, liigne alkoholi tarbimine, füüsiline mitteaktiivsus ning ebatervislik toitumine (*Depression* 2021; *Diabetes* 2021). Riskiteguriteks on veel vanus, meessugu, ülekaal, hüpertensioon ja kõrge kolesteroolitase (Elderon, L. ja Whooley, M. A., 2013; Fletcher, B., Gulanick, M. ja Lamendola, C., 2002).

Kardimetaboolsete haiguste suurema haigestumusega ja suremusega on seostatud ka depressiooni (Hare, D. L. *et al.*, 2014; Roy, T. ja Lloyd, C. E., 2012). Seejuures on leitud, et depressiooni ja kardimetaboolsete haiguste vaheline seos on kahe-suunaline: kardimetaboolsete haiguste patsientide seas on rohkem depressiooniga indiviide ning ka vastupidi (Hare, D. L. *et al.*, 2014; Roy, T. ja Lloyd, C. E., 2012). Osaliselt võib depressiooni mõju kardimetaboolsetele haigustele põhjendada asjaolu, et depressiooniga kaasnevad tihti ebatervislikud eluviisid nagu suitsetamine, liigne alkoholi tarbimine, füüsiline mitteaktiivsus ja ebatervislik toitumine. (Penninx, 2017). Riskifaktoreid arvesse võttes on uuringute tulemused viidanud, et depressioon on ka sõltumatuks riskiteguriks kardimetaboolsetele haigustele (Elderon, L. ja Whooley, M. A., 2013; Knol, M. J. *et al.*, 2006).

Uuringud näitavad, et leidub seos ka antidepressantide kasutuse ja kardimetaboolsete haiguste vahel (Dimoula, A. *et al.*, 2021; Barnard, K., Peveler, R. C. ja Holt, R. I. G., 2013). Antidepressantide ja kardiovaskulaarsete haiguste vaheliste seoste leidmisel on uuringuid läbi viidud peamiselt SSRI tüüpi antidepressantidega, kus tulemused on valdavalt näidanud, et SSRI antidepressandid vähendavad kardiovaskulaarhaiguste riski (Dimoula, A. *et al.*, 2021). Teiste antidepressantide klassidega on uuringuid tehtud vähem ning olemasolevad tulemused näitavad pigem kahjulikku mõju (Dimoula, A. *et al.*, 2021). Uuringud teist tüüpi diabeedi ja

antidepressantide kohta viitavad ka nende vahelisele seosele, seejuures on osa antidepressantidega leitud teist tüüpi diabeedi riski suurendav mõju (Barnard, K., Peveler, R. C. ja Holt, R. I. G., 2013). Selleks, et tulemuste põhjal luua kliiniliseks kasutuseks mõeldud soovitusi, on vaja teha veel rohkem uuringuid.

Siinses töös keskendutakse kardiometaboolsete haiguste seast järgmistele haigustele (koos kasutatud RHK-10 koodidega):

- südame isheemiatõbi (I21-I25);
- peajuinfarkt ehk insult (I63 v.a I63.6);
- insuliinisõltumatu suhkurtõbi ehk teist tüüpi diabeet (E11);
- kodade virvendus ja laperdus (I48);
- rinnaangiin (I20).

Depressiooni diagnoosina kasutatakse koodidele F32, F33, F41.2 vastavaid diagnoose. Antidepressantide klassid, millele töös keskendutakse on koos vastavate ATC koodidega järgmised: SSRI (N06AB), tritsüklilised (N06AA) ja ülejäänud antidepressandid (N06AF, N06AG ja N06AX).

2 Analüüsimetoodika

Magistritöös kasutatakse andmete analüüsimiseks elukestusanalüüsi meetodeid ja mitme olekuga mudeleid. Juhul, kui pole märgitud teisiti, põhineb elukestusanalüüsi kirjeldav alampeatükk 2.1 David Colleti raamatul „Modelling Survival Data in Medical Research“ (Collett, 2015) ning mitme olekuga mudeleid tutvustav alampeatükk 2.2 autorite Luís Meira-Machado, Jacobo de Uña-Álvarez, Carmen Cardarso-Suárez ja Per Kragh Andersen kirjutatud artiklil „Multi-state models for the analysis of time to event data“ (Meira-Machado, L. *et al.*, 2009). Alampeatükkides on välja toodud praktilised näited elukestusanalüüsi rakendamise kohta statistikatarkvara R paketti *survival* kasutades.

2.1 Elukestusanalüüs

Elukestusanalüüsi (*survival analysis*) kasutatakse, kui huvi pakub aeg kindlaks määratud ajahetkest ehk algmomendist kuni teatud sündmuse toimumiseni ehk lõppmomendini. Meditsiinilistes analüüsides on algmomendiks sageli indiviidi uurinuga liitumise hetk ning vaadeldavaks sündmuseks on näiteks surm, sümptomite teke või haiguse diagnoosimine.

Aega fikseeritud algmomendist kuni kindla lõppsündmuseni nimetatakse elukestuseks ning vastavaid andmeid kestusandmeteks. Üldjuhul ei kasutata kestusandmete analüüsimiseks standardseid statistilisi meetodeid. Üheks põhjuseks on see, et kestusandmete korral vaadeldakse ajavahemike pikkuseid, mis ei saa olla negatiivsed ning mille jaotus on tihti ebasümmeetriline ja paremalt pikema sabaga. Veel on üheks kestusandmete eripäraks tsenseeritus, mida tutvustab järgnev alampeatükk.

2.1.1 Tsenseeritud andmed

Indiviidi elukestust nimetatakse tsenseerituks, kui huvipakkuvat sündmust pole sellel indiviidil vaadeldud. Tsenseerimisega on tegemist juhul, kui indiviidil pole

uuringu lõppedes huvipakkuvat sündmust toimunud või on enne uuringu lõppemist vaatluse alt lahkunud. Antud töös on kasutatavad andmed paremalt tsenseeritud, mis tähendab, et teada on viimane ajahetk, millal sündmust ei toimunud.

Tsenseeritud kestusandmete analüüsimisel on oluliseks eelduseks, et tsenseerimisaeg ja tegelik elukestus oleksid üksteisest sõltumatud. See tähendab, et huvipakkuva sündmuse toimumine ei sõltu sellest, mis põhjustab indiviidi tsenseerimise. Seega sõltumatuse eeldusel ja paremalt tsenseeritud andmete korral teame vaid, et tsenseerimisaeg on väiksem tegelikust elukestusest ning muud informatsiooni tsenseerimine huvipakkuva sündmuse toimumise kohta ei anna. Sellist tsenseerimist nimetatakse mitteinformatiivseks tsenseerimiseks.

2.1.2 Ajaskaala valik

Elukestusanalüüsis on tähtis ajaskaala valik, kuna vale valik võib viia nihkega hinnanguteni. Üldjuhul kasutatakse elukestusanalüüsis ajaskaalana vanust või aega jälgimisaja algusest. Viimast neist kasutatakse rohkem kliinilistes uuringutes, kus jälgimisaja algus on näiteks haiguse diagnoosimine või ravi alustamine ja lõpp on näiteks surm, haiguse taasesinemine vms. Epidemioloogilistes uuringutes, kus individid on vaatluse all uuringu algusest kuni haiguse diagnoosimiseni, kasutatakse üldjuhul ajaskaalana vanust, kuna paljude haiguste esinemine on tugevalt seotud vanusega. Ka siinses magistritöös kasutatakse ajaskaalana vanust. (Thiébaud, A. C. M. ja Bénichou, J., 2004)

Ajaskaala valikust sõltub, milliseid indiviide omavahel võrreldakse. Kui ajaskaalaks on aeg jälgimisaja algusest, võrreldakse omavahel indiviide, kes on vaatluse all olnud sama kaua. Kui ajaskaalaks on vanus, siis võrreldakse igal ajamomendil indiviide, kes on sama vanad. Kindla vanuse korral aga ei pruugi olla kõik individid veel vaatluse alla jõudnud. Seega tuleb arvesse võtta, et uuringuga saavad liituda vaid individid, kes on uuringu alguses elus. Seda nimetatakse vasakult tõkestatuseks ning sellega mitte arvestamisel võib saada nihkega hinnangud. (Canchola, A. J.

et al., 2003; Klein, J. P. ja Moeschberger, M. L., 2003)

Tarkvaras R elukestusanalüüsi läbi viies paketiga *survival* saab vasakult tõkestatust arvesse võtta, kui luuakse objekt *Surv(time, time2, event)*, kus *time* on uuringuga liitumise aega ning *time2* on uuringust lahkumise aeg. Lisaks võetakse paremalt tsenseerimist arvesse argumendiga *event*, mis vastab sündmuse toimumise indikaatorile. Indikaatori väärtus on 1, kui indiviidil toimus sündmus, või 0, kui vaatlus on paremalt tsenseeritud. (Therneau, T. M. *et al.*, 2022)

2.1.3 Tähtsamad definitsioonid

Elukestusanalüüsis on tähtsad mõisted üleelamisfunktsioon ja riskifunktsioon. Tähistagu indiviidi elukestust pidev juhuslik suurus T , mille väärtused on mittenegatiivsed. Olgu juhusliku suuruse T tihedusfunktsioon $f(t)$ ja jaotusfunktsioon $F(t)$:

$$F(t) = P(T \leq t) = \int_0^t f(u)du.$$

Tõenäosust, et sündmus toimub peale ajamomenti t , nimetatakse üleelamisfunktsiooniks $S(t)$ ja see avaldub kujul:

$$S(t) = P(T > t) = 1 - F(t).$$

Üleelamisfunktsioon on mittekasvav funktsioon. Uuringu alguses ehk ajamomendil $t = 0$ ei ole ühelgi indiviidil sündmust toimunud, mistõttu kehtib $S(0) = 1$. Lisaks kehtib üleelamisfunktsiooni korral omadus $S(\infty) = \lim_{t \rightarrow \infty} S(t) = 0$. (Kleinbaum, 1996)

Riskifunktsioon $h(t)$ näitab riski, et sündmus toimub ajamomendil t tingimusel, et enne ajamomenti t ei ole sündmust toimunud. Riskifunktsioon on defineeritud järgmise valemiga:

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

Riskifunktsiooni väärtused on vahemikus $[0, \infty)$ (Kleinbaum, 1996). Kasutades tingliku tõenäosuse valemit ja tuletise definitsiooni saab näidata, et üleelamisfunktsiooni $S(t)$ ja riskifunktsiooni $h(t)$ on seotud järgnevalt:

$$h(t) = \frac{f(t)}{S(t)} = -\frac{d}{dt} \ln S(t).$$

Lisaks on epidemioloogilistes uuringutes, kus huvipakkuvaks sündmuseks on haigestumine, tihti kasutatav näitaja esmahaigestumus (*incidence rate*), mis iseloomustab uute haigusjuhtude arvu riskiaja kohta. Indiviidid on siinses töös jälgimise all kuni haigestumiseni (sündmuse toimumiseni), surmani või uuringuperioodi lõpuni. Olgu uuringuperioodi jooksul haigestunud indiviidide arv D ning vaatluse all olnud indiviidide jälgimisaegade summa Y , mida nimetatakse ka riskiajaks. Esmahaigestumuse IR hinnang avaldub kujul:

$$\hat{IR} = \frac{D}{Y}.$$

Lihtsaim lähend esmahaigestumuse 95% usaldusintervallile on järgmine:

$$\hat{IR} \pm 1,96 \cdot \frac{\hat{IR}}{\sqrt{D}} = \hat{IR} \pm 1,96 \cdot \sqrt{\frac{D}{Y^2}},$$

lähtudes, et D on juhuslik suurus Poissoni jaotusega $D \sim \text{Poisson}(IR \cdot Y)$, millest tulenevalt $SE(\hat{IR}) = \frac{\hat{IR}}{\sqrt{D}}$. (Fischer, K. *et al.*, 2019)

Tihti on mõistlik uurida ka kahe grupi esmahaigestumuste vahet (*incidence rate difference*). Olgu esimeses grupis uute haigusjuhtude arv D_0 ja riskiaeg Y_0 ning teises grupis uute haigusjuhtude arv D_1 ja riskiaeg Y_1 . Sel juhul on esmahaigestumuste vahe RD hinnang

$$\hat{RD} = \hat{IR}_1 - \hat{IR}_0 = \frac{D_1}{Y_1} - \frac{D_0}{Y_0}$$

ning selle 95% usaldusintervall kujul

$$\hat{RD} \pm 1,96 \cdot \sqrt{\frac{D_1}{Y_1^2} + \frac{D_0}{Y_0^2}}.$$

(Fischer, K. *et al.*, 2019)

2.1.4 Kaplan-Meieri hinnang üleelamisfunktsioonile

Olgu n indiviidi ja nendele vastavad sündmuse või tsenseerimise ajad t_1, \dots, t_n . Olgu r ($r \leq n$) unikaalset sündmuse toimumise aega ja tähistagu $t_{(1)} < \dots < t_{(r)}$ järjestatud sündmuse toimumise aegu. Olgu veel n_j indiviidide arv, kellel pole enne ajamomenti $t_{(j)}$ sündmust toimunud, ning d_j indiviidide arv, kellel toimub sündmus ajamomendil $t_{(j)}$, $j = 1, \dots, r$. Eeldusel, et indiviididel toimuvad sündmused üksteisest sõltumatult, on Kaplan-Meieri hinnang üleelamisfunktsioonile defineeritud järgmiselt:

$$\hat{S}(t) = \prod_{j=1}^k \left(1 - \frac{d_j}{n_j}\right),$$

kus $t_{(k)} \leq t < t_{(k+1)}$, $k = 1, \dots, r$ ning $t_{(r+1)} = \infty$. Seejuures on määratud, et $\hat{S}(t) = 1$, kui $t < t_{(1)}$.

Kaplan-Meieri hinnang on treppfunktsioon, mille väärtused kahanevad ajamomentidel, kus toimus sündmus, ja on konstantsed järjestikuste sündmuste toimumisegade vahel. Juhul, kui suurim vaadeldud elukestus (olgu selle tähis t^*) on tsenseeritud, siis ei ole $t > t^*$ korral $\hat{S}(t)$ defineeritud. Kui aga suurim vaadeldud elukestus on mittetsenseeritud ehk see on $t_{(r)}$, siis kehtib $\hat{S}(t) = 0$, kui $t \geq t_{(r)}$. Kaplan-Meieri hinnanguid saab leida R-i paketi *survival* funktsiooniga *survfit()* (Therneau, T. M. *et al.*, 2022).

2.1.5 Coxi võrdeliste riskide mudel

Olgu vaatluse all n indiviidi ja p argumenttunnust X_1, \dots, X_p . Olgu i -nda indiviidi argumenttunnuste väärtuste vektor $\mathbf{x}_i = (x_{1i}, \dots, x_{pi})'$ ja riskifunktsioon $h_i(t)$, kus $i = 1, \dots, n$. Eeldatakse, et kirjeldavate tunnuste väärtused on mõõdetud uuringu alguses ja need ei muutu ajas. Veel tähistame riskifunktsiooni indiviidil, kellel kõikide argumenttunnuste väärtused on 0, funktsiooniga $h_0(t)$, mida nimetatakse ka baasriskifunktsiooniks.

Võrdeliste riskide mudelis avaldub i -nda indiviidi riskifunktsioon kujul:

$$h_i(t) = \psi(\mathbf{x}_i)h_0(t),$$

kus $\psi(\mathbf{x}_i)$ on funktsioon, mis sõltub vektorist \mathbf{x}_i . Funktsioon $\psi(\mathbf{x}_i)$ näitab riskide suhet i -nda indiviidi ja indiviidi, kellel kõikide argumenttunnuste väärtused on 0, vahel ja see ei muutu ajas. Riskide suhe ei saa olla negatiivne, mistõttu saab võrdeliste riskide mudelis valida funktsiooniks $\psi(\mathbf{x}_i)$ eksponentfunktsiooni:

$$h_i(t) = \exp(\beta_1 x_{1i} + \dots + \beta_p x_{pi})h_0(t) = \exp(\boldsymbol{\beta}' \mathbf{x}_i)h_0(t), \quad (1)$$

kus β_1, \dots, β_p on argumenttunnustele vastavad parameetrid ja $\boldsymbol{\beta} = (\beta_1, \dots, \beta_p)'$ on parameetrite vektor. Funktsiooni $\psi(\mathbf{x}_i)$ valikuks on ka teisi võimalusi, kuid võrdeliste riskide mudel kujul (1) on elukestusanalüüsis kõige sagedamini kasutatav regressioonmudel. Coxi võrdeliste riskide mudeli korral ei hinnata võrdeliste riskide mudelis (1) baasriskifunktsiooni $h_0(t)$. Seega on tegemist poolparameetrilise mudeliga, kus hinnatakse vaid parameetreid $\boldsymbol{\beta}$, mida tehakse Coxi regressioonis suurima tõepära meetodil.

Olgu n indiviidi, kelle hulgast huvipakkuv sündmus on toimunud r indiviidil ning ülejäänud $n - r$ vaatlust on paremalt tsenseeritud. Eeldatakse, et sündmuste toimumise ajad on erinevad. Olgu järjestatud sündmuste toimumise ajad $t_{(1)} < \dots < t_{(r)}$, kus $t_{(j)}$ vastab järjekorras j -ndana toimunud sündmusele. Tähistagu $R(t_{(j)})$ indiviidi

dide hulka, kes on ajamomendil $t_{(j)}$ riskigrupis, st kellel pole enne ajamomenti $t_{(j)}$ sündmust ega tsenseerimist toimunud. Võrdeliste riskide mudeli (1) jaoks avaldas D. R. Cox tõepärafunktsiooni kujul:

$$L(\boldsymbol{\beta}) = \prod_{j=1}^r \frac{\exp(\boldsymbol{\beta}' \mathbf{x}_{(j)})}{\sum_{l \in R(t_{(j)})} \exp(\boldsymbol{\beta}' \mathbf{x}_l)}, \quad (2)$$

kus $x_{(j)}$ on argumenttunnuste väärtuste vektor indiviidil, kellel sündmus toimus ajamomendil $t_{(j)}$. Tegemist on osalise tõepärafunktsiooniga, kuna tõepärafunktsioon (2) ei sõltu täpsetest tsenseerimise ja sündmuse toimumise aegadest, vaid nende järjekorrast. Suurima tõepära hinnagud parameetritele $\boldsymbol{\beta}$ leitakse logaritmitud tõepärafunktsiooni maksimeerimisel, milleks kasutatakse numbrilisi meetodeid.

2.1.6 Coxi regressioon võrdsete sündmuste toimumise aegade korral

Coxi võrdeliste riskide mudel eeldab, et riskifunktsioon on pidev, mistõttu parameetrite $\boldsymbol{\beta}$ hindamiseks eeldatakse tõepärafunktsioonis (2), et kõik sündmuse toimumise ajad on erinevad. Praktikas on üldjuhul elukestused teada päeva, kuu või aasta täpsusega, seega võib leiduda ka võrdseid sündmuse toimumise aegu. Siinses töös kasutatavad elukestused on samuti leitud päeva täpsusega.

