

Tartu Ülikool
Loodus- ja täppisteaduste valdkond
Matemaatika ja statistika instituut

Pille-Riin Vare

**Geneetilise varieeruvuse mõju antidepressantide tarvitamisele TÜ Eesti
Geenivaramu andmete põhjal**

Matemaatika ja statistika õppekava
Matemaatilise statistika eriala
Magistritöö (30 EAP)

Juhendajad: PhD Krista Fischer,
MSc Kristi Krebs

Tartu 2019

Geneetilise varieeruvuse mõju antidepressantide tarvitamisele TÜ Eesti Geenivaramu andmete põhjal

Magistritöö

Pille-Riin Vare

Magistritöö eesmärk on välja selgitada, kuidas geneetiline varieeruvus antidepressantide metabolismis olulisi ensüüme kodeerivates geenides *CYP2C19* ja *CYP2D6* mõjutab ühe antidepressandi vahetamist teise vastu ja kas üldises antidepressantide kasutamise mustris on erinevusi. Lisaks geneetilistele aspektidele pööratakse tähelepanu ka tausttunnustele nagu vanus ravi alustamise ajal, pikkus ja kaal. Kõigepealt antakse ülevaade depressioonist ning vajalikest geneetika terminitest. Seejärel kirjeldatakse kasutatavaid statistilisi mudeleid, milleks on logistiline regressioonimudel, Poissoni regressioonimudel, negatiivne binoomimudel, Coxi võrdeliste riskide mudel ning konkureerivate riskide mudel (nn Fine ja Gray mudel). Töö viimase osa moodustab Eesti Geenivaramu andmete analüüs, mille käigus uuritakse seoseid geneetilise varieeruvuse ning tausttunnuste ja erinevate välja ostetud ravimite retseptide arvu, üldise retseptide arvu ja depressiooniravimi vahetamise vahel. Poissoni mudeli hindamise tulemusena selgus, et need, kellel esineb deletsioon geenis *CYP2C19*, on kasutanud rohkem erinevaid ravimeid. Fine ja Gray mudeli põhjal sai järeldada, et neil, kellel ensüümi *CYP2D6* aktiivsusgrupp viitab aeglasele metaboliseerimisele, on suurem risk ravimi vahetamisele kui normaalsetel metaboliseerijatel.

Märksõnad: depressioon, andmeanalüüs, geneetiline muutlikkus

CERCS teaduseriala: statistika, operatsioonanalüüs, programmeerimine, finants- ja kindlustusmatemaatika (P160)

The Impact of Genetic Variations for Antidepressant Use Based on Estonian Genome Center's Data

Master's thesis

Pille-Riin Vare

The aim of the master's thesis is to find out how genetic variability in the *CYP2C19* and *CYP2D6* genes (known to be associated with the antidepressant metabolism), affects switching from one antidepressant to another, and the general use of antidepressants. In addition to the genetic aspects, the effect of background characteristics like gender, age at the beginning of treatment, height and weight of the subject is also studied. Firstly, a general overview of depression and also some basic concepts of genetics is given. Secondly, the statistical models used in the thesis are introduced: the logistic regression model, the Poisson regression model, the negative binomial model, the Cox proportional hazards model, and a model of competing risks. In the final part of the thesis the data from Estonian Genome Center is analysed. The analysis examines the associations between genetic variability and background characteristics, the number of prescriptions for the different purchased drugs, the overall number of prescriptions and the change of treatment of depression. According to Poisson regression model, individuals who have a deletion in gene *CYP2C19* are purchasing different antidepressants more frequently. Based on the Fine and Gray model for competing risks, it can be concluded that individuals in the group of poor metabolizers of enzyme *CYP2D6* have a bigger risk of switching antidepressant than normal metabolizers of *CYP2D6*.

Keywords: depression, data analysis, genetic variation

CERCS research specialisation: statistics, operation research, programming, actuarial mathematics (P160)

Sisukord

Sissejuhatus	5
1 Depressioon.....	6
2 Geneetika	7
3 Analüüsimeetodid	10
3.1 Logistiline regressioonimudel.....	10
3.2 Loendusandmete mudelid	11
3.2.1 Poissoni regressioonimudel	11
3.2.2 Negatiivne binoommudel.....	12
3.3 Elukestusanalüüs.....	13
3.3.1 Mõisted	13
3.3.2 Coxi võrdeliste riskide mudel	14
3.3.3 Mudelid konkureerivate riskide korral.....	15
3.3.4 Fine ja Gray konkureerivate riskide mudel.....	16
3.4 Lõplike mudelite leidmine	18
4 Andmete analüüs.....	19
4.1 Analüüsimeetodite rakendamine tarkvara R abil.....	19
4.2 Ülevaade andmetest	19
4.3 Kirjeldav analüüs	22
4.4 Analüüs mudelite abil	27
4.4.1 Loendusandmete mudelid retseptide arvule.....	27
4.4.2 Logistiline regressioonimudel ravimi vahetamisele	30
4.4.3 Coxi võrdeliste riskide mudel ning konkureerivate riskide mudel ravimi vahetamisele.....	31
Kokkuvõte.....	37
Kirjanduse loetelu	39

Sissejuhatus

Depressioon on haigus, mis haarab enda alla nii inimese psüühika kui ka keha [1]. See võib põhjustada uusi tervisehäireid, süvendada varasemaid tervisega seotud probleeme ning takistada igapäevast töötegemist [2]. Depressiooniravi on pikk protsess – minimaalselt kuus kuud, pigem isegi aasta [3] ning seetõttu on tulemuslikuma ravi eesmärgil vaja lähemalt uurida, kuidas inimese geneetiline varieeruvus mõjutab depressiooniravimite kasutamist.

Magistritöö eesmärk on välja selgitada, kuidas geneetiline varieeruvus *CYP2C19* ja *CYP2D6* geenides, mis põhjustab antidepressantide metabolismis oluliste ensüümide ravimilagundamise kiiruse erinevusi, mõjutab ühelt antidepressandilt vahetamist teisele ning üldist antidepressantide kasutust. Täpsemalt uuritakse, millist mõju avaldavad antidepressantide vahetamisele ja kasutamisele geenide *CYP2C19* ja *CYP2D6* poolt kodeeritud ensüümide erinevad metaboliseerimise kiirusgrupid, deletsioonid nendes geenides ning erinevad tausttunnused.