Selleks, et Coxi regressioonis saaks käsitleda samal ajal toimunud sündmusi, leidsid Kalbfleisch ja Prentice sobiva tõepärafunktsiooni. See on aga keerukas ja arvutuslikult aeganõudev, eriti siis, kui ühel või mitmel ajamomendil on toimunud palju sündmusi. Seega täpse tõepärafunktsiooni asemel kasutakse üldjuhul lähendeid. Enamus statistikapakette kasutab Breslow lähendit, mis on lihtne ja suhteliselt täpne, kui samal ajal pole toimunud liiga palju sündmusi. R-i pakett *survival* kasutab selle asemel vaikimisi Efroni lähendit, mis annab võrdsete elukestuste korral täpsema lähendi ning on arvutuslikult sama efektiivne (Therneau, T. M. *et al.*, 2022).

Olgu indiviidide, kellel toimus sündmus ajamomendil $t_{(j)}$, arv d_j ja olgu nende indiviidide hulk $D(t_{(j)})$, $j = 1, \dots, r$. Tähistagu vektor \mathbf{s}_j argumenttunnuste väärtuste vektorite summat nendel indiviididel, kes kuuluvad hulka $D(t_{(j)})$. Efroni lähend tõepärafunktsioonile avadub kujul:

$$\prod_{j=1}^r \frac{\exp(\boldsymbol{\beta}' \mathbf{s}_j)}{\prod_{k=1}^{d_j} [\sum_{l \in R(t_{(j)})} \exp(\boldsymbol{\beta}' \mathbf{x}_l) - (k-1) d_j^{-1} \sum_{l \in D(t_{(j)})} \exp(\boldsymbol{\beta}' \mathbf{x}_l)]}.$$

Kui sündmuste toimumisajad on erinevad, siis $d_j = 1$ iga $j = 1, \dots, r$ korral ja Efroni lähend võrdsustub osalise tõepärafunktsiooniga (2).

2.1.7 Võrdeliste riskide eelduse kontrollimine

Võrdeliste riskide mudelis eeldatakse, et riskide suhted on võrdelised ja ajas muutumatud. Üheks võimaluseks, millega eeldust kontrollida, on kasutada Schoenfeldi jääke. Üldjuhul kasutatakse selleks kaalutud Schoenfeldi jääke, mis on eelduste rikumise tuvastamisel efektiivsemad. Kaalutud Schoenfeldi jääke kasutab vaikumisi ka R-i paketi *survival* funktsioon *cox.zph()* (Therneau, T. M. *et al.*, 2022).

Olgu p argumenttunnust ja n indiviidi. Olgu vaatluste põhjal hinnatud riskifunktsiooni Coxi võrdeliste riskide mudeliga (1), mille tulemusena on saadud parameetervectori $\boldsymbol{\beta}'$ hinnang $\hat{\boldsymbol{\beta}}'$. Schoenfeldi jäägid on defineeritud nii, et igal indiviidil leitakse p jääki, mis vastavad p argumenttunnusele. Olgu sündmuse toimumist näitav indikaator δ_i , mille väärtus on 1, kui i -ndal indiviidil toimus sündmus, ja 0, kui i -nda indiviidi elukestus on tsenseeritud. Schoenfeldi jääk, mis vastab indiviidile i ja argumenttunnusele X_j , on kujul:

$$r_{S_{ji}} = \delta_i \left(x_{ji} - \frac{\sum_{l \in R(t_i)} x_{jl} \exp(\hat{\boldsymbol{\beta}}' \mathbf{x}_l)}{\sum_{l \in R(t_i)} \exp(\hat{\boldsymbol{\beta}}' \mathbf{x}_l)} \right),$$

kus, x_{ji} on j -nda argumenttunnuse väärtus indiviidil i ja $R(t_i)$ tähistab jällegi riskigrupi ajamomendil t_i . Paneme tähele, et tsenseeritud vaatlustel on Schoenfeldi

jäägid nullid, kuna tsenseerimise korral on indikaatori δ_i väärtus 0. Selle tõttu kasutatakse tihti tsenseerimisest tingitud nulliliste jääkide asemel puuduvaid väärtusi, et oleks võimalik eristada n-ö tegelikke nulle.

Schoenfeldi jääkide abil avaldatakse kaalutud Schoenfeldi jäägid. Tähistame i -nda indiviidi Schoenfeldi jääkide vektorit $\mathbf{r}_{S_i} = (r_{S_{1i}}, \dots, r_{S_{pi}})'$. Kaalutud Schoenfeldi jääkide vektor i -ndal indiviidil avaldub järgmiselt:

$$\mathbf{r}_{S_i}^* = d\text{var}(\hat{\boldsymbol{\beta}})\mathbf{r}_{S_i},$$

kus d tähistab n indiviidi hulgas toimunud sündmuste arvu ja $\text{var}(\hat{\boldsymbol{\beta}})$ on Coxi regressioonis hinnatud parameetrite $\hat{\boldsymbol{\beta}}$ kovariatsioonimaatriks. Kaalutud Schoenfeldi jääk $r_{S_{ji}}^*$, mis vastab i -ndale indiviidile ja j -ndale argumenttunnusele, on vektori $\mathbf{r}_{S_i}^*$ j -s element.

Olgu $\beta_j(t)$ argumenttunnuse X_j ajas muutuv parameeter ning $\beta_j(t_i)$ selle väärtus i -nda indiviidi sündmuse toimumise hetkel t_i . Kaalutud Schoenfeldi jääkide keskväertuse kohta kehtib järgmine seos:

$$E(r_{S_{ji}}^*) \approx \beta_j(t_i) - \hat{\beta}_j,$$

kus $\hat{\beta}_j$ on parameetri β_j hinnang. Seega j -nda argumenttunnuse ajas muutuva kordaja $\beta_j(t)$ käitumist saab näha, kui kanda joonisele väärtused $r_{S_{ji}}^* + \hat{\beta}_j$, mis on seatud vastavusse sündmuste toimumisaegadega. Kui joonisele kantud punktid asuvad horisontaalsel sirgel, siis j -nda argumenttunnuse kordaja on ajas konstantne ja võrdeliste riskide eeldus on selle argumenttunnuse korral täidetud. Lisaks graafilisele kontrollile saab *survival* paketis *cox.zph()* funktsiooniga testida, kas graafikule sobitatud sirge tõus on võrdne nulliga (Therneau, T. M. *et al.*, 2022). Kui testi tulemusena selgub, et sirge tõus ei ole 0, siis võrdeliste riskide eeldus on rikutud.

2.1.8 Kihistatud Coxi võrdeliste riskide mudel

Võib juhtuda, et Coxi võrdeliste riskide mudelis ei ole võrdeliste riskide eeldus kõikide argumenttunnuste korral täidetud. Sel juhul kasutatakse tihti kihistatud võrdeliste riskide mudelit. Kihid moodustatakse eeldusi rikkunud argumenttunnuste põhjal. Kihistatud võrdeliste riskide mudelis eeldatakse, et ülejäänud argumenttunnuste korral on võrdeliste riskide eeldus kihtide siseselt täidetud, kuid kihtidele vastavad baasriskifunktsioonid võivad olla erinevad.

Kihid vastavad eeldusi rikkunud argumenttunnuste väärtuste kõikvõimalikele kombinatsioonidele. Seejuures kihistamiseks kasutatud tunnuseid enam argumenttunnusena mudelisse ei lisata. Juhul kui eeldusi rikkunud tunnus on pidev, siis tuleks kihtide moodustamiseks defineerida see ümber faktortunnuseks, kus faktorid vastavad kindlatele väärtuste intervallidele. (Kleinbaum, D. G. ja Klein, M., 2005)

Olgu g kihti ja olgu j -ndas kihis olevate indiviidide arv n_j ning baasriskifunktsioon $h_{0j}(t)$, kus $j = 1, \dots, g$. Tähistame j -ndas kihis oleva i -nda indiviidi riskifunktsiooni $h_{ij}(t)$, kus $i = 1, \dots, n_j$. Kihistatud võrdeliste riskide mudel p argumenttunnusega avaldub kujul

$$h_{ij}(t) = \exp(\boldsymbol{\beta}' \mathbf{x}_{ij}) h_{0j}(t),$$

kus \mathbf{x}_{ij} on argumenttunnuste X_1, \dots, X_p väärtuste vektor j -ndas kihis oleval i -ndal indiviidil.

Paneme tähele, et kihtidele vastavad baasriskifunktsioonid $h_{0j}(t)$ võivad olla erinevad, kuid parameetrid $\boldsymbol{\beta}$ on samad. See tähendab, et riskide suhted on igas kihis samad. Kihistatud Coxi regressioonis leitakse parameetrite $\boldsymbol{\beta}$ hinnangud jällegi osalise tõepära maksimeerimise abil. Maksimeeritav tõepärafunktsioon on kihtidele vastavate tõepärafunktsioonide korrutis. (Kleinbaum, D. G. ja Klein, M., 2005)

2.2 Mitme olekuga mudelid elukestusanalüüsis

Mitme olekuga mudeleid (*multi-state models*) kasutakse, et analüüsida andmeid, kus huvipakkuvaid sündmusi on rohkem kui üks ning ühel indiviidil võib toimuda mitu sündmust. Eelnevas alampeatükis 2.1 tutvustatud elukestusanalüüsi meetodid on mõeldud ühe huvipakkuva sündmuse uurimiseks. Siinses alampeatükis kirjeldatakse, kuidas elukestusanalüüsi meetodeid rakendatakse mitme olekuga mudelitele.

Näiteks on mitme olekuga mudeleid mõistlik kasutada juhul, kui ühe sündmuse toimumine takistab teise sündmuse toimumist. Juhul, kui huvipakkuvaks sündmuseks on haigestumine, siis ei saa indiviidil sündmust toimuda, kui ta sureb enne haigestumist. Kui antud näites surma käsitleda tsenseerimisena, siis rikutakse tähtsat eeldust, et tsenseerimine peab olema sõltumatu tegelikust elukestusest. Eelduse rikkumine viib aga nihkega Kaplan-Meieri hinnanguteni, mistõttu tuleks siin näites vaadelda surma eraldi sündmusena. (Putter, H., Fiocco, M. ja Geskus, R. B., 2007)

Mitme olekuga mudelite korral saab indiviid liikuda lõpliku arvu seisundite ehk olekute vahel. Sellisteks olekuteks on näiteks haiguse staadiumid. Ühest seisundist teise liikumist nimetatakse üleminekuks või sündmuseks. Kõige lihtsamaks mitme olekuga mudeliks on kahe olekuga mudel, kus võimalik on vaid üks üleminek. Selline mudel on juba tuttav elukestusanalüüsist, kus nimetatud kaheks seisundiks on *terve* ja *surm* ning võimalik on minna vaid olekust *terve* olekusse *surm*. Peatükkides 2.2.3 ja 2.2.4 tutvustatakse kahte tuntud mitme olekuga mudelit: konkureerivate riskide mudel ja haigus-surm mudel.

2.2.1 Tähistused ja definitsioonid

Juhuslikuks protsessiks $\{X(t), t \in \mathcal{T}\}$ nimetatakse juhuslike suuruste pere, kus $X(t)$ on juhuslik suurus iga $t \in \mathcal{T}$ korral (Ross, 2014). Mitme olekuga mudelid vastavad juhuslikule protsessile $\{X(t), t \in \mathcal{T}\}$, kus seisundite hulk on lõplik $S =$

$\{1, \dots, N\}$ ning $\mathcal{T} = [0, \tau]$, $\tau < \infty$. Seega $X(t) = j$ tähendab, et indiviid on ajamomendil t seisundis j . Tõenäosust olla ajamomendil t seisundis j ($j \in S$) tähistatakse

$$p_j(t) = P(X(t) = j).$$

Üldjuhul võivad tuleviku sündmused sõltuda minevikust. Olgu H_{t-} protsessi ajalugu, mis sisaldab intervallis $[0, t)$ toimunud vaatluste informatsiooni nagu varasemalt läbitud seisundid ja üleminekute toimumise ajad. Mitme olekuga mudeleid kirjeldavad üleminekutõenäosused on defineeritud järgmiselt:

$$p_{ij}(s, t) = P(X(t) = j \mid X(s) = i, H_{s-}),$$

kus $i, j \in S$ on võimalikud seisundid ja $s, t \in \mathcal{T}$, $s \leq t$. Seega üleminekutõenäosus näitab tõenäosust olla ajamomendil t seisundis j tingimusel, et ajamomendil s oldi seisundis i , ja arvestades protsessi ajalooga intervallis $[0, s)$.

Seisundid saavad olla mööduvad või neelavad. Mööduva seisundiga on tegemist siis, kui sellest olekust on võimalik üleminek mõnda teise olekusse. Vastupidiselt on neelav seisund selline, kust pole võimalik ühtegi seisundisse edasi liikuda. Juhul, kui seisund j on neelav, siis kehtib $p_{jj}(s, t) = 1$ iga $s, t \in \mathcal{T}$, $s \leq t$ korral. Neelavaks seisundiks on näiteks surm.

Mitme olekuga mudeleid saab kirjeldada ka ülemineku intensiivsuste abil. Ülemineku intensiivsus olekust i olekusse j näitab hetkelist riski minna olekusse j tingimusel, et eelnevalt ollakse seisundis i , ning seejuures võttes arvesse ka protsessi minevikku. Seisundist i seisundisse j ülemineku intensiivsus ajamomendil t on defineeritud järgmiselt:

$$h_{ij}(t) = \lim_{\Delta t \rightarrow 0} \frac{p_{ij}(t, t + \Delta t)}{\Delta t}.$$

2.2.2 Mittehomogeenne Markovi protsess

Tihti eeldatakse mitme olekuga mudelite puhul, et juhusliku protsessi näol on tegemist Markovi protsessiga. Juhuslikku protsessi nimetatakse Markovi protsessiks, kui kehtib järgmine omadus:

$$P(X(t) = j \mid X(s) = i, H_{s-}) = P(X(t) = j \mid X(s) = i)$$

kus $0 \leq s < t$. Seega Markovi protsess on juhuslik protsess, mille korral protsessi tulevik ei sõltu minevikust, vaid viimasest teadaolevast seisundist ehk olevikust. Selline eeldus tehakse ka siinses töös.

Veel üheks eelduseks, mida sageli tehakse, on aja homogeensuse eeldus, mille korral ülemineku intensiivsused on ajas konstantsed. Sel juhul kehtib ülemineku tõenäosuste korral omadus $p_{ij}(s, t) = p_{ij}(0, t - s)$. Antud töös aga käsitletakse mittehomogeenseid Markovi mudeleid. Üheks võimaluseks, kuidas mittehomogeenseid Markovi protsesse mudeldada, on mitteparameetriselt. Selliseks mitteparameetriseliseks meetodiks on näiteks Kaplan-Meieri hinnangu üldistus mitme olekuga mudelitele, mille pakkusid välja Aalen ja Johansen. Aalen-Johanseni hinnanguid üleminekutõenäosustele konkureerivate riskide mudeli ja haigus-surm mudeli korral tutvustatakse vastavalt peatükkides 2.2.3 ja 2.2.4.

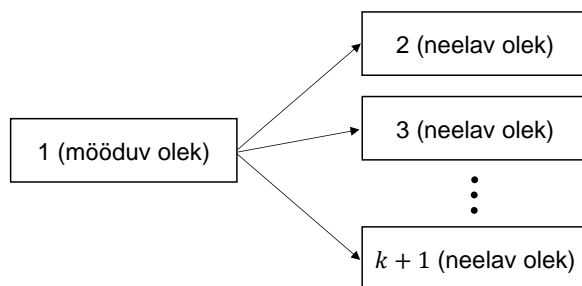
Aalen-Johanseni hinnangute abil saab leida ka hinnangud seisundis olemise tõenäosustele $p_j(t)$, mida mitme oleku mudelite puhul kujutatakse tihti graafiliselt. Lihtsamal juhul, kui andmed ei ole vasakult tõkestatud, avaldub seisundis olemise tõenäosus järgmiselt:

$$\hat{p}_j(t) = \sum_{i=1}^N \hat{p}_i(0) \hat{p}_{ij}(0, t),$$

kus $\hat{p}_i(0)$ on tõenäosus, et indiviid on algushetkel $t = 0$ seisundis i (Borgan, 2014). Tarkvara R pakettis *survival* saab hinnata seisundis olemise tõenäosusi Aalen-Johanseni hinnangute abil funktsiooniga *survfit()* ka vasakult tõkestatud andmete jaoks (Therneau, T. M. *et al.*, 2022).

2.2.3 Aalen-Johanseni hinnang konkureerivate riskide mudeli korral

Konkureerivate riskide mudel (*competing risks model*) on tuntud mitme olekuga mudel, mis sisaldab ühte mööduvat olekut ning lõplik arv neelavaid olekuid. Olgu konkureerivate riskide mudelis k neelavat olekut, siis seisundite hulk on $S = \{1, 2, \dots, k + 1\}$, kus 1 märgib mööduvat olekut ning $2, \dots, k + 1$ on neelavad olekud. Ühel indiviidil saab konkureerivate riskide mudeli korral toimuda vaid üks sündmus. Seega sündmused n-ö konkureerivad omavahel, millest tuleneb ka mudeli nimi. Konkureerivate riskide mudelit on kujutatud Joonisel 1.



Joonis 1: Konkureerivate riskide mudel.

Tihti on konkureerivate riskide mudelis mööduvaks olekuks elus olemine ning neelavateks olekuteks erinevad surmapõhjused. Antud töös konkureerivad omavahel erinevad kardiometaboolsed haigused, kuna huvi pakub vaid esimesena diagnoositud kardiometaboolne haigus. Seejuures konkureerib haigustega ka surm.

Näitame nüüd, kuidas avalduvad Aalen-Johanseni hinnangud konkureerivate riskide mudeli korral. Vaatleme konkureerivate riskide mudelit, kus on k neelavat olekut ja $S = \{1, \dots, k + 1\}$. Olgu n indiviidi ja r erinevat aega, millal toimus sündmus. Tähistame järjestatud sündmuse toimumise ajad $t_{(1)} < \dots < t_{(r)}$. Olgu n_{1j} indiviidide arv, kes olid enne ajamomenti $t_{(j)}$ seisundis 1, ning d_{1hj} indiviidide arv, kellel toimus ajamomendil $t_{(j)}$ üleminek seisundist 1 seisundisse h , $h = 2, \dots, k + 1$.

Üleminekutõenäosuse $p_{11}(s, t)$ hinnangu saab leida Kaplan-Meieri hinnangu abil:

$$\hat{p}_{11}(s, t) = \prod_{s < t_{(j)} \leq t} \left(1 - \frac{\sum_{h=2}^{k+1} d_{1hj}}{n_{1j}} \right).$$

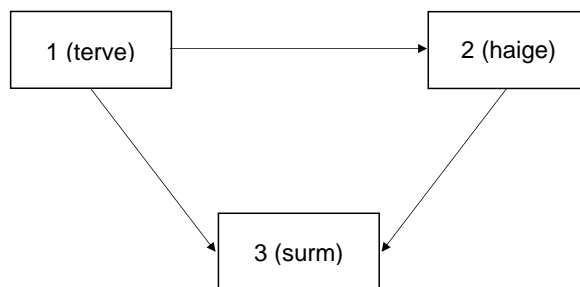
Üleminekutõenäosuste $p_{1h}(s, t)$ hinnangud saab leida Aalen-Johanseni hinnagutega, mis avalduvad kujul:

$$\hat{p}_{1h}(s, t) = \sum_{s < t_{(j)} \leq t} \hat{p}_{11}(s, t_{(j-1)}) \frac{d_{1hj}}{n_{1j}}.$$

(Borgan, 2014)

2.2.4 Aalen-Johanseni hinnang haigus-surm mudeli korral

Haigus-surm mudel (*illness-death model*) on mitme olekuga mudel, kus on kaks mööduvat seisundit *terve* ja *haige* ning üks neelav seisund *surm*. Seega vastavalt nendele olekutele on seisundite hulk $S = \{1, 2, 3\}$. Haigus-surm mudeli seisundid ja üleminekud on kujutatud Joonisel 2.



Joonis 2: Haigus-surm mudel.

Tihti on haigus-surm mudeli eesmärgiks hinnata, kas suuremise risk tervetel indiviididel on erinev haigetest. Mõnikord on võimalik ka haigusest tervenemine. Sel juhul lisatakse mudelisse veel üleminek olekust *haige* olekusse *terve*, kuid sellist

üleminekut antud töös ei käsitleta. Siinses töös on uuritavaks haiguseks esimene kardiometaaboolne haigus.

Näitame nüüd, kuidas avalduvad Aalen-Johanseni hinnangud üleminekutõenäosustele $p_{11}(s, t)$, $p_{12}(s, t)$ ja $p_{22}(s, t)$. Ülejäänud üleminekutõenäosused saab haigusurm mudelis leida nende abil. Olgu n indiviidi ja r erinevat aega, millal toimus sündmus (haigus või surm). Tähistame järjestatud sündmuse toimumise ajad $t_{(1)} < \dots < t_{(r)}$. Tähistagu n_{1j} tervete indiviidide arvu enne ajamomenti $t_{(j)}$ ning d_{12j} ja d_{13j} tervete indiviidide arvu, kes vastavalt haigestusid või surid ajamomendil $t_{(j)}$. Sarnaselt olgu n_{2j} haigete indiviidide arv enne ajamomenti $t_{(j)}$ ning d_{23j} haigete indiviidide arv, kes surid ajamomendil $t_{(j)}$. Hinnangud üleminekutõenäosustele $p_{11}(s, t)$ ja $p_{22}(s, t)$ on leitavad Kaplan-Meieri hinnangu abil:

$$\hat{p}_{11}(s, t) = \prod_{s < t_{(j)} \leq t} \left(1 - \frac{d_{12j} + d_{13j}}{n_{1j}} \right),$$

$$\hat{p}_{22}(s, t) = \prod_{s < t_{(j)} \leq t} \left(1 - \frac{d_{23j}}{n_{2j}} \right).$$

Aalen-Johanseni hinnang üleminekutõenäosusele $p_{12}(s, t)$ avaldub kujul:

$$\hat{p}_{12}(s, t) = \sum_{s < t_{(j)} \leq t} \hat{p}_{11}(s, t_{(j-1)}) \frac{d_{12j}}{n_{1j}} \hat{p}_{22}(t_{(j)}, t).$$

(Borgan, 2014)

2.2.5 Coxi regressioon mitme olekuga mudelite korral

Olgu p argumenttunnust ja tähistagu $h_{ij0}(t)$ baasintensiivsust üleminekul seisundist i seisundisse j . Seisundist i seisundisse j ülemineku intensiivsus $h_{ij}(t)$ avaldub võrdeliste riskide mudeliga kujul:

$$h_{ij}(t) = \exp(\beta'_{ij} \mathbf{x}) h_{ij0}(t), \quad (3)$$

kus β'_{ij} on parameetrite vektor üleminekul seisundist i seisundisse j , ning \mathbf{x} on argumenttunnuste väärtuste vektor. Coxi võrdeliste riskide mudeli (3) parameetrid hinnatakse jällegi osalise tõepärafunktsiooniga (2), kus kasutatakse vaid indiviide, kes on olnud seisundis i ning mudeldatavaks sündmuseks on üleminek seisundist i seisundisse j .

Paneme tähele, et ülemineku intensiivsuse h_{ij} mudeldamisel Coxi võrdeliste riskide mudeliga on parameetrite β_{ij} hindamiseks vaja osalise tõepärafunktsioonis (2) teada sündmuse j toimumiste aegadel, kes on riskigrupis ja kellel toimus sündmus j . Seega osalise tõepärafunktsioonis ei mängi rolli, kas individid on uuringust lahkunud tsenseerimise või teise sündmuse toimumise tõttu. Järelikult parameetrite hindamiseks tehakse sisuliselt tavaline ühe sündmusega Coxi võrdeliste riskide mudel, kus huvipakkuvaks sündmuseks on üleminek seisundist i seisundisse j ning ülejäänud võimalikud üleminekud on tsenseeritud.

R-i pakettis *survival* saab Coxi võrdeliste riskide mudeleid teha nii ühe kui ka mitme sündmuse korral funktsiooniga *coxph()*. Mitme olekuga mudelite korral saab ühel indiviidil toimuda mitu sündmust. Seega peab näiteks haigus-surm mudeli korral olema funktsioonile etteantavas andmestikus kaks rida kõigil indiviididel, kes jõudsid seisundisse *haige* (Therneau, T., Crowson, C. ja Atkinson, E., 2020). Selle tõttu tuleb mitme olekuga mudelite korral funktsioonis täpsustada argument *id*, mis on igal indiviidil unikaalne identifitseerija. Lisaks kasutatakse funktsioonis *coxph()* vähimisi parameetrite hinnangute standardvigade leidmisel robustseid standardvigu, et võtta arvesse olukorda, kus ühel indiviidil toimus mitu sündmust. (Therneau, T. M. *et al.*, 2022)

3 Andmete analüüs

Antud magistritöös uuritakse depressiooni kui kardiometaboolsete haiguste riskitegurit Tartu Ülikooli Eesti Geenivaramu andmete põhjal, rakendades elukestusanalüüsi meetodeid ja mitme olekuga mudeleid.

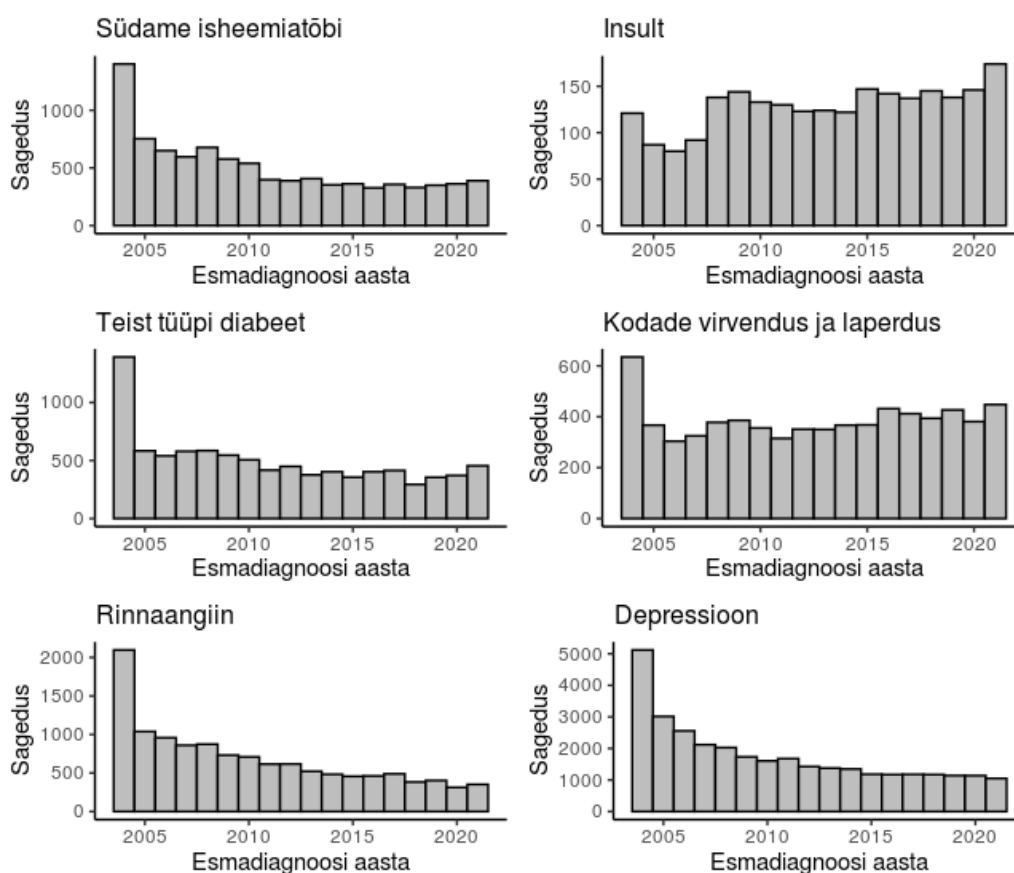
3.1 Andmete ülevaade

Tartu Ülikooli Eesti Geenivaramu andmebaasis on 2021. aasta lõpu seisuga üle 208 000 geenidoonori andmed. Seejuures on geenivaramuga liitumised toimunud kahes suuremas osas: valdavalt perearstikeskuste kaudu kuni aastani 2011 ($n = 51\,173$ geenidoonorit) ja kampaania „Kingime Eestile juubeliks 100 000 uut geenidoonorit“ raames aastatel 2018–2019 ($n = 150\,437$). Vahepealsetel aastatel on liitumisi toimunud vähem. Analüüsis kasutatakse andmeid 112 806 mittesuguluses geenidoonori kohta, kes olid liitumishetkel vähemalt 18-aastased ja kelle kohta on olemas informatsioon tervise andmebaasidest nagu Haigekassa. Kasutatakse vaid mittesuguluses indiviide, kuna töös rakendatavad elukesustanalüüsi meetodid eeldavad, et indiviidide elukestused on sõltumatud (Andersen, P. K. *et al.*, 1997).

Kõigi geenidoonorite kohta on teada sugu ja sünniaasta. Kuna täpset sünnikuupäeva ei väljastata doonorite anonüümsuse tagamiseks, siis kasutatakse sünnikuupäevana kõigil indiviididel vastava aasta 30. juunit. Analüüsis kasutatavatest geenidoonoritest on suuremal osal liitumishetke seisuga teada andmed ka teiste kardiometaboolsete haiguste riskitegurite kohta nagu kehamassiindeks, haridustase, suitsetamise staatus, alkoholi tarbimise staatus ja füüsiline aktiivsus. Haridustasemed on siinses töös jagatud kolmeks: kuni põhiharidus, keskharidus ja kõrgharidus. Suitsetamise staatusel ja alkoholi tarbimise staatusel on samuti kolm taset: mitte kunagi, endine ja praegune. Füüsilise aktiivsuse tunnus jaguneb kaheks: ei tegele tervisespordiga ja tegeleb tervisespordiga. Juhul, kui indiviid on surnud, siis on teada ka surmakuupäev. Üksikutel indiviididel on teada vaid surma-aasta, mille

korral kasutatakse surmakuupäevana vastava aasta 31. detsembrit.

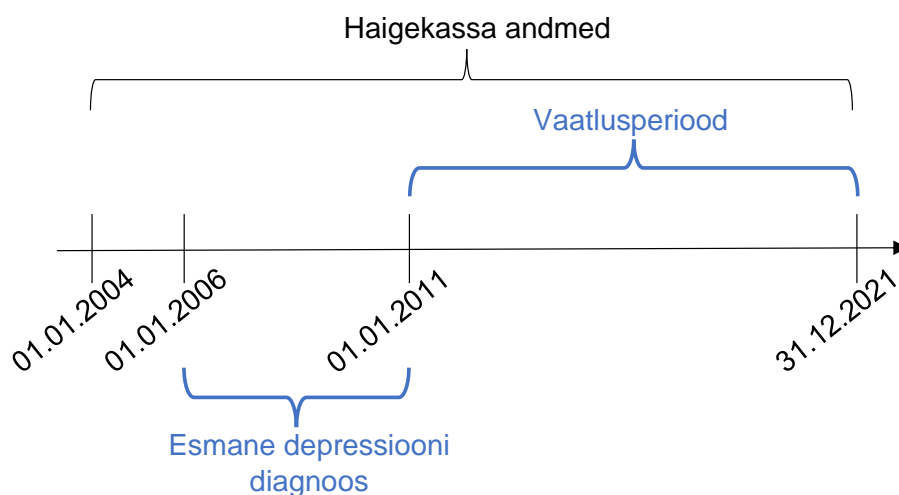
Andmed diagnooside ja ravimite kasutuse kohta on saadud linkimisel Haigekassa andmebaasiga, kus andmed on olemas alates 01.01.2004 kuni 31.12.2021. Varasemate diagnooside kohta täielikud andmed puuduvad, mistõttu ei saa täpselt määrata haiguse esmase diagnoosimise kuupäeva ega ka haiguste ajalist järjestust. Seda on näha ka Joonisel 3 kujutatud histogrammidelt, kus krooniliste haiguste puhul on 2004. aastal esmaste diagnooside arv väga kõrge võrreldes järgnevate aastatega. Kuna varasemate diagnooside kohta andmeid pole, siis leidub 2004. aasta „esmdiagnooside“ hulgas ka korduvaid diagnoose, mille tegelik esimene diagnoos on saadud enne aastat 2004.



Joonis 3: Huvipakkuvate kardiometaboolsete haiguste ja depressiooni esmdiagnoosid aastate kaupa TÜ Eesti Geenivaramu kohordis.

Lisaks on kardiometaboolsete haiguste hulgas südame isheemiatõve, teist tüüpi diabeedi ja rinnaangiini korral näha esmaste diagnooside arvus langevat trendi. See võib viidata kordusdiagnooside esinemisele esmadiagnooside hulgas ka pärast 2004. aastat. Südame isheemiatõve korral ühtlustub esmadiagnooside arv alates aastast 2011. Sealt alates on ka ülejäänud kardiometaboolsete haiguste esmadiagnooside arvud enam-vähem ühtlustunud. Selle tõttu kasutatakse esmaste diagnooside täpsemaks määratlemiseks uuringuperioodi algusena 2011. aastat.

Seega magistritöös uuritakse baastaseme depressiooni mõju kardiometaboolsete haiguste riskile 11-aastase vaatlusperioodi jooksul (01.01.2011–31.12.2021). Seejuures on indiviidil baastasemel depressioon, kui esmase depressiooni diagnoos on saadud aastatel 2006–2010, s.o kuni viis aastat enne uuringuperioodi algust (vt Joonis 4). Samamoodi toimitakse ka hüpertensiooni ja hüperkolesteroleemia, millele vastavad RHK-10 koodid on I10 ja E78.0, korral ehk leitakse, kas perioodil 2006–2010 on saadud esmane haiguse diagnoos. Ka erinevat tüüpi antidepressantide (SSRI, tritsüklilised, ülejäänud) kasutust vaadeldakse aastatel 2006–2010 ja leitakse retseptide väljaostmiste arv sellel perioodil.



Joonis 4: Vaatlusperiood ja baastaseme depressioon.

Indiviidid, kes said depressiooni diagnoosi enne 2006. aastat, eemaldatakse valimist. Lisaks saavad siinses analüüsis osaleda vaid indiviidid, kes olid elus uuringuperioodi alguses ehk 2011. aastal. Seega geenidonorid, kes surid enne 2011. aastat eemaldatakse valimist. Lisaks eemaldatakse kõigist analüüsides indiviidid, kellel on kõik viis kardiometaaboolset haigust esimest korda diagnoositud enne aastat 2011. Juhul, kui analüüsida haigusi eraldi, siis eemaldatakse valimist geenidonorid, kellel on vaadeldava haiguse esmadiagnoos saadud enne uuringuperioodi algust. Mitme olekuga mudelite korral vaatleme indiviidil esimesena saadud kardiometaaboolse haiguse diagnoosi. Seega eemaldatakse valimist indiviidid, kellel on esimene kardiometaaboolse haiguse diagnoos saadud enne 2011. aastat. Kui samal kuupäeval on kirjas rohkem kui ühe kardiometaaboolse haiguse esmadiagnoosi saamine, siis valitakse neist juhuslikult üks esimeseks kardiometaaboolseks haiguseks. Vaatluste eemaldamine on täpsemalt illustreeritud Lisas 1.

Väiksel osal indiviididest on kovariaatide (kehamassiindeksi, hariduse, suitetamise, alkoholi tarbimise ja füüsilise aktiivsuse) kohta andmed teada mitmest ajapunktist. Sellisel juhul võetakse kovariaatide andmed ajapunktist, mis on 2011. aastale kõige lähemal. Lisaks on kehamassiindeksi väärtused asendatud puuduvate väärtustega, kui nende väärtus on väiksem kui 16 või suurem kui 50.

3.2 Kirjeldav analüüs

Kirjeldava analüüsi jaoks kasutatakse 103 969 geenidonorist koosnevat andmestiku, mille hulka kuuluvad indiviidid on kaasatud vähemalt ühte analüüsi (vt Lisa 1). Nende keskmine vanus uuringu perioodi alguses on 39,1 aastat. Andmestikus on mehi 37 258 (35,8%) ning naisi 66 711 (64,2%).

Tabelis 1 on välja toodud vanuse, soo ja teiste kovariaatide näitajad depressioonita ja depressiooniga geenidonoritel. Depressiooni diagnoosiga geenidonorid on keskmiselt vanemad uuringuperioodi alguses ja suurema kehamassiindegiga. Seejuures on nende tunnuste korral keskmiste erinevus depressiooniga ja depressioonita indi-

viididel statistiliselt oluline. Depressiooniga indiviidide hulgas on naiste, endiste ja praeguste suitsetajate ning füüsiliselt mitteaktiivsete indiviidide osakaal suurem. Lisaks on depressiooniga indiviidide seas väiksem osakaal kõrgharitud ja praeguseid alkoholi tarbijaid. Hii-ruut testiga näidati, et hariduse, suitsetamise, alkoholi tarbimise ja füüsilise aktiivsuse korral on statistiliselt oluline seos depressiooniga. Sarnaseid seoseid depressiooni ja kovariaatidega on kirjeldatud ka teistes teadusuuringutes, kuid alkoholi tarbimisega on näidatud vastupidist seost, mis vajab täpsemat uurimist (nt Penninx, 2017).