Töö jaguneb kaheks osaks: teoreetiliseks ja praktiliseks. Teoreetilises osas antakse ülevaade depressiooni olemusest ning selle ravimisest. Järgnevalt räägitakse lähemalt antidepressantide farmakogenoomikast, andes ülevaate ka töös uuritud geenidest ning viimaks kirjeldatakse analüüsis kasutatavaid mudeleid.

Praktiline pool koosneb TÜ Eesti Geenivaramu andmete analüüsist, kus käsitletakse geenidonorite ravimireseptide välja ostmisega seotud andmeid. Esimene pool sisaldab kirjeldavat ülevaadet andmetest ning teises pooles on kokkuvõtte mudelitest, mida seoste leidmiseks kasutati.

Töö on vormistatud kasutades tekstitöötlustarkvara Microsoft Word 2016 ning andmete analüüsimiseks kasutati statistikatarkvara R.

Autor tänab töö juhendajaid Krista Fischerit ja Kristi Krebsi heade ideede ja kasulike nõuannete eest.

1 Depressioon

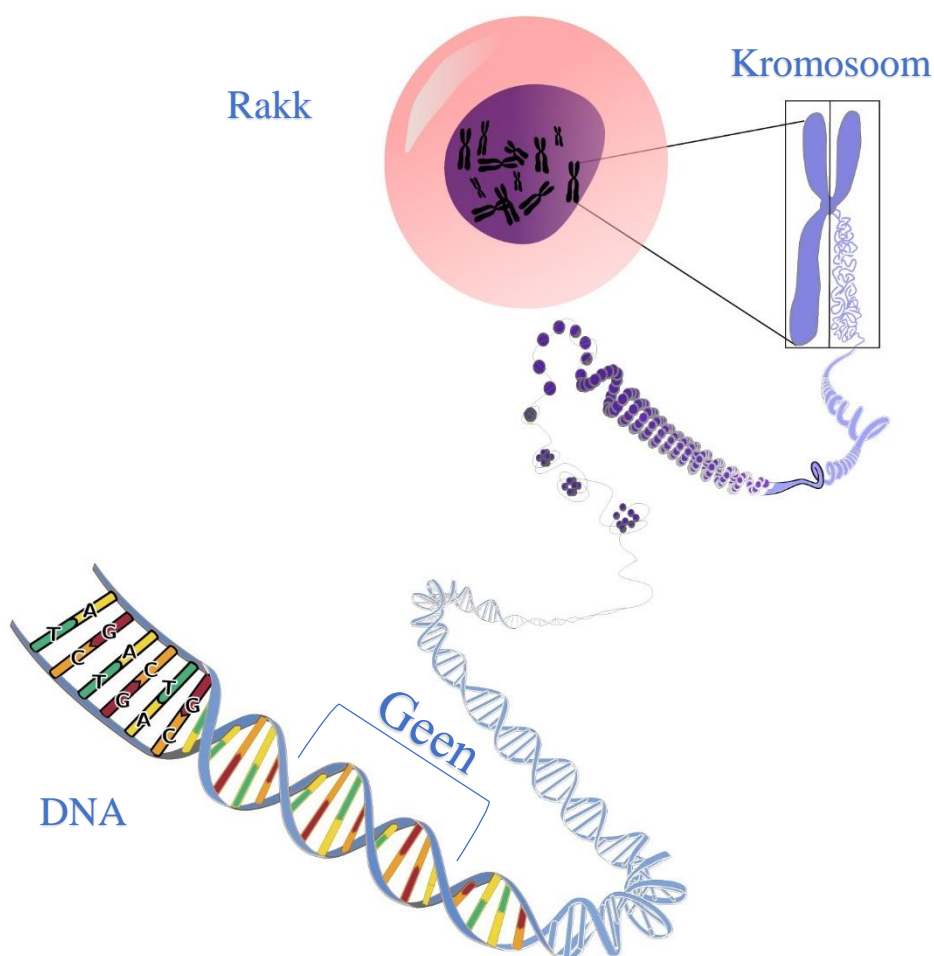
Depressioon on psüühiline haigus, mis vajab samamoodi ravi nagu iga teine haigus [4]. Depressioon võib hakata mõjutama inimese igapäevaelu ning selle tulemusena võib olla raske millestki rõõmu leida [2]. Tõsi on, et sageli ei taha depressioonis olevad inimesed arstiga ühendust võtta, sest nad ei usu, et neil on depressioon, või arvavad, et nende jaoks ei ole abistavaid ravivõimalusi [4]. Depressiooni võivad esile kutsuda erinevad stressi tekitavad sündmused nagu lein, lahkuminek või trauma [2]. Lisaks stressile võivad põhjustajateks olla ka haigused, alkohol, narkootikumid ja inimese iseloom [2].

Depressioonist tervenemine on iga inimese puhul erinev [4]. Tavaliselt kasutakse raviks antidepressante, mida sageli kombineeritakse psühholoogilise nõustamisega [4]. Esimesena välja kirjutatud antidepressant ei pruugi alati anda soovitud tulemusi [3]. Levinud on ka mitmete ravimite kombineerimine [5]. Raviga alustamine on inimeste jaoks sageli raske, esiteks seetõttu, et mõju avaldumine võtab aega [3], teiseks seetõttu, et ravi algusjärgus on sageli enesetunne halvem, kuid see on üldiselt mööduv nähtus [4]. Ravimi tegelik efekt võib ilmuda alles 3-4 nädalat pärast ravi algust [3]. Kuna enesetunne ei parane mõne päeva jooksul, kuid see on piisav aeg kõrvaltoimete tekkimiseks, siis just seetõttu jääbki tihti ravi pooleli [3]. Mõned kõrvaltoimed muutuvad leebemaks, kui ravi on kestnud regulaarselt piisavalt pikka aega [4].

Järgnev lõik on refereeritud allika [3] põhjal. Viimasel kümnendil on inimeste teadlikkus depressiooni olemuse ja antidepressantide tarvitamise osas paranenud. Siiski on üsna levinud eksiarvamus, et antidepressandid tekitavad sõltuvust. Inimeste jaoks on sõltuvus hirmutav, kuid kui antidepressantide kohta kõik ära selgitada, siis muutub hirm ehk nende ees ka väiksemaks. Tähtis roll ravimi toime saavutamiseks ongi inimesel endal, kes peab olema õigesti meelestatud ning ravi pooldama.