Kehamassiindeksi, hariduse, suitsetamise, alkoholi tarbimise ja füüsilise aktiivsuse andmed on teada liitumisaasta seisuga. Kaks suuremat geenivaramuga liitumise perioodi on aset leidnud enne 2011. aastat või alates 2018. aastast, mistõttu on nende kovariaatide informatsioon osal indiviididel teada enne uurimisperioodi algust (enne 2011. aastat) ning osal indiviididel uurimisperioodi lõpust. Lisas 2 on näha, et kuigi valdavalt on seosed kovariaatide ja depressiooni esinemise vahel olemas ja samasuunalised nii enne kui ka pärast 2011. liitunute puhul, siis on teatud erinevused hilisemate ja varasemate liitujate vahel. Kõige ilmekam on erinevus suitsetamise staatuse puhul: enne 2011. aastat liitunute hulgas on ~13% endiseid ja ~30% praeguseid suitsetajaid, aga pärast 2011. aastat liitunute hulgas on ~32% endiseid ja ~15% praeguseid suitsetajaid. Lisaks on võimalik, et hilisemate liitujate korral on kovariaatide andmed teada peale kardiometaboolse haiguse diagnoosimist, mille järel on võib-olla eluviise muudetud. Seega kasutatakse järgnevates kovariaatidega mudelites ainult enne 2011. aastat liitunute andmeid, et võimalikult hästi kirjeldada indiviidide elustiili baastaseme seisuga.

Tabelis 2 on näidatud kardiometaboolsete haiguste jälgimiselsete juhtude arv depressiooniga ja depressioonita geenidoonorite hulgas. Oodatavalt on depressiooniga indiviidide hulgas enne 2011. aastat diagnoositud kardiometaboolsete haiguste patsientide osakaal suurem võrreldes depressioonita indiviididega (hii-ruut testi p -väärtus $< 2,2 \cdot 10^{-16}$). See ilmestab depressiooni ja kardiometaboolsete haiguste vahelist komorbiidsust. Seejuures on suurimad erinevused rinnaangiini ja insuldi

Tabel 1: Vanuse, kehamassiindeksi, soo, hariduse ja elustiiliga seotud tunnuste näitajad depressiooniga ja depressioonita indiviididel TÜ Eesti Geenivaramu kohordis.

Tunnus	Depressioonita <i>n</i> = 94 001	Depressiooniga <i>n</i> = 9 968	<i>p</i> -väärtus
Vanus (01.01.2011 seisuga) Keskmine (sd)	38,7 (16,5)	43,5 (15,1)	$< 2,2 \cdot 10^{-16}$
Kehamassiindeks Keskmine (sd)	25,9 (5,0)	26,3 (5,3)	$1,1 \cdot 10^{-12}$
Sugu Mees Naine	34 823 (37,0%) 59 178 (63,0%)	2 435 (24,4%) 7 533 (75,6%)	$< 2,2 \cdot 10^{-16}$
Haridus Kuni põhiharidus Keskharidus Kõrgharidus	5 663 (8,7%) 32 371 (49,6%) 27 182 (41,7%)	749 (10,5%) 3 687 (51,5%) 2 717 (38,0%)	$2,0 \cdot 10^{-11}$
Suitsetamine Mitte kunagi Endine Praegune	37 104 (54,9%) 16 927 (25,0%) 13 567 (20,0%)	3 658 (49,7%) 1 959 (26,6%) 1 750 (23,8%)	$< 2,2 \cdot 10^{-16}$
Alkohol Mitte kunagi Endine Praegune	5 198 (8,1%) 4 077 (6,4%) 54 755 (85,5%)	642 (9,1%) 590 (8,4%) 5 803 (82,5%)	$1,9 \cdot 10^{-12}$
Füüsiline aktiivsus Ei ole aktiivne On aktiivne	20 943 (34,2%) 40 239 (65,8%)	2 529 (38,6%) 4 021 (61,4%)	$1,6 \cdot 10^{-15}$

Märkus. Kvantitatiivsete tunnuste korral testitakse kahe grupi keskmiste erinevust t-testiga. Kvalitatiivsete tunnuste puhul testitakse seose olemasolu tunnuse ja depressiooni vahel hii-ruut testiga. Seejuures kehamassiindeksi, haridustaseme, suitsetamise staatuse, alkoholi tarbimise staatuse ja füüsilise aktiivsuse näitajate leidmisel kasutatakse indiviide, kelle kohta on vastavad andmed olemas.

korral, kus jälgimiselsete juhtude arv depressiooniga indiviidide seas on vastavalt 2 ja 2,4 korda suurem.

Hüpertensioon on diagnoositud 9 522 (9,2%) geenidoonoril ning hüperkolesteroleemia 3 700 (3,6%) geenidoonoril. Antidepressante on aastatel 2006–2010 välja ostnud 10 829 (10,4%) indiviidi. Antidepressante kirjutatakse lisaks depressiooni

Tabel 2: Jälgimiseelsed kardiometaboolsete haiguste juhud depressiooniga ja depressioonita indiviididel TÜ Eesti Geenivaramu kohordis.

Haigus	Depressioonita <i>n</i> = 94 001	Depressiooniga <i>n</i> = 9 968	p-väärtus
Südame isheemiatõbi			$< 2,2 \cdot 10^{-16}$
Jälgimiseelsed juhud	3 550 (3,8%)	640 (6,4%)	
Jälgimiseelselt terved	90 451 (96,2%)	9 328 (93,6%)	
Insult			$< 2,2 \cdot 10^{-16}$
Jälgimiseelsed juhud	426 (0,5%)	119 (1,2%)	
Jälgimiseelselt terved	93 575 (99,5%)	9 849 (98,8%)	
Teist tüüpi diabeet			$< 2,2 \cdot 10^{-16}$
Jälgimiseelsed juhud	3 422 (3,6%)	575 (5,8%)	
Jälgimiseelselt terved	90 579 (96,4%)	9 393 (94,2%)	
Kodade virvendus ja laperdus			$< 2,2 \cdot 10^{-16}$
Jälgimiseelsed juhud	1 827 (1,9%)	356 (3,6%)	
Jälgimiseelselt terved	92 174 (98,1%)	9 612 (96,4%)	
Rinnaangiin			$< 2,2 \cdot 10^{-16}$
Jälgimiseelsed juhud	4 926 (5,2%)	1 041 (10,4%)	
Jälgimiseelselt terved	89 075 (94,8%)	8 927 (89,6%)	

Märkus. Seose olemasolu testitakse hii-ruut testiga.

ravile välja ka muudel põhjustel (nt unehäirete ja migreeni korral). See kajastub ka andmestikus, kus antidepressante väljaostnud 10 829 indiviidist on depressiooni diagnoos 6 467 indiviidil. SSRI tüüpi antidepressante on välja ostnud 7 522 (7,2%) indiviidi, kelle hulgas keskmine väljaostetud retseptide arv on 5,2 ja mediaan on 3. Tritsüklilisi antidepressante on kasutanud 1 697 (1,6%) indiviidi, kelle seas keskmine väljaostetud retseptide arv on 2,9 ja mediaan on 1. Ülejäänud tüüpi antidepressante on kasutanud 4 114 (4,0%) geenidoonorit, seejuures nende seas on keskmine väljaostetud retseptide arv 4,2 ning mediaan on 2.

3.3 Naiivne elukestusanalüüs

Naiivses elukestusanalüüsis vaadeldakse korraga vaid üht kardiometaboolset haigust ning teisi kardiometaboolsete haiguste diagnoose arvesse ei võeta. Tegemist on n-ö naiivse lähenemisega, kuna ignoreeritakse teisi kardiometaboolseid haigusi, mis suures osas jagavad samu riskifaktoreid ning konkureerivad omavahel. Sellist lähenemist kasutatakse sageli ka teadusartiklites (nt Honigberg, M. C. *et al.*, 2022). Juhul, kui huvipakkuvaks sündmuseks on südame isheemiatõve diagnoos, on sündmus toimunud kõigil indiviididel, kellel on aastatel 2011–2021 diagnoositud südame isheemiatõbi. Ülejäänud indiviidide elukestused on paremalt tsenseeritud. Seejuures enne 2011. aastat südame isheemiatõve diagnoosi saanud indiviidid on valimist eemaldatud nagu kirjeldatud peatükis 3.1. Teiste kardiometaboolsete haigustega viiakse analüüs läbi analoogselt.

Kardiometaboolseid haigusi eraldi analüüsidest on tehtud iga haiguse kohta eraldi andmestik, kus on uuringuperioodi jooksul diagnoositud südame isheemiatõve uusi juhte 3 602 (3,6% 99 779 indiviidist), insuldi uusi juhte 1 362 (1,3% 103 424 indiviidist), teist tüüpi diabeedi uusi juhte 3 853 (3,9% 99 972 indiviidist), kodade virvenduse ja laperduse uusi juhte 3 818 (3,8% 101 786 indiviidist) ning rinnaangiini uusi juhte 4 547 (4,6% 98 002 indiviidist). Nende andmete korral on indiviidide keskmine jälgimisaeg 10,7 aastat. Haigustele vastavate valimite suurused, koos juhtude arvu ja keskmise jälgimisaajaga on näidatud Lisas 1.

Nendelt andmetelt on leitud esmahaigestumused 1000 inimaasta kohta depressioonita ja depressiooniga geenidonoritel (vt Tabel 3). Kõige suurem vahe depressiooniga ja depressioonita indiviidide esmahaigestumustes on rinnaangiini korral, kus vahe on 2,17 juhtu 1000 inimaasta kohta. Seejuures on kõikide haiguste korral esmahaigestumuste vahe depressiooniga ja depressioonita indiviidide hulgas erinev, kuna vahede 95% usaldusintervallid ei sisalda arvu 0.

Joonisel 5 on esitatud rinnaangiini korral Kaplan-Meieri hinnangud üleelamisfunktsioonile depressiooni diagnoosi ja soo lõikes. Seejuures on ajaskaalana kasutatud

Tabel 3: Esmahaigestumused 1000 inimaasta kohta depressiooniga ja depressioonita geenidonoritel ning esmahaigestumuste vahe koos 95% usaldusintervalliga.

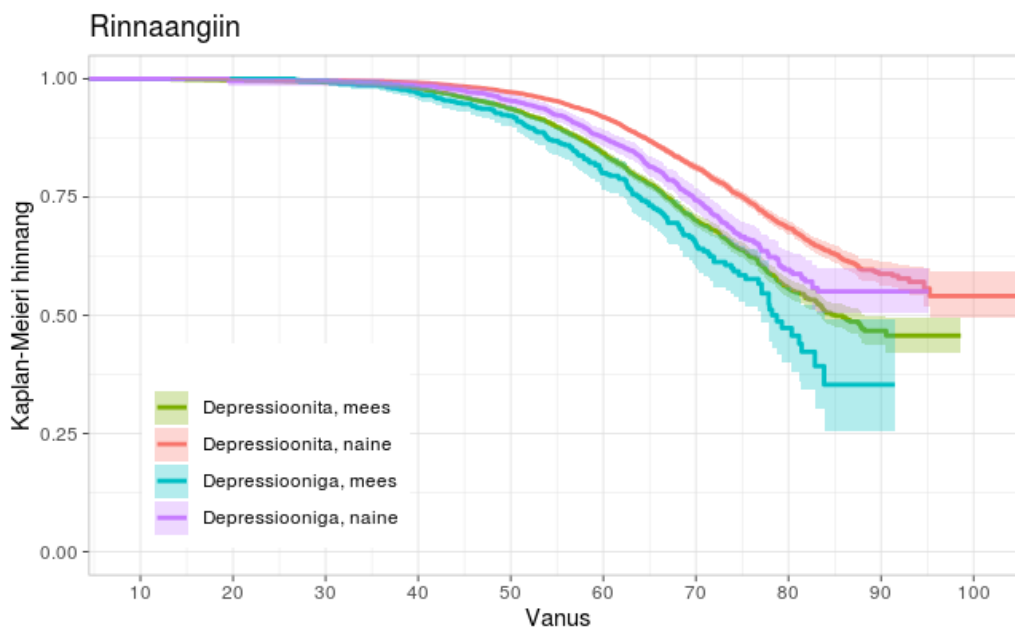
Haigus	Depr.-ta	Depr.-ga	Vahe	Usaldusint.
Südame isheemiatõbi	3,27	4,39	1,12	(0,69; 1,55)
Insult	1,19	1,49	0,29	(0,05; 0,54)
Teist tüüpi diabeet	3,50	4,74	1,25	(0,80; 1,69)
Kodade virvendus ja laperdus	3,40	4,50	1,09	(0,67; 1,52)
Rinnaangiin	4,18	6,35	2,17	(1,64; 2,70)

vanust ning arvestatud vasakult tõkestatusega. Jooniselt 5 on näha, et depressiooniga indiviididel diagnoositakse rinnaangiini varem nii meeste kui ka naiste seas. Seejuures on meestel elukestused väiksemad võrreldes naistega. Tõenäosus, et 80. eluaastaks on saadud rinnaangiini diagnoos, on depressioonita naistel 32%, depressiooniga naistel 40%, depressioonita meestel 44% ning depressiooniga meestel 53%.

Lisas 3 on välja toodud ka teiste kardiometaaboolsete haiguste Kaplan-Meieri hinnangud üleelamisfunktsioonile. Südame isheemiatõve, insuldi ning kodade virvenduse ja laperduse korral on samuti näha, et meeste ja depressiooniga indiviidide elukestused on väiksemad. Teist tüüpi diabeedi korral on kõverate vahelised erinevused väiksemad.

Kõikide kardiometaaboolsete haiguste korral tehakse kolm Coxi võrdeliste riskide mudelit:

1. mudelis on depressioon ja kihistamiseks kasutatakse sugu (soo korral ei olnud võrdeliste riskide eeldus täidetud ühegi kardiometaaboolse haiguse korral);



Joonis 5: Kaplan-Meieri hinnangud rinnaangiini üleelamisfunktsioonile koos 95% usaldusintervalliga depressiooni diagnoosi ja soo lõikes TÜ Eesti Geenivaramu andmete põhjal

2. mudelis on depressioon ja peamised riskifaktorid (kehamassiindeks, haridustase, suitsetamise staatus, alkoholi tarbimise staatus, füüsiline aktiivsus), seejuures kasutatakse sugu ja kehamassiindeksit (ala- või normaalkaalus ($KMI \leq 25$) ja ülekaalus või rasvunud ($KMI > 25$)) kihistamiseks;
3. mudelis on depressioon, peamised riskifaktorid, hüpertensioon, hüperkolesteroleemia ja erinevat tüüpi antidepressantide (SSRI, tritsüklilised, ülejäänud) väljaostmiste arv, seejuures kasutatakse sugu ja kehamassiindeksit (ala- või normaalkaalus ($KMI \leq 25$) ja ülekaalus või rasvunud ($KMI > 25$)) kihistamiseks.

Esimese mudeli puhul saab kasutada kõiki andmestikus olevaid indiviide. Teise ja kolmanda mudeli puhul kasutatakse vaid indiviide, kellel on kovariaatide andmed olemas ning teada enne 2011. aastat. Andmestike suurused koos juhtude arvudega on välja toodud Tabelis 4.

Tabel 4: Kõikide indiviidide arv ja juhtude arv kogu andmestikus ning kovariaatidega andmestikus.

Haigus	Kogu andmestik		Kovariaatidega andmestik	
	Indiviide	Juhte	Indiviide	Juhte
Südame isheemiatõbi	99 779	3 602 (3,6%)	20 728	1 212 (5,8%)
Insult	103 424	1 362 (1,3%)	22 461	639 (2,8%)
Teist tüüpi diabeet	99 972	3 853 (3,9%)	21 184	1 207 (5,7%)
Kodade virvendus ja laperdus	101 786	3 818 (3,8%)	21 738	1 422 (6,5%)
Rinnaangiin	98 002	4 547 (4,6%)	20 251	1 395 (6,9%)

Nii depressiooni, hüpertensiooni kui ka hüperkolesteroleemia puhul kasutatakse indikaatortunnust, mille väärtus on 1, kui diagnoos saadi aastatel 2006–2010, ning 0 vastasel juhul. Antidepressantide väljaostmiste arvud on samuti leitud perioodil 2006–2010. Kõikides Coxi võrdeliste riskide mudelites kasutatakse ajaskaalana vanust ning arvestatakse vasakult tõkestatusega. Kovariaatidega mudelites vastavad baastasemetele indiviidid, kes on depressioonita, keskharidusega, mitte kunagi suitsetanud, mitte kunagi alkoholi tarbinud, füüsiliselt mitteaktiivsed, hüpertensioonita ja hüperkolesteroleemiata. Võrdeliste riskide eelduste täidetust testitakse funktsiooniga *cox.zph()*, seejuures kontrollitakse eeldusi ka graafiliselt. Statistiliselt oluliste seoste leidmisel võetakse arvesse mitmest testimist Bonferroni korrektsiooni rakendades ehk olulisus nivooks on $\alpha = \frac{0,05}{5} = 0,01$. Tulemuste interpreteerimisel võrreldakse samast kihist indiviide, kellel teiste argumenttunnuste väärtused on samad.

Depressiooni mõju sooga kohandatud mudelis ja kovariaatidega kohandatud mudelis

Sooga kihistatud mudelis on depressiooni mõju statistiliselt oluline südame is-

heemiatõve, kodade virvenduse ja laperduse ning rinnaangiini korral ($p < 0,01$, vt Tabel 5). Depressiooniga indiviididel on hinnanguliselt 1,2 korda suurem risk saada südame isheemiatõbi, 1,2 korda suurem risk saada kodade virvendus ja laperdus ning 1,4 korda suurem risk saada rinnaangiin. Depressioon jääb oluliseks riskifaktoriks südame isheemiatõve (1,3 korda suurem risk) ja rinnaangiini (1,4 korda suurem risk) puhul ka mudelites, kus peamised riskifaktorid on arvesse võetud. Kodade virvenduse ja laperduse korral ei olnud depressiooni mõju kovariaatidega mudelis enam oluline ning ainuke oluline mõju on haridusel: kõrgharitudel on 1,2 ($\approx \frac{1}{0,83}$) korda madalam risk haigestuda võrreldes keskharitutega (vt Tabel 5). Südame isheemiatõve korral leiti teadaolevatest riskifaktoritest geenidoonorite andmetega oluline mõju suitsetamisel ja alkoholi tarbimisel. Praegustel suitsetajatel on 1,3 korda suurem risk haigestuda südame isheemiatõvesse võrreldes mitte kunagiste suitsetajatega. Praegustel alkoholi tarbijatel on 1,3 ($\approx \frac{1}{0,78}$) korda madalam risk kui mitte kunagi alkoholi tarbinutel. Alkoholi tarbimine on oluline ka rinnaangiini mudelis. Rinnaangiini saamise risk on praegustel alkoholi tarbijatel 1,4 ($\approx \frac{1}{0,73}$) korda väiksem võrreldes mitte kunagi alkoholi tarbinutega. Põhjuseks, miks südame isheemiatõve ja rinnaangiini korral on praegustel alkoholi tarbijatel väiksem risk, võib olla asjaolu, et antidepressantide (ja ka paljude teiste) ravimite koostarvitamine alkoholiga ei ole soovitatav. Seega võib praeguste alkoholi tarbijate hulgas olla vähem tervisehädadega indiviide.