2 Geneetika

Järgmises lõigus on kasutatud materjali A. Heinaru õpikust „Geneetika” [6]. DNA on polümeer, mis kannab endas pärilikkuse infot, ning on koos DNA-ga seondunud valkudega koondunud kromosoomidesse. Geen on DNA segment ning pärilikkuse ühik, mis asub kromosoomis. Geeni erivormid on alleelid, mis määravad genotüübi. Genotüüp hoiab endas organismi geneetilist struktuuri ehk geenilookuste alleelset koosseisu. Tihedalt aheldunud alleelide kombinatsioon kromosoomis on haplotüüp, mis pärandub järglastele tervikuna. Inimese fenotüübi moodustavad organismi vaadeldavad tunnused, mis määratakse genotüübi ja ka keskkonnategurite poolt. Mitmekesisuse tagavad muutused DNA järjestuses. Deletsioon on muutuse tüüp, mille tulemusena on osa kromosoomi segmendist puudu. Geenide kaudu antakse edasi infot valkude moodustamiseks, mis on olulised kõikides bioloogilistes protsessides. Valk, mis kiirendab kindlat biokeemilist reaktsiooni, on ensüüm. Joonisel 1 on kujutatud rakk, DNA, kromosoom ja geen.



Joonis 1. Rakk, DNA, kromosoom ja geen [7]

Käesolev lõik ning järgnev lõik on refereeritud allika [8] põhjal. Farmakogenoomika on teadusharu, mis uurib geenide mõju inimese ravimivastusele. Farmakogenoomika on kombinatsioon farmakoloogiast ehk ravimiteadusest ning genoomikast ehk gene ja nende funktsioone hõlmavast teadusest. Farmakogenoomika eesmärk on luua efektiivseid ja turvalisi ravimeid ning selgitada välja optimaalsed ravimidooside suurused, mis sobiksid inimese geneetilise taustaga.

Tänapäeval toodetakse ravimeid keskmisele inimesele, eesmärgiga, et need sobiksid kõigile. Tegelikult mõjuvad ravimid erinevatele inimestele tihti erinevalt. Teadlased üritavad järjest rohkem kindlaks teha, kuidas pärilikud iseärasused mõjutavad organismi vastust ravimitele. Soovitakse ära kasutada geneetilisi erinevusi, hinnata, kas ravimil võib konkreetse inimese jaoks olla positiivne efekt ning mil moel oleks võimalik ennetada ravimi negatiivseid kõrvalmõjusid.

Tsütokroom P450 on nn valkude superperekonid [9]. Tsütokroomi P450 geenide poolt toodetud ensüümid on seotud peamiselt molekulide ja kemikaalide sünteesi ja metabolismiga [10]. Need ensüümid on tähtsal kohal ka mitmete ravimite metabolismis [11]. Kuigi selles klassis on üle 50 ensüümi, siis kuus neist metaboliseerivad üle 90% ravimitest [11]. Iga tsütokroomi P450 geen sisaldab tähistust *CYP*, millele on lisatud kindel grupinumber, alamgrupi täht ning number selles alamgrupis [10]. Näiteks tsütokroomi P450 geen tähisega 1, mis kuulub gruppi 27, alamgruppi A, omab nimetust *CYP27A1* [10]. Peamiselt on P450 ensüüme leitud maksarakkudest, kuid need asuvad kehas ka mujal [10]. Variatsioonid ehk polümorfismid neis geenides võivad mõjutada ensüümide tegevust, põhjustades erinevat metabolismi fenotüüpi, mille tulemusena metaboliseeritakse ravimeid sõltuvalt geenist ja variatsioonist erineva kiirusega [10].

Ravimi lagundamise ehk metabolismi fenotüübi tasemel on võimalik inimesed jaotada nelja *CYP* ensüümi aktiivsusrühki ehk kiirusrühki: aeglased metaboliseerijad, keskmised metaboliseerijad, tavapärased ehk normaalsed ja kiired [12]. Üldjuhul on aeglastel metaboliseerijatel veres ravimi kõrgem kontsentratsioon, mis on tingitud ravimi aeglasest lagundamisest selle ensüümi raja poolt [12]. Seetõttu võivad need inimesed kõrvaltoimete riski vältimiseks vajada madalamat ravimidoosi [12]. Kiirete metaboliseerijate korral on vajalik suurem doos, et ravimi efekt tekiks [12]. Psühhotroopsete ravimite metabolismis on tähtsal kohal maksaensüümid *CYP2D6* ja *CYP2C19* [5]. Need ensüümid kuuluvad eespool mainitud kuue ensüümi hulka, mis metaboliseerivad üle 90% ravimitest [11].

Lõik põhineb materjalil [13], kui pole teisiti kirjas. Kirjeldamaks ADME-geenide (geenid, mis on seotud ravimi absorptsiooni, jaotuse ning metabolismi ja eritusega) varieeruvust, on loodud erinevaid nomenklatuurisüsteeme. Kõige levinum süsteem on nn „star“ (*) süsteem, mis loodi 1990. aastatel ning on laialt levinud. Star-alleeli nomenkatuur tekkis, kui sooviti standardiseerida geneetiliste polümorfismide annotatsioone tsütokroomi *P450* geenide puhul [14]. Iga star-alleel on defineeritud haplotüübi tasandil spetsiifilise kombinatsiooniga ühenukleotiidsetest polümorfismidest ja/või geneetilistest variantidest *CYP2D6* või *CYP2C19* geenilookustest. Alleel *1 on määratud referentsjärjestuseks, millega polümorfseid kohti võrreldakse [14]. See on tavaliselt esimene kirjeldatud järjestus, mis kodeerib funktsionaalse valgu [14].

Lõigus on kasutatud materjali [15]. Varasemalt on uuritud CYP ensüümide ja antidepressantide vahel olevaid erinevaid seoseid. Näiteks on leitud, et aeglased metaboliseerijad puutusid rohkem kokku ebasoodsate efektidega võrreldes kiirete metaboliseerijatega. Näiteks oli aeglastel rohkem kõrvalmõjusid, ning samuti oli kaalutõus suurem. Lisaks on näidatud, et aeglased metaboliseerijad katkestasid võrreldes normaalsete metaboliseerijatega *estsitalopraami/tsitalopraami* kasutamise sagedamini.