Insuldile ja teist tüüpi diabeedile vastavad tulemused on esitatud Tabelis 6. Insuldi korral ei olnud depressiooni mõju oluline kummaski mudelis. Teist tüüpi diabeedi korral esimest mudelit ei hinnatud, sest võrdeliste riskide eeldus oli rikutud, ning kovariaatidega mudelis depressioonil olulist mõju ei ole.

Teadaolevatest riskifaktoritest on statistiliselt oluline efekt insuldile vaid suitsetamisel, kus praegustel suitsetajatel on 1,4 korda kõrgem risk võrreldes mitte kunagiste suitsetajatega. Teist tüüpi diabeedile on oluline mõju haridusel, suitsetamisel ja füüsilisel aktiivsusel. Kuni põhiharitudel on 1,3 korda kõrgem risk ning kõrg-

Tabel 5: Südame isheemiatõve, kodade virvenduse ja laperduse ning rinnaangiini korral riskide suhte hinnangud koos p -väärtustega naiivses lähenemises TÜ Eesti Geenivaramu andmete põhjal.

	1. mudel		2. mudel	
	$\exp(\hat{\beta}_i)$	p -väärtus	$\exp(\hat{\beta}_i)$	p -väärtus
Südame isheemiatõbi				
Depressioon	1,22	$1,4 \cdot 10^{-4}$	1,28	$3,4 \cdot 10^{-3}$
Haridus: kuni põhiharidus			1,08	0,34
Haridus: kõrgharidus			0,92	0,25
Suitsetamine: endine			0,98	0,82
Suitsetamine: praegune			1,32	$2,7 \cdot 10^{-4}$
Alkohol: endine			1,01	0,97
Alkohol: praegune			0,78	$5,6 \cdot 10^{-3}$
Füüsiline aktiivsus			0,92	0,20
Kodade virvendus ja laperdus				
Depressioon	1,18	$8,6 \cdot 10^{-4}$	1,04	0,67
Haridus: kuni põhiharidus			1,09	0,24
Haridus: kõrgharidus			0,83	$4,8 \cdot 10^{-3}$
Suitsetamine: endine			1,10	0,21
Suitsetamine: praegune			1,17	0,031
Alkohol: endine			1,15	0,28
Alkohol: praegune			0,99	0,87
Füüsiline aktiivsus			0,90	0,064
Rinnaangiin				
Depressioon	1,36	$5,5 \cdot 10^{-12}$	1,41	$6,8 \cdot 10^{-6}$
Haridus: kuni põhiharidus			0,95	0,56
Haridus: kõrgharidus			0,91	0,15
Suitsetamine: endine			1,00	0,99
Suitsetamine: praegune			1,07	0,34
Alkohol: endine			0,95	0,70
Alkohol: praegune			0,73	$1,5 \cdot 10^{-4}$
Füüsiline aktiivsus			0,92	0,13

haritutel 1,4 ($\approx \frac{1}{0,73}$) korda väiksem risk haigestuda teist tüüpi diabeeti võrreldes keskharitutega. Praegustel suitsetajatel on 1,3 korda suurem risk kui mitte kunagistel suitsetajatel. Füüsiliselt aktiivsetel indiviididel on 1,2 ($\approx \frac{1}{0,82}$) korda madalam risk kui füüsiliselt mitteaktiivsetel indiviididel.

Tabel 6: Insuldi ja teist tüüpi diabeedi korral riskide suhte hinnangud koos p -väärtustega naiivses lähenemises TÕ Eesti Geenivaramu andmete põhjal.

	1. mudel		2. mudel	
	$\exp(\hat{\beta}_i)$	p -väärtus	$\exp(\hat{\beta}_i)$	p -väärtus
Insult				
Depressioon	1,11	0,22	1,00	0,99
Haridus: kuni põhiharidus			1,09	0,38
Haridus: kõrgharidus			0,84	0,083
Suitsetamine: endine			1,07	0,55
Suitsetamine: praegune			1,37	$4,5 \cdot 10^{-3}$
Alkohol: endine			1,16	0,41
Alkohol: praegune			0,83	0,14
Füüsiline aktiivsus			0,90	0,21
Teist tüüpi diabeet				
Depressioon	(eeldus rikutud)		1,09	0,31
Haridus: kuni põhiharidus			1,25	$7,5 \cdot 10^{-3}$
Haridus: kõrgharidus			0,73	$3,2 \cdot 10^{-5}$
Suitsetamine: endine			1,04	0,63
Suitsetamine: praegune			1,28	$7,0 \cdot 10^{-4}$
Alkohol: endine			0,97	0,82
Alkohol: praegune			0,89	0,20
Füüsiline aktiivsus			0,82	$1,8 \cdot 10^{-3}$

Hüpertensiooni, hüperkolesteroleemia ja antidepressantide mõju

Kõikidel haigustel on 2. mudelis leitud olulised efektid säilinud ning kovariaatidele vastavad riskide suhte hinnangud jäänud ligikaudselt samaks. Küll aga on depressiooni efekti suurus üldjuhul kahanenud. Seejuures ei ole südame isheemiatõve korral enam depressiooni mõju olulisuse nivool $\alpha = 0,01$ oluline (vt Tabel 7). Rinnaangiini korral on depressiooni oluline mõju alles jäänud. Täpsemalt on depressiooniga ja depressioonita indiviidide riskide suhe hinnaguliselt 1,3. Varasemates rinnaangiini mudelites oli vastav riskide suhe ligikaudu 1,4 (vt Tabel 7).

Insuldi, teist tüüpi diabeedi ning kodade virvenduse ja laperduse korral ei ole hüpertensioon, hüperkolesteroleemia ega ühegi antidepressandi tüübi kasutus olulisuse nivool $\alpha = 0,01$ oluline (vt Lisa 4). Südame isheemiatõve mudelis on oluline

mõju hüpertensioonil, mis suurendab hinnanguliselt 1,3 korda südame isheemiatõve saamise riski (vt Tabel 7). Veel on oluline väljaostetud SSRI retseptide arv rinnaangiini mudelis (vt Tabel 7). Seejuures, mida suurem on väljaostetud SSRI retseptide arv, seda suurem on rinnaangiini haigestumise risk. Näiteks võrreldes SSRI antidepressante mittekasutanud indiviide nendega, kes on SSRI tüüpi antidepressante välja ostnud viiel korral, on viimastel 1,1 ($\approx 1,02^5$) korda suurem risk saada rinnaangiini diagnoos. Kuid indiviidil, kes on SSRI antidepressante väljaostnud kümnel korral, on rinnaangiini risk juba 1,2 ($\approx 1,02^{10}$) korda suurem võrreldes indiviidiga, kes pole SSRI antidepressante kasutanud. Kuna teaduskirjanduses on SSRI tüüpi antidepressantidega näidatud vastupidist efekti, siis tuleks teha täpsustavaid analüüse.

Tabel 7: Südame isheemiatõve ning rinnaangiini korral riskide suhte hinnangud koos p -väärtustega naiivses lähenemises TÜ Eesti Geenivaramu andmete põhjal.

	3. mudel			
	Südame isheemiatõbi		Rinnaangiin	
	$\exp(\hat{\beta}_i)$	p -väärtus	$\exp(\hat{\beta}_i)$	p -väärtus
Depressioon	1,25	0,015	1,33	$4,2 \cdot 10^{-4}$
Haridus: kuni põhiharidus	1,09	0,28	0,96	0,63
Haridus: kõrgharidus	0,93	0,30	0,91	0,17
Suitsetamine: endine	0,97	0,72	1,00	0,96
Suitsetamine: praegune	1,32	$2,1 \cdot 10^{-4}$	1,07	0,35
Alkohol: endine	1,00	0,99	0,93	0,62
Alkohol: praegune	0,78	$5,2 \cdot 10^{-3}$	0,72	$1,3 \cdot 10^{-4}$
Füüsiline aktiivsus	0,93	0,23	0,91	0,11
Hüpertensioon	1,28	$2,2 \cdot 10^{-4}$	1,13	0,067
Hüperkolesteroleemia	1,20	0,063	1,21	0,043
SSRI antidepress.	0,99	0,62	1,02	$4,7 \cdot 10^{-3}$
Tritsükliilised antidepress.	1,01	0,60	1,00	0,92
Teised antidepress.	1,01	0,51	0,99	0,64

3.4 Mitme olekuga mudelid

Esmalt tehakse konkureerivate riskide mudelid, kus vaadeldakse esimesena saadud kardiometaboolset haigust, kus konkureerivad südame isheemiatõbi, insult, teist tüüpi diabeet, kodade virvendus ja laperdus ning rinnaangiin. Lisaks võetakse konkureerivaks riskina arvesse ka surma. Järgmisena tehakse haigus-surm mudel, kus indiviid on seisundis *haige*, kui tal on diagnoositud ükskõik milline uuritavast viiest kardiometaboolsest haigusest.

Mitme olekuga mudelite jaoks kasutatakse 92 611 geenidoonorist koosnevat andmestikku, kust on eemaldatud indiviidid, kellel on enne 2011. aastat diagnoositud mingi kardiometaboolne haigus (vt Lisa 1). Nende hulgast 9 275 (10,0%) indiviidil on uuringuperioodi jooksul diagnoositud vähemalt üks huvipakkuvatest kardiometaboolsetest haigustest. Selle andmestiku korral on indiviidide keskmine jälgimisaeg 10,4 aastat.

Nii konkureerivate riskide mudelite kui ka haigus-surm mudelite korral tehakse Coxi võrdeliste riskide mudeleid kolmel erineval juhul (nagu naiivses lähenemises):

1. mudelis on depressioon ja kihistamiseks kasutatakse sugu (valdavalt oli soo korral võrdeliste riskide eeldus rikutud);
2. mudelis on depressioon ja peamised riskifaktorid, seejuures kasutatakse sugu ja kehamassiindeksit (ala- või normaalkaalus ($KMI \leq 25$) ja ülekaalus või rasvunud ($KMI > 25$)) kihistamiseks.
3. mudelis on depressioon, peamised riskifaktorid, hüpertensioon, hüperkolesteroleemia ja antidepressandid, seejuures kasutatakse sugu ja kehamassiindeksit (ala- või normaalkaalus ($KMI \leq 25$) ja ülekaalus või rasvunud ($KMI > 25$)) kihistamiseks.

Jällegi 1. mudeli korral kasutatakse kõiki andmestikus olevaid indiviide ning 2. ja 3. mudeli korral vaid indiviide, kellel on kovariaatide andmed olemas ja need teada enne 2011. aastat. Andmeid analüüsid kasutatakse ajaskaalana vanust ning

arvestatakse vasakult tõkestatusega. Tulemuste tõlgendamiseks kasutatakse konkureerivate riskide mudelite korral olulisuse nivood $\alpha = \frac{0,05}{5} = 0,01$ ning haigus-surm mudelite korral olulisuse nivood $\alpha = \frac{0,05}{3} \approx 0,017$. Tulemuste interpreteerimisel võrreldakse samast kihist indiviide, kellel teiste argumenttunnuste väärtused on samad.

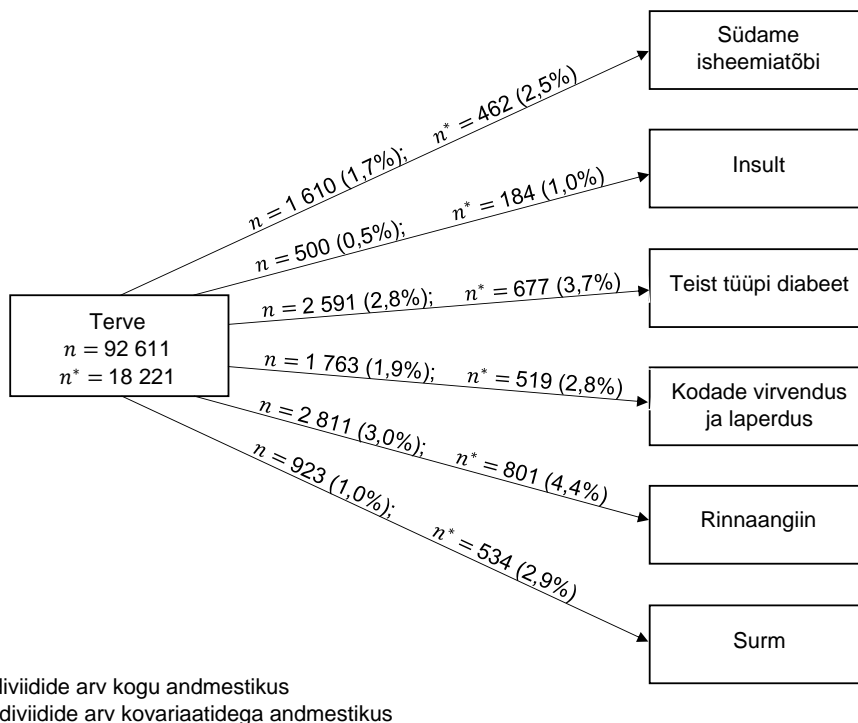
Üleminek seisundist *terve* seisundisse *surm* on olemas nii konkureerivate riskide kui ka haigus-surm mudelis. Kuna Coxi võrdeliste riskide hinnangute leidmisel teised sündmused tsenseeritakse, siis on mudeli tulemused sellele üleminekule samad mõlemas mitme olekuga mudelis. Seega on üleminekule *terve* \rightarrow *surm* vastavad mudelid välja toodud vaid haigus-surm mudelite juures.

3.4.1 Konkureerivate riskide mudel

Töös kasutatav konkureerivate riskide mudel koos indiviidide arvudega vastavatel üleminekutel on kujutatud Joonisel 6. Näeme, et nii kogu andmestikus, kui ka kovariaatidega andmestikus, on kõige rohkem esimese kardiometaaboolse haigusena diagnoositud rinnaangiini, kelle arv on kogu andmestikus 2 811 ja kovariaatidega andmestikus 801. Kõige vähem on geenidoonorid saanud esimese kardiometaaboolse haigusena insulti, mis on kogu andmestikus diagnoositud 500-l indiviidil ning kovariaatidega andmestikus 184-l indiviidil. Ilma kardiometaaboolse haiguseta ehk tervena suri kogu andmestikus 923 geenidoonorit ning kovariaatidega andmestikus 534 geenidoonorit.

Aalen-Johanseni hinnangute abil leitud seisundis olemise tõenäosused on kõikide sündmuste korral näidatud Joonisel 7. Tõenäosus saada vanuseks 80 esimese kardiometaaboolse haigusena rinnaangiin on 20%, teist tüüpi diabeet on 17%, kodade virvendus ja laperdus on 14%, südame isheemiatõbi on 12% ning insult 4%. Tõenäosus, et 80. eluaastaks ollakse elus ja terve (ilma kardiometaaboolse haiguseta) on 27%.

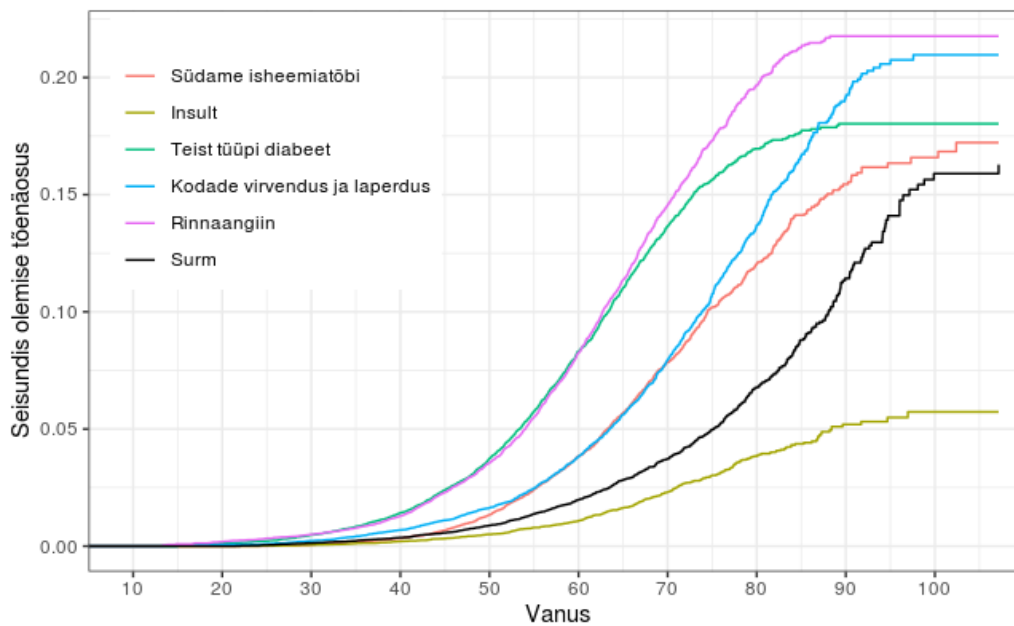
Sellise konkureerivate riskide mudeli korral on leitud seisundis olemise tõenäosuste



Joonis 6: Konkureerivate riskide mudel koos indiviidide arvudega üleminekutes kogu andmestikus ja kovariaatidega andmestikus.

hinnangud ka depressiooni diagnoosi ja soo lõikes. Vastavad tõenäosused on rinnaangiini korral välja toodud Joonisel 8. Ülejäänud kardiometaboolsetele haigustele vastavad joonised on välja toodud Lisas 5. Surma kohta pole eraldi joonist välja toodud, kuid seda on konkureeriva riskina arvesse võetud. Rinnaangiini diagnoositakse mõlema soo korral depressiooniga indiviididel varem kui depressioonita indiviididel. Seda on näha ka kodade virvenduse ja laperduse korral, kuid teistel haigustel on erinevused väiksemad (vt Lisa 5). Küll aga on valdavalt kõigil haigustel näha, et meestel diagnoositakse kardiometaboolset haigust varem võrreldes naistega.