Käesolev lõik põhineb allikal [16]. Kliiniliste tõendite ja kirjanduse põhjal on mõnede depressiooniravimite ja nende dooside määramiseks koostatud üldiseid soovitusi erinevat tüüpi ensüümide CYP2C19 ja CYP2D6 aktiivsusgruppide jaoks. Näiteks soovitatakse CYP2D6 aeglaste metaboliseerijate grupi korral tritsüklilisi antidepressante kõrvalmõjude tõttu vältida ja kasutamise korral võiks alustamiskoos olla poole väiksem soovitatust. Tavapärase ravimilagundamise korral soovitatakse tavalist algdoosi. Kiirete metaboliseerijate korral ei soovitata samuti tritsüklilisi ravimeid, kuna need ei ole nii efektiivsed. Kuid kui siiski on ette nähtud kasutamine, siis soovitatavalt võiks algdoos olla tavapärasest doosist suurem. CYP2C19 korral on soovitusi antud tertsiarsete amiinide korral, näiteks võiks aeglaste ja kiirete metaboliseerijate korral seda tüüpi ravimeid vältida, kuna esineb võimalus, et ei saavutata optimaalset ravimivastust.

3 Analüüsimeetodid

3.1 Logistiline regressioonimudel

Antud alapeatükk on refereeritud kursuse „Andmeanalüüs II“ konsekti põhjal [17].

Olgu uurimise all funktsioontunnus Y , millel on ainult kaks võimalikku väärtust 0 ja 1, tihti tähistamaks kategooriad „jah“ ja „ei“ vastavalt teatud sündmuse esinemisele. Olgu tähistatud $P(Y = 1) = \pi$ ja $P(Y = 0) = 1 - \pi$.

Sellisel juhul on tegu Bernoulli jaotusega, $Y \sim B(1, \pi)$. Sellise jaotuse korral on keskväärts $EY = \pi$ ning dispersioon $DY = \pi(1 - \pi)$, mis tähendab seda, et mudel keskväärtusele annab hinnangu sündmuse toimumise tõenäosusele. Huviobjektiks on seos valitud seletavate tunnuste ning uuritava tunnuse esinemise tõenäosuse π vahel.

Uuritava tunnuse jaoks kasutatakse *logit* seosefunktsiooni järgneval kujul,

$$\eta = \text{logit}(\pi) = \ln \frac{\pi}{1 - \pi},$$

kus $\frac{\pi}{1 - \pi}$ on sündmuse esinemise šanss.

Logit seosefunktsioonist saab välja kirjutada sündmuse esinemise tõenäosuse, mis avaldub kujul:

$$\pi = \frac{e^\eta}{1 + e^\eta}.$$

Logistilise mudeli abil hinnatakse šansi logaritmi

$$\ln \frac{\pi}{1 - \pi} = \beta_0 + \beta_1 x_1 + \dots + \beta_k x_k,$$

kus $\pi = P(Y = 1)$ tähistab sündmuse esinemise tõenäosust ning x_i ($i = 1, \dots, k$) on seletavad tunnused ja β_i ($i = 0, \dots, k$) tähistavad tundmatuid parameetreid. Parameetrid hinnatakse suurima tõepära meetodil.

Siit järeldub, et kui tunnus x_i ($i = 1, \dots, k$) suureneb ühe ühiku võrra, siis sündmuse esinemise šanss suureneb e^{β_i} korda. Seega on parameeter β_i tõlgendatav kui tunnuse x_i ühele ühikule vastav logaritmitud šansisuhe.

3.2 Loendusandmete mudelid

Tihti on vaja modelleerida sündmuste arvu, mis tähendab, et tunnuse väärtuseks on täisarvud 0, 1, 2, ..., näiteks liiklusõnnetuste arv või arstivisiitide arv [18]. Siis öeldakse, et tegemist on loendusandmetega [18]. Käesolevas töös kasutatakse kahte sorti mudeleid loendusandmete jaoks, Poissoni regressioonimudelit ning negatiivset binoommudelit.

3.2.1 Poissoni regressioonimudel

Antud alapeatükis on kasutatud allikaid [19] ja [20], kui ei ole teisiti kirjas.

Öeldakse, et diskreetne juhuslik suurus Y on Poissoni jaotusega, kui tema tõenäosusfunktsioon avaldub kujul

$$P(Y = y) = p(y, \mu) = \frac{\mu^y}{y!} e^{-\mu},$$

kus $y = \{0, 1, 2, \dots\}$ ja $\mu > 0$ on Poissoni jaotuse parameeter. Poissoni jaotuse korral on keskväärtus ja dispersioon võrdsed, $EY = DY = \mu$.

Käesolev lõik põhineb materjalil [17]. Poissoni jaotusega üldistatud lineaarse mudeli hindamisel eeldatakse enamasti logaritmilist seosefunktsiooni

$$\eta = \ln(\mu) = \beta_0 + \beta_1 x_1 + \dots + \beta_k x_k,$$

kus β_0, \dots, β_k on mudeli tundmatud parameetrid ning x_1, \dots, x_k on argumenttunnused ja μ on uuritava tunnuse keskväärtus. Seega saab öelda, et kui x_i suureneb ühe ühiku võrra, suureneb funktsioontunnuse Y keskväärtus (μ) e^{β_i} korda.

Kui soovitakse hinnata sündmuste intensiivsust ajas st suurust $\frac{\mu}{t}$, siis on vaja mudelisse lisada ka nn lisatunnus *offset*, mis võtab arvesse ajaperioodi. Sellisel juhul kehtib

$$\ln\left(\frac{\mu}{t}\right) = \beta_0 + \beta_1 x_1 + \dots + \beta_k x_k$$

ehk

$$\eta = \ln(\mu) = \beta_0 + \beta_1 x_1 + \dots + \beta_k x_k + \ln(t),$$

kus $\ln(t)$ on *offset*.

Poissoni regressioonimudelil kasutatakse tavaliselt parameetrite hindamiseks suurima tõepära meetodit. Poissoni tõenäosusjaotus on eksponentsiaalse jaotuse pere liige ning võttes aluseks eksponentsiaalse pere kuju, saab Poissoni tõepärafunktsiooni välja kirjutada järgnevalt

$$L(y, \mu) = \prod_{i=1}^n \exp\{y_i \ln(\mu_i) - \mu_i - \ln(y_i!)\}.$$

Poissoni regressiooni korral on tihti probleemiks ülehajuvus ehk $EY = \mu$, aga $DY > \mu$. See tähendab seda, et dispersiooni hinnang on liiga suur ning parameetrite korrektse hindamise osas võib probleeme tekkida [17]. Korrektse mudeli korral oleks jagatise $\frac{D \text{ ehk hälbumus jääk}}{df \text{ ehk vabadusastmed}}$ väärtus ligikaudu üks. Kui jagatis on aga suurem kui kaks, siis on ülehajuvuse probleem selge ja kindlasti tuleb see mudelis arvesse võtta.

3.2.2 Negatiivne binoommudel

Peatükis on kasutatud materjale kursusel „Üldistatud lineaarsed mudelid“ [20], kui ei ole teisiti märgitud.

Suure ülehajuvuse korral on üheks lahenduseks kasutada mudelit, mis eeldab funktsioontunnuse negatiivset binoomjaotust.

Tõenäosusfunktsioon negatiivse binoomjaotusega juhusliku suuruse ($Y \sim NB(k, \pi)$) korral omab kuju

$$P(Y = y) = p(y; k; \pi) = \frac{\Gamma(k+y)}{y! \Gamma(k)} \pi^k (1 - \pi)^y,$$

kus jaotuse parameetrid $k > 0$ ja $0 < \pi < 1$.

Negatiivne binoomjaotus kirjeldab katsete arvu, mida on vaja, et teatud sündmus toimuks k korral, kusjuures π on sündmuse toimumise tõenäosus ühel korral. Negatiivse binoomjaotuse keskvärtus $EY = \mu = \frac{k(1-\pi)}{\pi}$ ja dispersioon $DY = \frac{k(1-\pi)}{\pi^2} = \mu + \frac{\mu^2}{k}$. Seega näeme, et alati $DY > EY$.

Võrreldes negatiivset binoomjaotust ja Poissoni jaotust, saab öelda, et negatiivse binoomjaotuse korral on nullidel suurem tõenäosus ja parempoolne saba on raskem.

Negatiivse binoommudeli korral kasutatakse sarnaselt Poissoni regressioonimudelile tavaliselt logaritmilist seosefunktsiooni

$$\eta = \ln(\mu) = \beta_0 + \beta_1 x_1 + \dots + \beta_k x_k,$$

kus β_0, \dots, β_k on mudeli tundmatud parameetrid, x_1, \dots, x_k on argumenttunnused ning μ on uuritava tunnuse keskväärtnus. Argumenttunnuse x_i suurenedes ühe ühiku võrra suureneb keskväärtnus (μ) e^{β_i} korda. Ka negatiivse binoommudeli korral kasutatakse *offsetti* samamoodi nagu Poissoni regressioonimudeli korral ning kehtib

$$\eta = \ln(\mu) = \beta_0 + \beta_1 x_1 + \dots + \beta_k x_k + \ln(t),$$

kus $\ln(t)$ on *offset*.

Negatiivne binoomjaotus kuulub samuti eksponentsiaalse jaotuse perre. Negatiivse binoommudeli eelis Poissoni mudeli ees on see, et paindlikum dispersioonistruktuur lubab mudelit hinnata suurema varieeruvusega andmetele. Siiski tuleb tähele panna, et negatiivne binoomimudel ei ole sobilik alahajuvuse korral.

Negatiivse binoommudeli hindamiseks kasutatakse suurima tõepära meetodit, kusjuures log-tõepärafunktsioon on järgnev

$$l(\beta_j; y; \alpha) = \sum_{i=1}^n y_i \ln \left(\frac{\alpha \exp(x_i^T \beta)}{1 + \alpha \exp(x_i^T \beta)} \right) - \frac{1}{\alpha} \ln(1 + \alpha \exp(x_i^T \beta)) + \ln \Gamma \left(y_i + \frac{1}{\alpha} \right) - \ln \Gamma(y_i + 1) - \ln \Gamma \left(\frac{1}{\alpha} \right),$$

kus β_j tähistavad parameetreid [19].

3.3 Elukestusanalüüs

3.3.1 Mõisted

Käesolev lõik põhineb raamatul [21]. Elukestusanalüüsiks nimetatakse andmeanalüüsi meetodikat, kus huvi pakub aeg mingi kindla sündmuse toimumiseni. See sündmus võib olla surm, aga ka näiteks valu kadumine, kõrvaltoime või mõne sümptomi esinemine.

Elukestuseks nimetatakse aega mingist kindlast algmomendist (nt uuringuga liitumine) sündmuse toimumiseni, olgu see tähistatud kui T [22]. Enamasti ei ole T kõigi uuritavate jaoks vaadeldud [22]. Sellisel juhul on vaadeldud nn tsenseerimisaeg C ehk aeg, mil uuritav oli vaatluse all ning kõigi uuritavate puhul vaadeldud $Z = \min(T, C)$ [22]. Tsenseerimise põhjus

võib olla näiteks uuringu lõppemine enne sündmuse toimumist või uuringu katkestamine indiviidi poolt [22]. Tsenseerimise korral on peamiselt tegemist paremalt poolt tsenseerimisega, mis tähendab seda, et uuritav sündmus ei toimunud enne uuringu lõppu [21]. Vasakult poolt tsenseerimisega on tegemist siis, kui uuritav sündmus toimus enne uuringu algust [21].

Antud lõik põhineb materjalil [21]. Elukestusanalüüsis on oluline ka üleelamisfunktsioon $S(t)$, mis on defineeritud kui tõenäosus, et sündmus ei toimu enne aega t . Üleelamisfunktsiooni saab välja kirjutada kui

$$S(t) = P(T \geq t) = 1 - F(t),$$

kus $F(t)$ on juhusliku suuruse T jaotusfunktsioon, mis omakorda on defineeritud kui

$$F(t) = P(T < t) = \int_0^t f(u) du,$$

kus $f(t)$ on juhusliku suuruse T tihedusfunktsioon.