Tõenäosus, et vanuseks 80 on saanud esimese kardiometaboolse haigusena rinnaangiin on depressiooniga naistel 24%, depressioonita naistel 18% depressiooniga meestel 23% ja depressioonita meestel 21% (vt Joonis 8). Naiivse elukestusanalüüsi



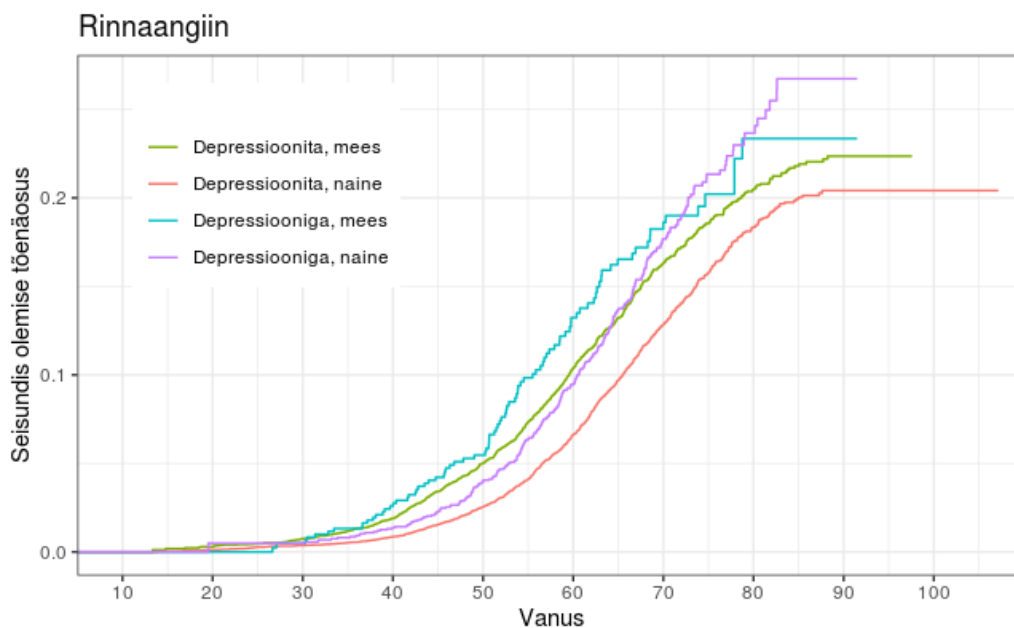
Joonis 7: Seisundis olemise tõenäosuste hinnangud konkureerivate riskide mudeli korral TÜ Eesti Geenivaramu andmete põhjal.

si korral olid vastavad tõenäosused suuremad (vt Joonis 5).

Põhjuseid, miks vastavad tõenäosused on konkureerivate riskide korral märgatavalt väiksemad, on mitmeid. Esiteks, naiivses lähenemises ignoreeritakse teisi konkureerivaid sündmusi nagu teised kardiometaaboolsed haigused ja surm. Teiseks mõjutab tulemusi ka see, et konkureerivate riskide korral on rinnaangiini juhte vähem (siin 2 811 juhtu, naiivses lähenemises 4 547 juhtu). Nimelt vaadeldi naiivse elukestusanalüüsi korral kõiki rinnaangiini diagnoose, kuid konkureerivate riskide korral, vaid neid rinnaangiini diagnoose, mis on ühtlasi ka esimene kardiometaaboolse haiguse diagnoos.

Depressiooni mõju sooga kohandatud mudelis ja kovariaatidega kohandatud mudelis

Depressiooni mõju oli sooga mudelis oluline kodade virvenduse ja laperduse ning rinnaangiini korral (vt Tabel 8). Seejuures depressiooniga indiviididel on 1,3 korda



Joonis 8: Seisundis olemise tõenäosuste hinnangud rinnaangiini korral depressiooni diagnoosi ja soo lõikes TÜ Eesti Geenivaramu andmete põhjal.

suurem risk saada kodade virvendus ja laperdus ning 1,4 korda suurem risk saada rinnaangiin. Kovariaatidega mudelis jäi depressiooni mõju oluliseks vaid rinnaangiini korral, seejuures riskide suhte hinnang on 1,5.

Võrreldes naiivse lähenemisega ei ole konkureerivate riskide lähenemise korral rinnaangiini mudelis enam alkoholi mõju oluline (naiivses lähenemises oli praeguste alkoholitarbijate risk väiksem). Kodade virvenduse ja laperduse korral on hariduse mõju säilinud. Kõrgaharitute risk on keskharitutega võrreldes 1,4 ($\approx \frac{1}{0,73}$) korda väiksem. Naiivses lähenemises oli kõrgharitute risk 1,2 ($\approx \frac{1}{0,83}$) korda väiksem võrreldes keskharitutega.

Naiivses lähenemises oli nii 1. mudeli kui ka 2. mudeli korral depressioonil oluline mõju südame isheemiatõvele. Konkureerivate riskide lähenemisega ei ole südame isheemiatõve korral depressiooni mõju kummaski mudelis oluline (vt Tabel 9). Sarnaselt naiivse lähenemisega ei ole depressioon mõju kummaski mudelis oluline insuldi korral. Teist tüüpi diabeedi korral oli võrdeliste riskide eeldus depressiooni

Tabel 8: Kodade virvenduse ja laperduse ning rinnaangiini riskide suhte hinnangud koos p -väärtustega konkureerivate riskide mudelis TÜ Eesti Geenivaramu andmete põhjal.

	1. mudel		2. mudel	
	$\exp(\hat{\beta}_i)$	p -väärtus	$\exp(\hat{\beta}_i)$	p -väärtus
Kodade virvendus ja laperdus				
Depressioon	1,31	$4,0 \cdot 10^{-4}$	1,14	0,36
Haridus: kuni põhiharidus			1,13	0,31
Haridus: kõrgharidus			0,73	$6,3 \cdot 10^{-3}$
Suitsetamine: endine			1,37	0,021
Suitsetamine: praegune			1,29	0,032
Alkohol: endine			0,82	0,45
Alkohol: praegune			0,95	0,72
Füüsiline aktiivsus			0,96	0,63
Rinnaangiin				
Depressioon	1,44	$1,3 \cdot 10^{-10}$	1,51	$3,5 \cdot 10^{-5}$
Haridus: kuni põhiharidus			0,93	0,59
Haridus: kõrgharidus			0,91	0,29
Suitsetamine: endine			0,93	0,55
Suitsetamine: praegune			0,99	0,95
Alkohol: endine			1,22	0,31
Alkohol: praegune			0,86	0,19
Füüsiline aktiivsus			0,94	0,41

korral rikitud, mistõttu depressiooni mõju seal ei hinnatud, kuid kovariaatidega mudelis depressioonil jällegi olulist mõju ei ole.

Südame isheemiatõve korral on kovariaatidega mudelis oluline vaid kordaja, mis vastab praegustele suitsetajatele (vt Tabel 9). Praeguste suitsetajate risk on 1,4 korda suurem võrreldes mitte kunagiste suitsetajatega. Naiivses lähenemises oli vastav riskide suhe 1,3 ning lisaks oli oluline ka praegustele alkoholi tarbijatele vastav kordaja. Insuldi mudelis oli naiivse lähenemise korral oluline vaid praeguste suitsetajate kordaja, mis konkureerivate riskide lähenemise korral enam oluline ei ole. Teist tüüpi diabeedi korral ei ole enam olulised kuni põhiharidusele vastav kordaja ning praeguste suitsetajate kordaja. Teised olulised tulemused on teist

Tabel 9: Südame isheemiatõve, kodade virvenduse ja laperduse ning rinnaangiini korral riskide suhte hinnangud koos p -väärtustega konkureerivate riskide mudelis TÜ Eesti Geenivaramu andmete põhjal.

	1. mudel		2. mudel	
	$\exp(\hat{\beta}_i)$	p -väärtus	$\exp(\hat{\beta}_i)$	p -väärtus
Südame isheemiatõbi				
Depressioon	1,17	0,061	1,12	0,47
Haridus: kuni põhiharidus			1,26	0,090
Haridus: kõrgharidus			1,05	0,69
Suitsetamine: endine			1,00	0,97
Suitsetamine: praegune			1,44	$1,9 \cdot 10^{-3}$
Alkohol: endine			1,29	0,28
Alkohol: praegune			0,71	0,027
Füüsiline aktiivsus			0,91	0,35
Insult				
Depressioon	1,05	0,73	0,79	0,39
Haridus: kuni põhiharidus			1,38	0,13
Haridus: kõrgharidus			0,93	0,68
Suitsetamine: endine			1,06	0,78
Suitsetamine: praegune			1,25	0,22
Alkohol: endine			1,46	0,29
Alkohol: praegune			0,74	0,20
Füüsiline aktiivsus			0,72	0,040
Teist tüüpi diabeet				
Depressioon	(eeldus rikutud)		1,12	0,35
Haridus: kuni põhiharidus			1,32	0,018
Haridus: kõrgharidus			0,65	$2,5 \cdot 10^{-5}$
Suitsetamine: endine			1,06	0,63
Suitsetamine: praegune			1,25	0,018
Alkohol: endine			1,12	0,62
Alkohol: praegune			0,90	0,42
Füüsiline aktiivsus			0,76	$1,2 \cdot 10^{-3}$

tüüpi diabeedi korral säilinud. Kõrgharitud teist tüüpi diabeedi risk on võrreldes keskharitudega $1,5 (\approx \frac{1}{0,65})$ korda väiksem ning füüsiliselt aktiivsetel on risk $1,3 (\approx \frac{1}{0,76})$ korda madalam võrreldes mitteaktiivsetega. Varasemalt olid vastavad riskide suhete hinnangud $1,4 (\approx \frac{1}{0,73})$ ja $1,2 (\approx \frac{1}{0,82})$.

Seega valdavalt on tulemused naiivse lähenemisega võrreldes küllaltki sarnased. Uusi olulisi mõjusid juurde ei tulnud, kuid mitmed varasemalt olulised efektid on konkureerivate riskide lähenemise korral muutunud mitteolulisteks. Seejuures on kovariaatidega mudelites depressioonil oluline mõju vaid rinnaangiinil. Tähelepanu tuleb pöörata erinevustele andmestikes, mida kasutatakse naiivse lähenemise ja konkureerivate riskide lähenemise korral. Konkureerivate riskide mudelis vaadeldakse vaid esmast kardiometaboolset diagnoosi ning lisaks on valimist välja jäetud inividid, kellel enne uuringuperioodi oli diagnoositud vähemalt üks kardiometaboolne haigus.

Hüpertensiooni, hüperkolesteroleemia ja antidepressantide mõju

Südame isheemiatõve korral on jällegi oluline hüpertensiooni mõju (vt Tabel 10). Hüpertensioon suurendab südame isheemiatõve riski 1,4 korda, seega efekti suurus on veidi suurenenud, sest varasemalt oli riskide suhteks 1,3. Lisaks on rinnaangiini korral säilinud oluline mõju SSRI tüüpi antidepressantide väljaostmiste arvule. Seejuures on hinnatud riskide suhe sama suur naiivses lähenemises saadud hinnanguga, st ühe võrra suurem SSRI retseptide väljaostmiste arv suurendab rinnaangiini riski 1,02 korda.

Võrreldes omavahel konkureerivate riskide lähenemisel tehtud 2. ja 3. mudelit, siis paneme tähele, et rinnaangiini mudelis on depressiooni mõju suurus vähenenud. Depressioon suurendab rinnaangiini riski 1,4 korda. Mudelis 2 oli vastav suurus 1,5. Südame isheemiatõve korral on depressioonile vastava kordaja sarnaselt 2. mudeliga ebaoluline, kuid jällegi on kordaja suurus vähenenud riskide suhte hinnangult 1,1 hinnangule 1,0. Ülejäänud kovariaatide efekti suurused on jäänud 2. mudeliga ligikaudselt samaks.

Sarnaselt naiivse lähenemisega ei ole ülejäänud kolme kardiometaboolse haiguse (insult, teist tüüpi diabeet ning kodade virvendus ja laperdus) korral hüpertensioonil, hüperkolesteroleemial ega antidepressantidel olulist mõju. Nende haiguste

Tabel 10: Südame isheemiatõve ja rinnaangiini korral riskide suhte hinnangud koos p -väärtustega konkureerivate riskide mudelis TÜ Eesti Geenivaramu andmete põhjal.

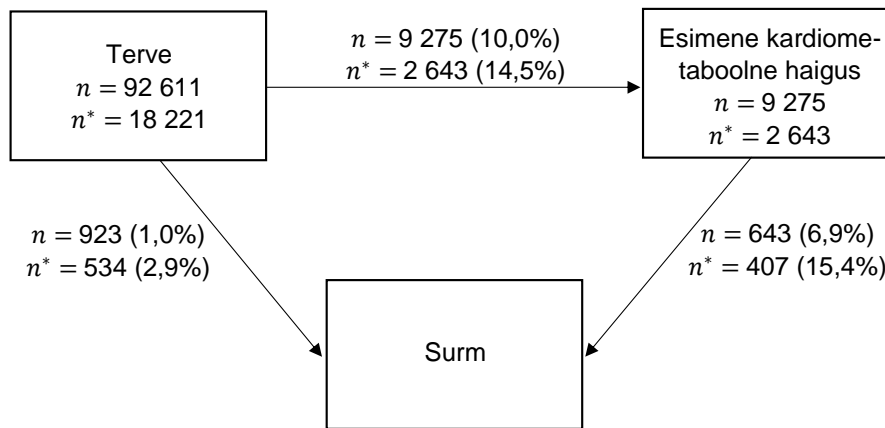
	3. mudel			
	Südame isheemiatõbi		Rinnaangiin	
	$\exp(\hat{\beta}_i)$	p -väärtus	$\exp(\hat{\beta}_i)$	p -väärtus
Depressioon	1,01	0,93	1,38	$3,4 \cdot 10^{-3}$
Haridus: kuni põhiharidus	1,28	0,067	0,94	0,64
Haridus: kõrgharidus	1,05	0,64	0,92	0,31
Suitsetamine: endine	1,00	0,98	0,93	0,51
Suitsetamine: praegune	1,44	$1,7 \cdot 10^{-3}$	0,99	0,93
Alkohol: endine	1,29	0,29	1,21	0,34
Alkohol: praegune	0,70	0,023	0,86	0,19
Füüsiline aktiivsus	0,91	0,35	0,94	0,40
Hüpertensioon	1,35	$8,9 \cdot 10^{-3}$	1,18	0,071
Hüperkolesteroleemia	1,37	0,071	1,24	0,12
SSRI antidepr.	1,01	0,36	1,02	$2,9 \cdot 10^{-3}$
Tritsüklilised antidepr.	0,96	0,46	1,01	0,73
Teised antidepr.	1,03	0,35	1,00	0,93

korral on 3. mudeli tulemused välja toodud Lisas 6. Depressiooni ning kovariaatide riskide suhte hinnangud on jäänud 2. mudeliga ligikaudselt samaks.

3.4.2 Haigus-surm mudel

Haigus-surm mudeli seisundid koos üleminekutele vastavate indiviidide arvudega on välja toodud Joonisel 9. Kuna haigus-surm mudeli korral kasutame samu indiviide, kes olid kaasatud ka konkureerivate riskide analüüsi, siis indiviidide arv on kogu andmestikus jälle 92 611 ning kovariaatidega andmestikus 18 221. Indiviidide arvud on tuttavad ka, kui vaadata üleminekut olekust *terve* olekusse *surm*. Lisaks näeme, et olekust *terve* olekusse *haige* üleminek on toimunud kogu andmestikus 9 275 geenidoonoril ning kovariaatidega andmestikus 2 643 geenidoonoril, s.o erinevatesse kardiometaboolsetesse haigustesse haigestunud indiviidide arvude summa konkureerivate riskide mudelis (vt Joonis 6). Kogu andmestikus on haiges-

tunud 9 275 geenidoonorist toimunud 643 indiviidil üleminek olekusse *surm* ning ülejäänud 8 632 indiviidi on sellel üleminekul paremalt tsenseeritud.



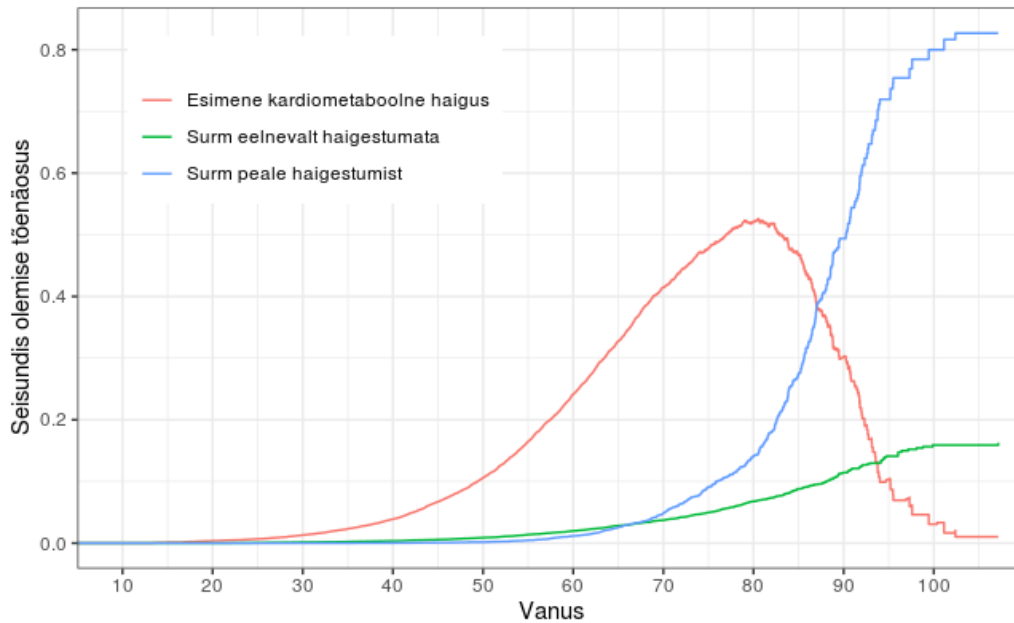
n – indiviidide arv kogu andmestikus
 n^* – indiviidide arv kovariaatidega andmestikus

Joonis 9: Haigus-surm mudel koos indiviidide arvudega üleminekutes kogu andmestikus ja kovariaatidega andmestikus.

Joonisel 10 on haigus-surm mudeli seisundis olemise tõenäosuste hinnagud. Üleminekule *terve* \rightarrow *haige* vastav joon esmalt tõuseb ja seejärel langeb, sest tegu on mööduva olekuga: olekutõenäosust suurendab see, kui keegi haigestub kardiometaboolsesse haigusesse, ning vähendab see, kui keegi peale haigestumist sureb ehk lahkub seisundist *haige*. Vanuses 80 on kõige rohkem indiviide seisundis *haige* ($p_{haige}(80) = 0,52$), mille järel kardiometaboolse haiguse seisundis olevate indiviidide osakaal hakkab langema, kuna rohkem indiviide liigub edasi olekusse *surm*. Vanuses 80 on tõenäosus olla surnud eelnevat haigestumata 7% ning surnud peale haigestumist 14%.

Depressiooni mõju sooga kohandatud mudelis ja kovariaatidega kohandatud mudelis

Tabelist 11 näeme, et baastaseme depressiooni mõju on oluline kardiometaboolse-



Joonis 10: Seisundis olemise tõenäosuste hinnangud haigus-surm mudeli korral TÜ Eesti geenivaramu andmete põhjal.

le haigusele ning surmale eelnevalt haigestumata. Kardiometaboolse haiguse risk on depressiooniga indiviididel 1,3 korda suurem. Kardiometaboolse haiguseta indiviididel on depressiooni diagnoosi korral suuremise risk 1,4 korda suurem. Peale haigestumist ei ole baastaseme depressiooni mõju surmale oluline ($p = 0,62$).

Kovariaatidega mudeli korral on depressiooni oluline mõju säilinud üleminekutele *terve* \rightarrow *haige* ja *terve* \rightarrow *surm* (vt Tabel 11). Seejuures on hinnatud riskide suhted sarnased 1. mudeliga. Tervetel indiviididel on depressiooni diagnoosi korral 1,2 korda suurem risk saada kardiometaboolne haigus ning 1,5 korda suurem risk surra. Baastaseme depressioonil jällegi peale haigestumist suuremisele olulist mõju ei ole.