Riskifunktsioon $h(t)$ kirjeldab tinglikku tõenäosust, et sündmus jääb poollõiku $[t, t + \delta t)$ tingimusel $T \geq t$ ja riskifunktsioon avaldub kui

$$h(t) = \lim_{\delta t \rightarrow 0} \left\{ \frac{P(t \leq T < t + \delta t | T \geq t)}{\delta t} \right\} [21].$$

Saab näidata, et riskifunktsioon avaldub üleelamisfunktsiooni kaudu kui

$$h(t) = -\frac{d[\ln(S(t))]}{dt} = \frac{f(t)}{S(t)} [22].$$

3.3.2 Coxi võrdeliste riskide mudel

Järgnev alapeatükk põhineb raamatul [22], kui ei ole teisiti kirjas.

Coxi võrdeliste riskide mudel on üks populaarsemaid meetodeid elukestusanalüüsis. Võrdeliste riskide mudeli abil on võimalik uurida seda, kuidas grupi patsientide elukestus sõltub ühe või mitme seletava tunnuse väärtustest [21]. See mudel kasutab riskifunktsiooni $h(t|\mathbf{X})$, kus \mathbf{X} on argumenttunnuste vektor.

Coxi võrdeliste riskide mudel pannakse tihti kirja kasutades riskifunktsiooni järgnevalt

$$h(t|\mathbf{X}) = h_0(t) \exp\{\boldsymbol{\beta}^T \mathbf{X}\} = h_0(t) \exp\{\beta_1 X_1 + \dots + \beta_k X_k\},$$

kus $\boldsymbol{\beta}$ on parameetrite vektor ja $h_0(t)$ on baasriskifunktsioon.

Argumenttunnuste vektori väärtuste \mathbf{x}_1 ja \mathbf{x}_2 korral kehtib

$$\frac{h(t|\mathbf{x}_1)}{h(t|\mathbf{x}_2)} = \frac{h_0(t) \exp\{\boldsymbol{\beta}^T \mathbf{x}_1\}}{h_0(t) \exp\{\boldsymbol{\beta}^T \mathbf{x}_2\}} = \exp\{\boldsymbol{\beta}^T (\mathbf{x}_1 - \mathbf{x}_2)\},$$

mis tähendab, et riskisuhe ei sõltu ajast t ega baasriskist $h_0(t)$.

Parameetrite hindamiseks kasutatakse osalise tõepära meetodit. On näidatud, et tõepärafunktsioon avaldub kujul

$$L(\boldsymbol{\beta}) = \prod_{j=1}^r \frac{\exp(\boldsymbol{\beta}^T \mathbf{x}_j)}{\sum_{\mathbf{l} \in R(t_j)} \exp(\boldsymbol{\beta}^T \mathbf{x}_l)},$$

kus \mathbf{x}_j on argumenttunnuste vektor indiviidil, kellel esineb sündmus hetkel t_j , $R(t_j)$ on riskigrupp inimestest, kes on hetkel t_j vaatluse all [21]. Kokku vaadeldakse uuritavas andmestikus r erinevat sündmuse toimumise aega $t_1 < t_2 < \dots < t_r$ [21].

3.3.3 Mudelid konkureerivate riskide korral

Järgnevas kahes alapeatükis on kasutatud materjale raamatust [21].

Konkureerivatest riskidest räägitakse olukorras, kus uuritava vaatluse all olek võib lõppeda ühega mitmest erinevast sündmusest, kusjuures iga sündmuse toimumine välistab teiste sündmuste toimumise võimaluse. Sellisel kujul andmeid esineb erinevates valdkondades. Näiteks vähiga seotud kliiniliste uuringute korral võivad huvi pakkuda surmad, mille põhjuseks on erinevad vähitüübid ning sellisel juhul on surm näiteks südameinfarkti või insuldi tõttu võistlevad riskifaktorid.

Lõplikud andmed sisaldavad elukestust T ja lõppsündmuse tüüpi C ning indiviidi i jaoks kirjutatakse (t_i, c_i) , $i = 1, 2, \dots, n$ ja c_i võimalikud väärtused on $0, 1, \dots, m$ ja $c_i = 0$, kui lõppsündmust ei ole vaadeldud ehk vaatlus on tsenseeritud.

Seega kui $c_i = j$ ($j = 1, \dots, m$), siis teame, et ükski sündmus k , kus $k \neq j$ ei olnud esinenud enne aega t_i ja nad oleksid võinud esineda pärast t_i kui põhjus j ei oleks esinenud.

Põhjuse-spetsiifiline (*cause-specific*) riskifunktsioon $h_j(t)$ põhjusele j , $j = 1, \dots, m$, on defineeritud kui

$$h_j(t) = \lim_{\delta t \rightarrow 0} \left\{ \frac{P(t \leq T \leq t + \delta t, C = j \mid T \geq t)}{\delta t} \right\}. \quad (1)$$

See riskifunktsioon on põhjuse j riskimäär hetkel t kõigi teiste riskide olemasolul tingimusel, et ükski sündmus polnud enne aega toimunud. Riskifunktsioon on esitatav põhjuse-spetsiifilise tihedusfunktsiooni $f_j(t)$ ja üldise üleelamisfunktsiooni $S(t) = P(T \geq t)$ kaudu järgnevalt

$$h_j(t) = \frac{f_j(t)}{S(t)}.$$

Ainult üks võimalikust m põhjusest võib viia huvipakkuva sündmuseni, seega kehtib

$$P(t \leq T \leq t + \delta t \mid T \geq t) = \sum_{j=1}^m P(t \leq T \leq t + \delta t, C = j \mid T \geq t).$$

Üldine riskifunktsioon avaldub aga kujul

$$h(t) = \sum_{j=1}^m h_j(t).$$

3.3.4 Fine ja Gray konkureerivate riskide mudel

Järgnevalt on kirjeldatud Fine ja Gray konkureerivate riskide mudelit.

Põhjuse-spetsiifiline kumulatiivne alamjaotuse funktsioon (*cause-specific cumulative incidence function, subdistribution function*) avaldub põhjuse j jaoks $F_j(t) = P(T < t, C = j)$.