Üleminekule *terve* \rightarrow *haige* on oluline mõju kõikidel kovariaatidel. Kuni põhiharidusega indiviididel on kardiometaboolse haiguse risk 1,2 korda suurem ning kõrgharitutel 1,2 ($\approx \frac{1}{0,83}$) korda väiksem võrreldes keskharitutega. Praegustel suitsetajatel on 1,2 korda suurem risk haigestuda võrreldes mitte kunagi suitsetanutega. Praegustel alkoholi tarbijatel on risk 1,2 ($\approx \frac{1}{0,85}$) korda väiksem kui mitte kunagi

Tabel 11: Haigus-surm mudelis riskide suhte hinnangud koos p -väärtustega TÜ Eesti Geenivaramu andmete põhjal.

	1. mudel		2. mudel	
	$\exp(\hat{\beta}_i)$	p -väärtus	$\exp(\hat{\beta}_i)$	p -väärtus
Kardiometaboolne haigus				
Depressioon	1,26	$5,4 \cdot 10^{-12}$	1,22	$1,4 \cdot 10^{-3}$
Haridus: kuni põhiharidus			1,16	0,015
Haridus: kõrgharidus			0,83	$1,5 \cdot 10^{-4}$
Suitsetamine: endine			1,06	0,33
Suitsetamine: praegune			1,20	$2,3 \cdot 10^{-4}$
Alkohol: endine			1,15	0,19
Alkohol: praegune			0,85	0,014
Füüsiline aktiivsus			0,87	$1,5 \cdot 10^{-3}$
Surm eelnevalt haigestumata				
Depressioon	1,42	$8,6 \cdot 10^{-4}$	1,46	$3,8 \cdot 10^{-3}$
Haridus: kuni põhiharidus			1,82	$3,4 \cdot 10^{-8}$
Haridus: kõrgharidus			0,83	0,12
Suitsetamine: endine			1,29	0,090
Suitsetamine: praegune			2,72	$< 2 \cdot 10^{-6}$
Alkohol: endine			1,17	0,47
Alkohol: praegune			0,82	0,18
Füüsiline aktiivsus			0,84	0,070
Surm peale haigestumist				
Depressioon	1,07	0,62	1,05	0,79
Haridus: kuni põhiharidus			1,39	0,013
Haridus: kõrgharidus			0,71	0,016
Suitsetamine: endine			1,24	0,18
Suitsetamine: praegune			1,45	$8,8 \cdot 10^{-3}$
Alkohol: endine			1,25	0,39
Alkohol: praegune			1,02	0,88
Füüsiline aktiivsus			0,86	0,21

alkoholi tarbinutel. Veel on füüsiliselt aktiivsete indiviidide risk 1,2 ($\approx \frac{1}{0,87}$) korda väiksem.

Üleminekule *terve* \rightarrow *surm* on teadaolevatest riskifaktoritest oluline mõju haridusel ja suitsetamisel. Tervetel on suuremise risk 1,8 korda suurem neil, kes on kuni põhiharidusega, võrreldes keskharitutega. Suremise (eelnevalt kardiometaboolset

haigust saamata) risk on praegustel suitsetajatel 2,7 korda suurem kui mittekunagistel suitsetajatel.

Üleminekule *haige* \rightarrow *surm* on oluline mõju haridusel ja suitsetamisel. Peale haigestumist on suremise risk kuni põhiharidusega indiviididel 1,4 korda suurem ning kõrgharitudel 1,4 ($\approx \frac{1}{0,71}$) korda väiksem võrreldes keskharitutega. Lisaks on peale kardiometaboolse haiguse diagnoosi praegustel suitsetajatel hinnanguliselt 1,5 korda suurem risk surra võrreldes mittekunagiste suitsetajatega.

Hüpertensiooni, hüperkolesteroleemia ja antidepressantide mõju

Hüperkolesteroleemia ja erinevat tüüpi antidepressantide kasutus ei ole oluline ühegi ülemineku korral (vt Tabel 12). Hüpertensiooni mõju ei saanud üleminekul *terve* \rightarrow *haige* usaldusväärset hinnata, sest võrdeliste riskide eeldus oli rikutud. Teistel üleminekul hüpertensiooni mõju ei olnud oluline.

Mudelites on depressioonil oluline mõju vaid surmale ilma haigestumiseta (vt Tabel 12). Võrreldes 2. mudeliga ei ole enam depressiooni mõju oluline kardiometaboolsele haigusele. Seejuures on kõikides kolmes mudelis depressioonile vastavate riskide suhte hinnangud vähenenud. Kardiometaboolse haiguse risk on depressiooniga indiviididel 1,1 korda suurem (2. mudelis oli riskide suhte hinnanguks 1,2). Tervete indiviidide suremise riski suurendab depressioon 1,4 korda (varasemalt oli vastav suurus 1,5).

Võrreldes 2. mudeliga on kovariaatide mõju suurused väga sarnased ning tunnused, mis olid olulised, on jäänud oluliseks ka 3. mudelis. Ainut ülemineku *haige* \rightarrow *surm* korral on kõrgharitudel vastav kordaja statistiliselt oluline olulisusnivool $\alpha = 0,05$, aga mitte arvestades mitmest testimist.

Tabel 12: Haigus-surm mudelis riskide suhte hinnangud koos p -väärtustega TÜ Eesti Geenivaramu andmete põhjal.

	3. mudel			
	Kardiometaboolne haigus		Surm eelnevalt haigestumata	
	$\exp(\hat{\beta}_i)$	p -väärtus	$\exp(\hat{\beta}_i)$	p -väärtus
Depressioon	1,14	0,053	1,43	$9,6 \cdot 10^{-3}$
Haridus: kuni põhiharidus	1,17	0,012	1,81	$5,4 \cdot 10^{-8}$
Haridus: kõrgharidus	0,83	$1,8 \cdot 10^{-4}$	0,83	0,13
Suitsetamine: endine	1,06	0,34	1,27	0,10
Suitsetamine: praegune	1,20	$2,4 \cdot 10^{-4}$	2,71	$< 2 \cdot 10^{-16}$
Alkohol: endine	1,15	0,20	1,18	0,45
Alkohol: praegune	0,85	0,014	0,83	0,21
Füüsiline aktiivsus	0,87	$1,5 \cdot 10^{-3}$	0,85	0,082
Hüpertensioon	1,15*	$8,0 \cdot 10^{-3}$ *	1,25	0,043
Hüperkolesteroleemia	1,09	0,30	0,78	0,24
SSRI antidepr.	1,01	0,065	1,00	0,81
Tritsüklilised antidepr.	0,99	0,75	1,04	0,32
Teised antidepr.	1,02	0,052	0,99	0,83
	Surm peale haigestumist			
	$\exp(\hat{\beta}_i)$	p -väärtus		
Depressioon	1,04	0,81		
Haridus: kuni põhiharidus	1,41	0,010		
Haridus: kõrgharidus	0,72	0,023		
Suitsetamine: endine	1,23	0,20		
Suitsetamine: praegune	1,44	0,011		
Alkohol: endine	1,27	0,36		
Alkohol: praegune	1,04	0,83		
Füüsiline aktiivsus	0,87	0,24		
Hüpertensioon	1,22	0,11		
Hüperkolesteroleemia	1,06	0,77		
SSRI antidepr.	0,98	0,48		
Tritsüklilised antidepr.	1,05	0,49		
Teised antidepr.	0,99	0,79		

* - võrdeliste riskide eeldus on rikutud. Märkus: kontrolliks tehti hüpertensioonita mudelid, kus tulemused olid väga sarnased.

Kokkuvõte

Töö eesmärgiks oli välja selgitada, kas depressioon on riskifaktoriks kardiometaboolsetele haigustele (südame isheemiatõbi, insult, teist tüüpi diabeet, kodade virvendus ja laperdus ning rinnaangiin). Selleks kasutatakse 11-aastast uuringuperioodi (2011–2021) ning hinnatakse kuni viis aastat enne uuringuperioodi (2006–2010) saadud esmase depressiooni diagnoosi mõju kardiometaboolsete haiguste tekkeriskile. Depressiooni mõju hinnatakse ka surmale nii kardometaboolse haiguse patsientidel kui ka tervetel indiviididel. Lisaks uuritakse hüpertensiooni, hüperkolesteroleemia ja erinevat tüüpi antidepressantide (SSRI, tritsüklilised ja ülejäänud) mõju kardiometaboolsetele haigustele.

Esmalt viidi läbi iga haiguse korral n-ö naiivne elukestusanalüüs, kus teisi haigusi ega surma konkureeriva riskina arvesse ei võetud. Sellisel juhul on sooga kohandatud mudelites depressioonil oluline mõju südame isheemiatõvele, kodade virvendusele ja laperdusele ning rinnaangiinile. Peamiste riskifaktoritega kohandatud mudelis jääb depressioon oluliseks riskifaktoriks südame isheemiatõvele ja rinnaangiinile. Seejärel uuriti depressiooni mõju kardiometaboolsetele haigustele konkureerivate riskide mudeliga, kus vaadeldi vaid esimesena saadud kardiometaboolset haigust ning konkureeriva riskina võeti arvesse ka surma. Valdavalt olid tulemused naiivse lähenemisega sarnased, kuid mitmed naiivse lähenemisega leitud olulised mõjud ei olnud konkureerivate riskide mudeliga enam olulised. Sooga kohandatud mudelis oli depressioonil oluline mõju vaid kodade virvendusele ja laperdusele ning rinnaangiinile. Peamiste riskifaktoritega kohandatud mudelis säilib depressiooni oluline mõju vaid rinnaangiinile. Depressioon suurendab rinnaangiini diagnoosi saamise riski hinnanguliselt 1,5 korda.

Depressiooni mõju hinnatakse surmale nii kardiometaboolse haiguse diagnoosiga kui ka diagnoosita indiviidide jaoks haigus-surma mudeli abil. Siinkohal on uurimise all esimesena diagnoositud kardiometaboolne haigus. Nii sooga kohandatud mudelis kui ka peamisi riskifaktoreid arvesse võtvas mudelis on baastaseme dep-

ressioonil mõju kardiometaboolsele haigusele ning surmale eelnevalt haigestumata. Depressioon suurendab kardiometaboolse haigus tekkeriski 1,2 korda ning kardiometaboolse haiguseta indiviide surma riski 1,5 korda.

Lisaks leiti töös, et hüpertensioonil on oluline mõju südame isheemiatõvele nii naiivse lähenemise kui ka konkureerivate riskide mudeli korral. Konkureerivate riskide mudeliga leiti, et hüpertensioon suurendab hinnanguliselt 1,4 korda südame isheemiatõvesse haigestumise riski. Lisaks leiti mõlemas lähenemises, et suurem väljaostetud SSRI retseptide arv on seotud rinnaangiini suurenenud riskiga (1,02 korda suurem risk ühe võrra suurema retseptide arvu kohta). Seejuures on depressioonile vastav riskide suhte hinnang vähenenud: depressioon suurendab rinnaangiini haigestumise riski 1,4 korda.

Seega depressioon on oluline riskifaktor teatud kardiometaboolsete haiguste tekkes sõltumata teadaolevatest riskifaktoritest, mida kinnitavad naiivne elukestusanalüüs ja komplekssem mitme olekuga mudelite rakendamine. Kui eesmärgiks on uurida depressiooni ja esmase kardiometaboolse haiguse diagnoosi saamise riski, siis on korrektsem kasutada mitme olekuga mudelit, mis võtab arvesse kardiometaboolsete haiguste komorbiidsust.

Töö edasiarendusena saaks depressiooni käsitleda ajas muutuva tunnusena. Lisaks võiks proovida Coxi segamudeleid või *frailty* mudeleid, mis võimaldaks depressiooni mõju kardiometaboolsetele haigustele hinnata kaasates ka suguluses olevaid indiviide. SSRI antidepressantide efekti täpsemaks hindamiseks võiks võrrelda antidepressantide kasutust depressiooniga indiviidide hulgas.

Kasutatud allikad

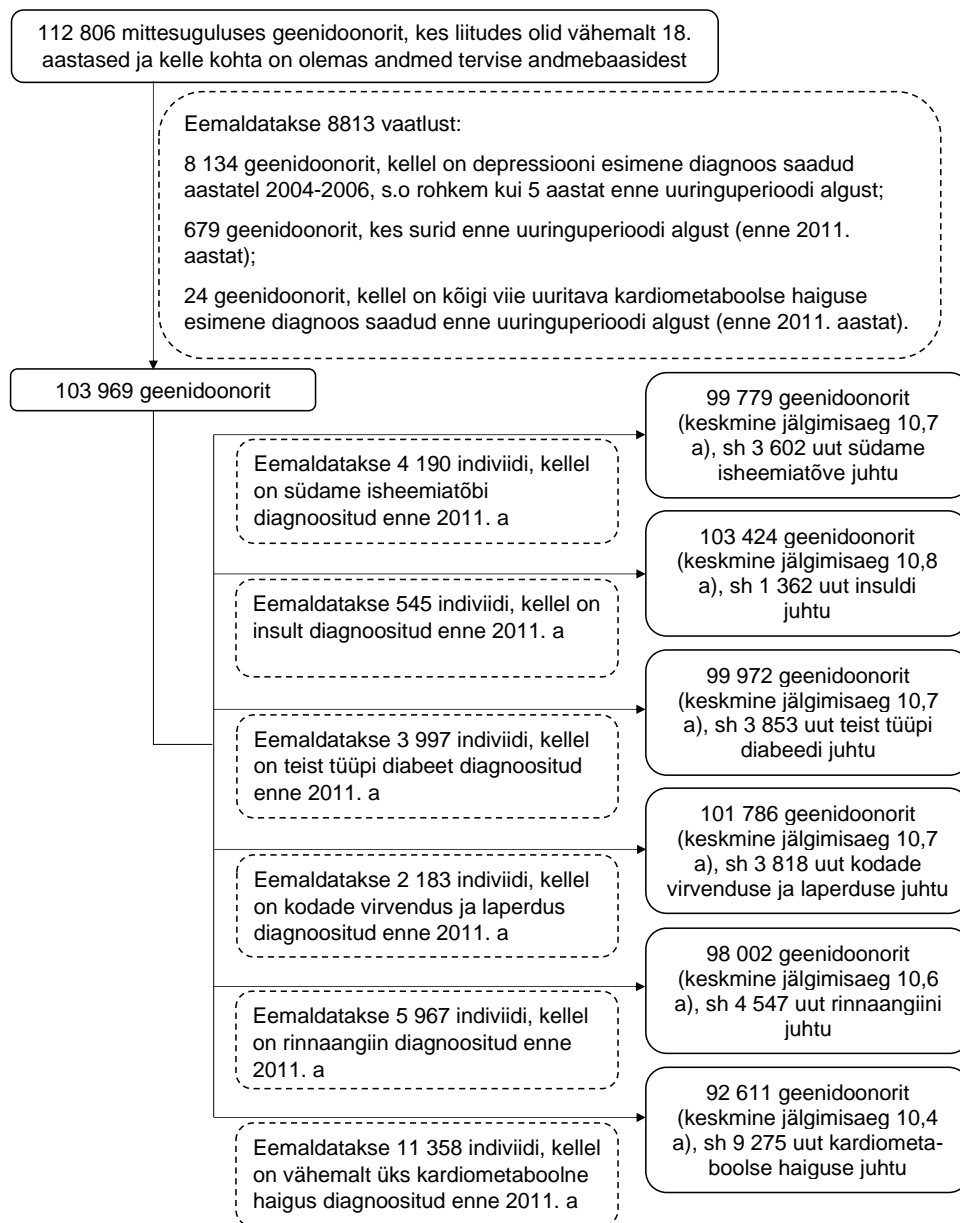
- Andersen, P. K., Klein, J. P., Knudsen, K. M. ja y Palacios, R. T. (1997). “Estimation of Variance in Cox’s Regression Model with Shared Gamma Frailties”. *Biometrics*, lk. 1475–1484. DOI: [10.2307/2533513](https://doi.org/10.2307/2533513).
- Barnard, K., Peveler, R. C. ja Holt, R. I. G. (2013). “Antidepressant medication as a Risk Factor for Type 2 Diabetes and Impaired Glucose Regulation: Systematic review”. *Diabetes care* 36.10, lk. 3337–3345. DOI: [10.2337/dc13-0560](https://doi.org/10.2337/dc13-0560).
- Borgan, Ø. (2014). “Aalen–Johansen Estimator”. *Wiley StatsRef: Statistics Reference Online*. DOI: [10.1002/9781118445112.stat06001.pub2](https://doi.org/10.1002/9781118445112.stat06001.pub2).
- Canchola, A. J., Stewart, S. L., Bernstein, L., West, D. W., Ross, R. K., Deapen, D., Pinder, R., Reynolds, P., Wright, W., Anton-Culver, H., Peel, D., Ziogas, A. ja Horn-Ross, P. L. (2003). “Cox regression using different time-scales”. *Western Users of SAS Software. San Francisco, California*.
- Collett, D. (2015). *Modelling survival data in medical research*. CRC press.
- De Waard, A.-K. M., Hollander, M., Korevaar, J. C., Nielen, M. M. J., Carlsson, A. C., Lionis, C., Seifert, B., Thilsing, T., de Wit, N. J., Schellevis, F. G. *et al.* (2019). “Selective prevention of cardiometabolic diseases: activities and attitudes of general practitioners across Europe”. *European Journal of Public Health* 29.1, lk. 88–93. DOI: [10.1093/eurpub/cky112](https://doi.org/10.1093/eurpub/cky112).
- Depression* (2021). Vaadatud 09.05.2022. URL: [https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-\(cvds\)](https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds)).
- Diabetes* (2021). Vaadatud 14.05.2022. URL: <https://www.who.int/news-room/fact-sheets/detail/diabetes>.
- Dimoula, A., Fotellis, D., Aivalioti, E., Delialis, D., Polissidis, A., Patras, R., Kokras, N. ja Stamatelopoulos, K. (2021). “Off-Target Effects of An-

- tidepressants on Vascular Function and Structure”. *Biomedicines* 10.1, lk. 56. DOI: [10.3390/biomedicines10010056](https://doi.org/10.3390/biomedicines10010056).
- Elderon, L. ja Whooley, M. A. (2013). “Depression and cardiovascular disease”. *Progress in cardiovascular diseases* 55.6, lk. 511–523. DOI: [10.1016/j.pcad.2013.03.010](https://doi.org/10.1016/j.pcad.2013.03.010).
- Fischer, K., Plummer, M., Pitkaniemi, J., Carstensen, B., Geroges, D. ja Läärä, E. (2019). “Statistical Practice in Epidemiology with R. Computer exercises”. URL: <https://bendixcarstensen.com/SPE/2019/pracs.pdf>.
- Fletcher, B., Gulanick, M. ja Lamendola, C. (2002). “Risk factors for type 2 diabetes mellitus”. *Journal of Cardiovascular Nursing* 16.2, lk. 17–23.
- Hare, D. L., Toukhsati, S. R., Johansson, P. ja Jaarsma, T. (2014). “Depression and cardiovascular disease: a clinical review”. *European heart journal* 35.21, lk. 1365–1372. DOI: [10.1093/eurheartj/eh462](https://doi.org/10.1093/eurheartj/eh462).
- Honigberg, M. C., Ye, Y., Dattilo, L., Sarma, A. A., Scott, N. S., Smoller, J. W., Zhao, H., Wood, M. J. ja Natarajan, P. (2022). “Low depression frequency is associated with decreased risk of cardiometabolic disease”. *Nature Cardiovascular Research* 1.2, lk. 125–131. DOI: [10.1038/s44161-021-00011-7](https://doi.org/10.1038/s44161-021-00011-7).
- Klein, J. P. ja Moeschberger, M. L. (2003). *Survival Analysis: Techniques for Censored and Truncated Data*. Kõide 2. Springer.
- Kleinbaum, D. G. (1996). “Introduction to Survival Analysis”. Teoses: *Survival Analysis: A Self-Learning Text*. New York, NY: Springer New York, lk. 1–44. DOI: [10.1007/978-1-4757-2555-1_1](https://doi.org/10.1007/978-1-4757-2555-1_1).
- Kleinbaum, D. G. ja Klein, M. (2005). “The Stratified Cox Procedure”. *Survival Analysis: A Self-Learning Text*, lk. 173–210. DOI: [10.1007/0-387-29150-4_5](https://doi.org/10.1007/0-387-29150-4_5).