Alamjaotuse riskifunktsioon põhjuse j jaoks on kujul

$$\lambda_j(t) = -\frac{d}{dt} \log\{1 - F_j(t)\} = \frac{1}{1 - F_j(t)} \cdot \frac{dF_j(t)}{dt}.$$

$1 - F_j(t)$ on tõenäosus, et indiviid elas üle aja t või suri mõne teise põhjuse kui j pärast ning kehtib

$$\frac{dF_j(t)}{dt} = \lim_{\delta t \rightarrow 0} \left\{ \frac{F_j(t + \delta t) - F_j(t)}{\delta t} \right\}.$$

Seega järeldub, et alamjaotuse riskifunktsioon $\lambda_j(t)$ on kujul

$$\lambda_j(t) = \lim_{\delta t \rightarrow 0} \left\{ \frac{P(t \leq T \leq t + \delta t, C=j | T \geq t \text{ või } \{T \leq t \text{ ja } C \neq j\})}{\delta t} \right\}.$$

See on sündmuse j esinemise kordaja hetkel t , juhul kui indiviidil ei ole eelnevalt esinenud sündmust j . Kuna selle riskifunktsiooni definitsioonis toodud tingliku tõenäosuse tingimus sisaldab neid, kellel on varem toimunud mõni teine sündmus kui j enne aega t , siis alamjaotuse riskifunktsioon erineb põhjus-spetsiifilisest riskifunktsioonist (1) nii definitsiooni kui interpretatsiooni poolest.

Et modelleerida põhjus-spetsiifilist alamjaotuse funktsiooni, siis eeldatakse Coxi regressiooni mudelit põhjuse j alamriski funktsiooni jaoks. Risk põhjuse j jaoks ajahetkel t indiviidi i jaoks n indiviidist on

$$\lambda_{ij}(t) = \exp(\boldsymbol{\beta}_j^T \mathbf{x}_i) \lambda_{0j}(t),$$

kus $\lambda_{0j}(t)$ on alamjaotuse baasväärtuse riskifunktsioon põhjuse j jaoks, \mathbf{x}_i on argumenttunnuste vektor indiviidi i jaoks ja $\boldsymbol{\beta}_j$ sisaldab argumentide hinnanguid põhjuse j jaoks.

Parameetrite hinnangud leitakse ka sellisel juhul osalise tõepära meetodil, kusjuures osalise tõepära funktsioon põhjuse j jaoks on

$$\prod_{h=1}^{r_j} \frac{\exp(\boldsymbol{\beta}_j^T \mathbf{x}_h)}{\sum_{l \in R(t_h)} w_{hl} \exp(\boldsymbol{\beta}_j^T \mathbf{x}_l)}, \quad (2)$$

kus korrutis on üle r_j indiviidi, kes surevad põhjusesse j aegadel $t_1 < t_2 < \dots < t_{r_j}$ ja \mathbf{x}_h on argumenttunnuste vektor indiviidi jaoks, kellel esineb sündmus j hetkel t_h , $h = 1, 2, \dots, r_j$. Riskigrupp $R(t_h)$ on hulk neist, kes pole sündmust kogunud enne aega t_h , ja neist, kes on kogunud konkureerivat riski ajaks t_h . Esimesse hulka kuuluvate indiviidide elukestvus on suurem või võrdne t_h ning teisse hulka kuuluvatel indiviididel on elukestvus väiksem või võrdne t_h . See riskigrupp ei ole otseselt interpreteeritav, kuna indiviid, kellel on enne aega t esinenud muu sündmus kui j , ei ole enam vaatluse all ajal t , kuid ikkagi kuulub ta selles mudelis riskigrupi.

Kaalud valemis (2) on defineeritud kui

$$w_{hl} = \frac{\hat{S}_c(t_h)}{\hat{S}_c(\min\{t_h, t_l\})},$$

kus $\hat{S}_c(t_h)$ on Kaplan-Meieri hinnang nn tsenseerimisaegade üleelamisfunktsioonile.

3.4 Lõplike mudelite leidmine

Käesolev ning järgnev lõik põhinevad materjalil [17], kui ei ole teisiti kirjas. Parimate mudelite leidmiseks on mitmeid meetodeid. Automaatsed valikute põhimõtted jagunevad enamjaolt kolmeks: ettepoole ehk kasvav valik, tahapoole ehk kahanev valik ja segavalik. Kasvava valiku korral hakatakse mudelisse kindlate kriteeriumite alusel tunnuseid juurde lisama. Kahaneva valiku korral hakatakse vastupidiselt eelmisele meetodile tunnuseid järjest ära võtma. Segavaliku korral kasutatakse mõlemat strateegiat segamini.

AIC ehk Akaike informatsioonikriteerium on üks mudeli headuse näitajatest. See näitaja saadakse hälbumusele teatud parandusliikme lisamisel. Mida väiksem on Akaike informatsioonikriteeriumi väärtus, seda parem on mudel. Kriteeriumi valem avaldub järgnevalt:

$$AIC = -2\ln L + 2p,$$

kus L tähistab suurima tõepära funktsiooni ning p on mudelis hinnatavate parameetrite arv [20].

Käesolevas töös kasutatakse lõplike mudelite leidmiseks tahapoole ehk kahanevat valikut. Kriteerium, mille järgi tunnuseid välja hakatakse jätma, on AIC. Igal sammul vaadatakse, millise tunnuse välja jätmisel oleks uue mudeli AIC kõige parem ehk kõige väiksem. Tunnus, mille korral see tingimus kehtib, jäetakse välja. Sama põhimõtet kasutatakse seni, kuni ühegi tunnuse eemaldamisel AIC enam ei vähene.

4 Andmete analüüs

Käesolevas peatükis tehtud analüüsid kasutati olulisuse nivood $\alpha = 0,05$. Selguse huvides olgu täpsustatud, et edasises on erinevatest ravimitest või antidepressantidest rääkides mõeldud ravimi toimeaineid.