- Knol, M. J., Twisk, J. W. R., Beekman, A. T. F., Heine, R. J., Snoek, F. J. ja Pouwer, F. (2006). “Depression as a risk factor for the onset of type 2 diabetes mellitus. A meta-analysis”. *Diabetologia* 49.5, lk. 837–845. DOI: [10.1007/s00125-006-0159-x](https://doi.org/10.1007/s00125-006-0159-x).
- Meira-Machado, L., de Uña-Álvarez, J., Cadarso-Suárez, C. ja Andersen, P. K. (2009). “Multi-state models for the analysis of time-to-event data”. *Statistical methods in medical research* 18.2, lk. 195–222. DOI: [10.1177/0962280208092301](https://doi.org/10.1177/0962280208092301).
- Penninx, B. W. J. H. (2017). “Depression and cardiovascular disease: Epidemiological evidence on their linking mechanisms”. *Neuroscience & Biobehavioral Reviews* 74, lk. 277–286. DOI: [10.1016/j.neubiorev.2016.07.003](https://doi.org/10.1016/j.neubiorev.2016.07.003).
- Putter, H., Fiocco, M. ja Geskus, R. B. (2007). “Tutorial in biostatistics: competing risks and multi-state models”. *Statistics in medicine* 26.11, lk. 2389–2430. DOI: [10.1002/sim.2712](https://doi.org/10.1002/sim.2712).
- Ross, S. M. (2014). *Introduction to probability models*. Academic press.
- Roy, T. ja Lloyd, C. E (2012). “Epidemiology of depression and diabetes: A systematic review”. *Journal of affective disorders* 142, S8–S21. DOI: [10.1016/S0165-0327\(12\)70004-6](https://doi.org/10.1016/S0165-0327(12)70004-6).
- Therneau, T., Crowson, C. ja Atkinson, E. (2020). “Multi-state models and competing risks”. *CRAN-R* (<https://cran.r-project.org/web/packages/survival/vignettes/compete.pdf>).
- Therneau, T. M., Lumley, T., Elizabeth, A. ja Cynthia, C. (2022). *Package ‘survival’*. URL: <https://cran.r-project.org/web/packages/survival/survival.pdf>.
- Thiébaud, A. C. M. ja Bénichou, J. (2004). “Choice of time-scale in Cox’s model analysis of epidemiologic cohort data: a simulation study”. *Statistics in medicine* 23.24, lk. 3803–3820. DOI: [10.1002/sim.2098](https://doi.org/10.1002/sim.2098).

Lisad

Lisa 1. Vaatluste eemaldamine.



Joonis 11: Vaatluste eemaldamine.

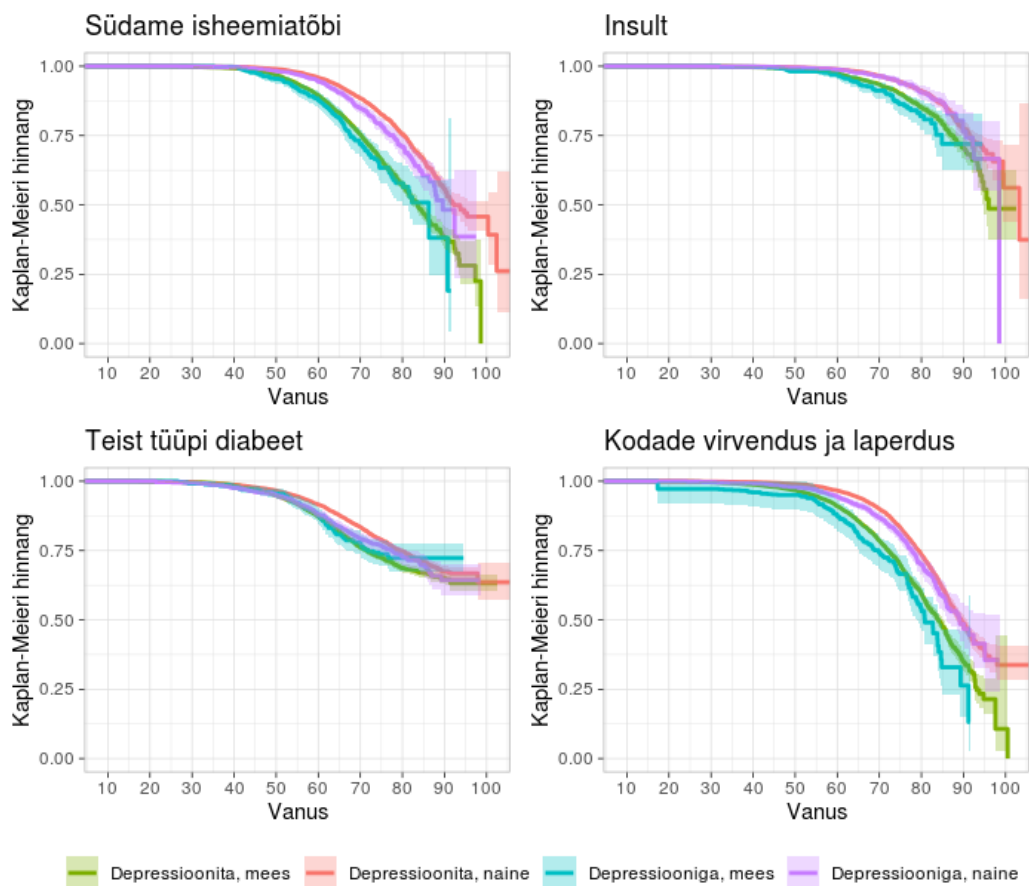
Lisa 2. Kovariaatide näitajad liitumisaasta ja depressiooni diagnoosi lõikes

Tabel 13: Vanuse, kehamassiindeksi, soo, hariduse ja elustiiliga seotud tunnuste näitajad liitumisaasta ja depressiooni diagnoosi lõikes TÜ Eesti Geeniaramu kohordis.

Tunnus	Liitumisaasta < 2011		Liitumisaasta ≥ 2011	
	Depr.-ta n = 23 837	Depr.-ga n = 3 330	Depr.-ta n = 70 164	Depr.-ga n = 6 638
Vanus liitumisel				
Keskmine (sd)	43,3 (16,9)	45,4 (16,0)	44,0 (15,5)	48,9 (14,0)
Kehamassiindeks				
Keskmine (sd)	26,2 (5,2)	26,4 (5,3)	25,7 (4,9)	26,3 (5,3)
Sugu				
Mees	8 782 (36,8%)	800 (24,0%)	26 041 (37,1)	1 635 (24,6)
Naine	15 055 (63,2%)	2 530 (76,0%)	44 123 (62,9)	5 003 (75,4)
Haridus				
Kuni põhiharidus	3 729 (15,7%)	549 (16,6%)	1 934 (4,7%)	200 (5,2%)
Keskharidus	14 038 (59,1%)	1 942 (58,5%)	18 333 (44,2%)	1 745 (45,5%)
Kõrgharidus	5 973 (25,2%)	826 (24,9%)	21 209 (51,1%)	1 891 (49,3%)
Suitsetamine				
Mitte kunagi	13 450 (56,6%)	1 806 (54,4%)	23 654 (53,9%)	1 852 (45,7%)
Endine	3 152 (13,3%)	483 (14,6%)	13 775 (31,4%)	1 476 (36,4%)
Praegune	7 146 (30,1%)	1 028 (31,0%)	6 421 (14,6%)	722 (17,8%)
Alkohol				
Mitte kunagi	2 696 (11,4%)	402 (12,3%)	2 502 (6,2%)	240 (6,4%)
Endine	1 290 (5,5%)	230 (7,0%)	2 787 (6,9%)	360 (9,6%)
Praegune	19 581 (83,1%)	2 646 (80,7%)	35 174 (86,9%)	3 157 (84,0%)
Füüsiline aktiivsus				
Ei ole aktiivne	9 392 (46,7%)	1 360 (49,7%)	11 551 (28,1%)	1169 (30,7%)
On aktiivne	10 699 (53,3%)	1 378 (50,3%)	29 540 (71,9%)	2643 (69,3%)

Märkus. Kehamassiindeksi, haridustaseme, suitsetamise staatuse, alkoholi tarbimise staatuse ja füüsilise aktiivsuse näitajate leidmisel kasutatakse indiviide, kelle kohta on vastavad andmed olemas.

Lisa 3. Kaplan-Meieri hinnangud



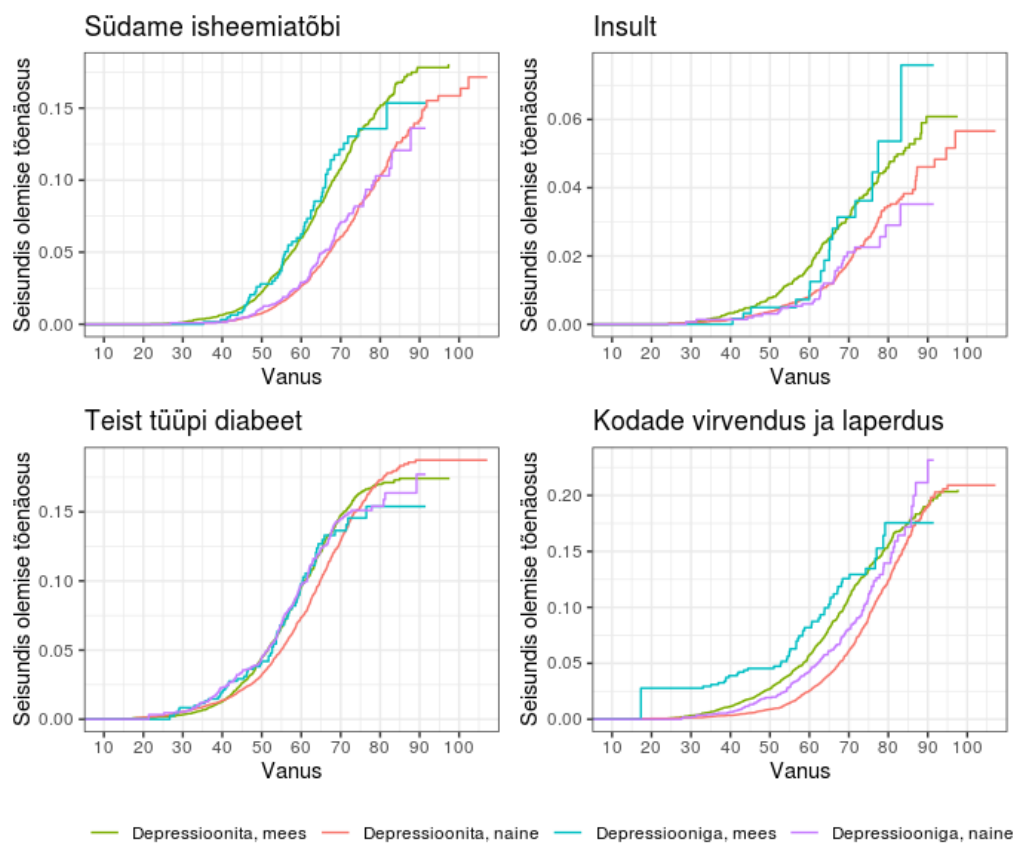
Joonis 12: Kaplan-Meieri hinnangud 95%-lise usaldusintervalliga depressiooniga ja depressioonita geenidoonoritel sugude lõikes.

Lisa 4. Mudelid naiivses elukestusanalüüsis

Tabel 14: Insuldi, teist tüüpi diabeedi ning kodade virvenduse ja laperduse korral riskide suhte hinnangud koos p -väärtustega naiivses lähenemises TÜ Eesti Geenivaramu andmete põhjal.

	3. mudel			
	Insult		Teist tüüpi diabeet	
	$\exp(\hat{\beta}_i)$	p -väärtus	$\exp(\hat{\beta}_i)$	p -väärtus
Depressioon	1,02	0,90	1,02	0,83
Haridus: kuni põhiharidus	1,09	0,38	1,26	$6,3 \cdot 10^{-3}$
Haridus: kõrgharidus	0,84	0,095	0,73	$3,8 \cdot 10^{-5}$
Suitsetamine: endine	1,07	0,55	1,04	0,63
Suitsetamine: praegune	1,37	$4,5 \cdot 10^{-3}$	1,28	$5,8 \cdot 10^{-4}$
Alkohol: endine	1,17	0,37	0,96	0,77
Alkohol: praegune	0,85	0,17	0,88	0,19
Füüsiline aktiivsus	0,90	0,23	0,82	$1,8 \cdot 10^{-3}$
Hüpertensioon	1,15	0,15	1,18	0,018
Hüperkolesteroleemia	0,97	0,82	1,10	0,35
SSRI antidepress.	0,98	0,37	1,00	0,71
Tritsükklilised antidepress.	1,05	0,051	1,00	0,94
Teised antidepress.	0,99	0,81	1,03	0,032
	Kodade virvendus ja laperdus			
	$\exp(\hat{\beta}_i)$	p -väärtus		
Depressioon	1,04	0,68		
Haridus: kuni põhiharidus	1,09	0,23		
Haridus: kõrgharidus	0,82	$4,3 \cdot 10^{-3}$		
Suitsetamine: endine	1,10	0,21		
Suitsetamine: praegune	1,18	0,030		
Alkohol: endine	1,15	0,30		
Alkohol: praegune	0,98	0,79		
Füüsiline aktiivsus	0,90	0,057		
Hüpertensioon	0,96	0,57		
Hüperkolesteroleemia	1,05	0,59		
SSRI antidepress.	1,00	0,59		
Tritsükklilised antidepress.	0,96	0,18		
Teised antidepress.	1,00	0,87		

Lisa 5. Seisundis olemise tõenäosuste hinnangud



Joonis 13: Seisundis olemise tõenäosused depressiooniga ja depressioonita geenidoonoritel sugude lõikes.

Lisa 6. Mudelid konkureerivate riskide korral

Tabel 15: Insuldi, teist tüüpi diabeedi ning kodade virvenduse ja laperduse korral riskide suhte hinnangud koos p -väärtustega konkureerivate riskide mudelis TÜ Eesti Geenivaramu andmete põhjal.

	3. mudel			
	Insult		Teist tüüpi diabeet	
	$\exp(\hat{\beta}_i)$	p -väärtus	$\exp(\hat{\beta}_i)$	p -väärtus
Depressioon	0,83	0,53	1,07	0,58
Haridus: kuni põhiharidus	1,39	0,12	1,33	0,017
Haridus: kõrgharidus	0,93	0,71	0,65	$3,6 \cdot 10^{-5}$
Suitsetamine: endine	1,06	0,81	1,06	0,62
Suitsetamine: praegune	1,26	0,20	1,24	0,018
Alkohol: endine	1,47	0,28	1,12	0,62
Alkohol: praegune	0,73	0,17	0,91	0,46
Füüsiline aktiivsus	0,72	0,039	0,77	$1,4 \cdot 10^{-3}$
Hüpertensioon	1,23	0,28	1,23	0,031
Hüperkolesteroleemia	1,32	0,33	0,90	0,55
SSRI antidepr.	0,96	0,30	1,00	0,80
Tritsüklilised antidepr.	0,94	0,65	1,02	0,60
Teised antidepr.	1,02	0,64	1,02	0,27
	Kodade virvendus ja laperdus			
	$\exp(\hat{\beta}_i)$	p -väärtus		
Depressioon	1,11	0,50		
Haridus: kuni põhiharidus	1,12	0,35		
Haridus: kõrgharidus	0,72	$4,3 \cdot 10^{-3}$		
Suitsetamine: endine	1,39	0,015		
Suitsetamine: praegune	1,29	0,032		
Alkohol: endine	0,82	0,45		
Alkohol: praegune	0,95	0,73		
Füüsiline aktiivsus	0,95	0,62		
Hüpertensioon	0,82	0,098		
Hüperkolesteroleemia	0,78	0,24		
SSRI antidepr.	1,00	0,97		
Tritsüklilised antidepr.	0,95	0,42		
Teised antidepr.	1,05	0,020		

Lihtlitsents lõputöö reprodutseerimiseks ja üldsusele kättesaadavaks tegemiseks

Mina, Laura Birgit Luitva,

1. annan Tartu Ülikoolile tasuta loa (lihtlitsentsi) minu loodud teose „Mitme olekuga mudelid depressiooni ja kardiometaboolsete haiguste vaheliste seoste hindamiseks TÜ Eesti Geenivaramu andmete põhjal“, mille juhendajad on Silva Kasela ja Lili Milani, reprodutseerimiseks eesmärgiga seda säilitada, sealhulgas lisada digitaalarhiivi DSpace kuni autoriõiguse kehtivuse lõppemiseni.
2. Annan Tartu Ülikoolile loa teha punktis 1 nimetatud teos üldsusele kättesaadavaks Tartu Ülikooli veebikeskkonna, sealhulgas digitaalarhiivi DSpace kaudu Creative Commons'i litsentsiga CC BY NC ND 3.0, mis lubab autorile viidates teost reprodutseerida, levitada ja üldsusele suunata ning keelab luua tuletatud teost ja kasutada teost ärieesmärgil, kuni autoriõiguse kehtivuse lõppemiseni.
3. Olen teadlik, et punktides 1 ja 2 nimetatud õigused jäävad alles ka autorile.
4. Kinnitan, et lihtlitsentsi andmisega ei riku ma teiste isikute intellektuaalomandi ega isikuandmete kaitse õigusaktidest tulenevaid õigusi.

Laura Birgit Luitva

17.05.2022