4.1 Analüüsimeetodite rakendamine tarkvara R abil

Logistilise regressiooni ning Poissoni regressiooni jaoks kasutati R paketi *stats* funktsiooni *glm* [23]. Selle funktsiooni nõutavaks parameetrik on nn mudeli valem ning üldistatud lineaarse mudeli hindamisel tuleks lisada nn perekonna tingimus *family* [23]. Valemi vasak pool peab sisaldama uuritavat tunnust, parem pool argumenttunnuseid ning pooled eraldatakse märgiga „~“ [23]. Logistilise regressiooni korral on vaja lisada tingimus *family*=“*binomial*“ ning Poissoni regressiooni korral *family*=“*poisson*“ [24]. Negatiivse binoomimudeli jaoks kasutati R paketi *MASS* funktsiooni *glm.nb*, kuhu sarnaselt *glm* funktsiooniga tuleb lisada samal kujul valem [25]. Loendusandmete mudelite korral on võimalik lisada ka *offset* tunnus, mis määrab ära selle, millises ajavahemikus sündmusi vaadeldi [20]. Coxi võrdeliste riskide mudeli hindamiseks kasutati R paketi *survival* funktsiooni *coxph*, mille nõutavaks parameetrik on valem kujul *Surv(aeg, sündmus) ~ arg1 + arg2 + ... + arg3*, kus märgist „~“ vasakul pool on tunnus, kus *aeg* tähistab vaadeldavat aega ning *sündmus* seda, kas uuritavat sündmust vaadeldi või mitte (sündmuse väärtus on üks, kui vaadeldi ning null, kui ei vaadeldud) [26]. Coxi mudeli võrdeliste riskide eelduse kontrollimise jaoks on R paketi *survival* funktsioon *cox.zph*, millele tuleb ette anda mudel, mida kontrollida soovitakse [26]. Konkureerivate riskide mudelit hinnati R paketi *cmprsk* oleva funktsiooni *crr* abil [27]. Funktsioonile *crr* tuleb ette anda sündmuse aja tunnus, sündmuse tüübi tunnus ning maatriks argumenttunnustega [27]. Sündmuse tüübi väärtus on üks, kui uuritav sündmus toimus vaadeldava aja jooksul, väärtus on null, kui vaatlus on tsenseeritud, ning väärtus on kaks, kolm jne siis, kui tegu on mõne konkreetse konkureeriva sündmusega ning aja tunnus sisaldab aega, mil indiviidi vaadeldi mingi sündmuse toimumiseni [27].

4.2 Ülevaade andmetest

Töös kasutati Tartu Ülikooli Eesti Geenivaramu geenidoonorite terviseandmeid. Selleks, et analüüsida geenidoonorite retseptiandmeid, kasutati kahte andmestikku: otse Eesti Haigekassast saadud andmestikku ning koondandmestikku, kuhu oli lisaks Haigekassa infole

lisatud ka informatsioon epikriisidest, geenivaramu küsimustikust, Põhja-Eesti Regionaalhaiglast ning Tartu Ülikooli Kliinikumist. Eespool mainitud kahe andmestiku põhjal moodustati unikaalsete ridadega ühine andmestik geenidoonorite retseptiandmete informatsiooniga ainult antidepressantide ravimiklassi kohta (ravimi ATC – anatoomilis-terapeutilis-keemiline – kood N06A). Geenidoonorite retseptiandmed pärinesid ajavahemikust 2004-2016. Töö tegemise hetkel oli viimane linkimine andmebaasidega toimunud 18. aprillil 2016.

Esialgne andmestik koosnes 183 724 reast ning unikaalseid indiviide, kellele arst mõne analüüsitava depressiooniravimi retsepti välja kirjutas, oli 10 639. Järgmise sammuna jäeti välja realiseerimata retseptid, kuna eesmärk oli jälgida seda, kuidas erinevad tegurid mõjutavad depressiooniravimite tarbimist ning vahetamist. Realiseeritud retseptidega andmestik koosnes 148 652 reast ning unikaalseid indiviide oli 10 457. Seega 182 inimest olid sellised, kellele küll retsept kirjutati, aga kes ravimi tarbimist ei alustanud.

Andmestik sisaldas patsiendi identifitseerimiskoodi, välja ostetud ravimi nimetust ja ravimi välja ostmise kuupäeva. Identifitseerimiskood oli igal patsiendil unikaalne.

Geenidoonorite retseptiandmetega puhastatud andmestik ühendati indiviidide geneetilise informatsiooniga. Antud andmestik oli juba varem TÜ Eesti Geenivaramu teadlaste poolt koostatud, kajastades informatsiooni ravimi vastuses oluliste geenide metabolismi fenotüüpide kohta [28]. Kuna käesolevas töös on peamine eesmärk vaadelda just nende geenide varieeruvust, mis on olulised antidepressantide metabolismis, siis keskenduti andmetes *CYP2C19* ja *CYP2D6* geeni varieeruvusele. Täpsemalt oli võimalik igale indiviidile lisada mõlema ensüümi aktiivsusgrupp ja deletsiooniga seotud informatsioon. Deletsioonitunnus oli 0/1 tunnus, kus üks tähistas, et indiviidil oli geenis deletsioon ja null, et ei olnud.

Eelnevale andmestikule lisati juurde ka andmestik järgnevate tausttunnustega: sugu, sünniaasta, surmakuupäev, pikkus, kaal ning kehamassiindeks.

Lisaks olid kättesaadavad ka andmed geenidoonorite suguluse kohta (vastav info oli saadud geeniandmete põhjal). Selleks, et töö tulemused oleksid täpsemad, võeti sugulust arvesse. Ideeks oli, et suguluses olevate indiviidide seast ei jääks üle ühe indiviidi lõplikku andmestikku. Tulemus saavutati nii, et hakati järjest sugulaspäri läbi vaatama ning kui vähemalt üks paarilistest esines edaspidiste ridade hulgas, siis esialgne rida kustutati ära. Nii kontrolliti läbi kõik read, kuni lõpuks jäi alles andmestik suguluses isikute paaridega. Viimase

sammuna jäeti üks igast paarist alles ja teine eemaldati andmestikust. Selleks, et üldist juhtude arvu suurendada, eemaldati andmestikust eelisjärjekorras kontrollid. Selle tulemusena indiviidide arv vähenes – selle sammu järel alles jäänud andmestik sisaldas andmeid 9446 indiviidi kohta.

Et osade indiviidide korral oli mõne tausttunnuse või aktiivsusgrupi väärtus puudu, jäeti lõplikku andmestikku alles need individid, kellel olid kõigi vajalike tunnuste väärtused olemas. Lõplik andmestik sisaldas 5709 rida. Peamiselt langesid individid välja seetõttu, et neil ei olnud võimalik täpselt aktiivsusgruppi määrata.

Edasiseks analüüsiks loodi andmestikku juurde mõned tunnused. Esiteks lisati retseptide koguarvu tunnus ning erinevate ravimite tunnus. Nende tunnuste jaoks loeti iga indiviidi puhul vastavalt kokku, mitu retsepti oli välja ostetud ning mitme erineva ravimi retsepte oli realiseeritud.

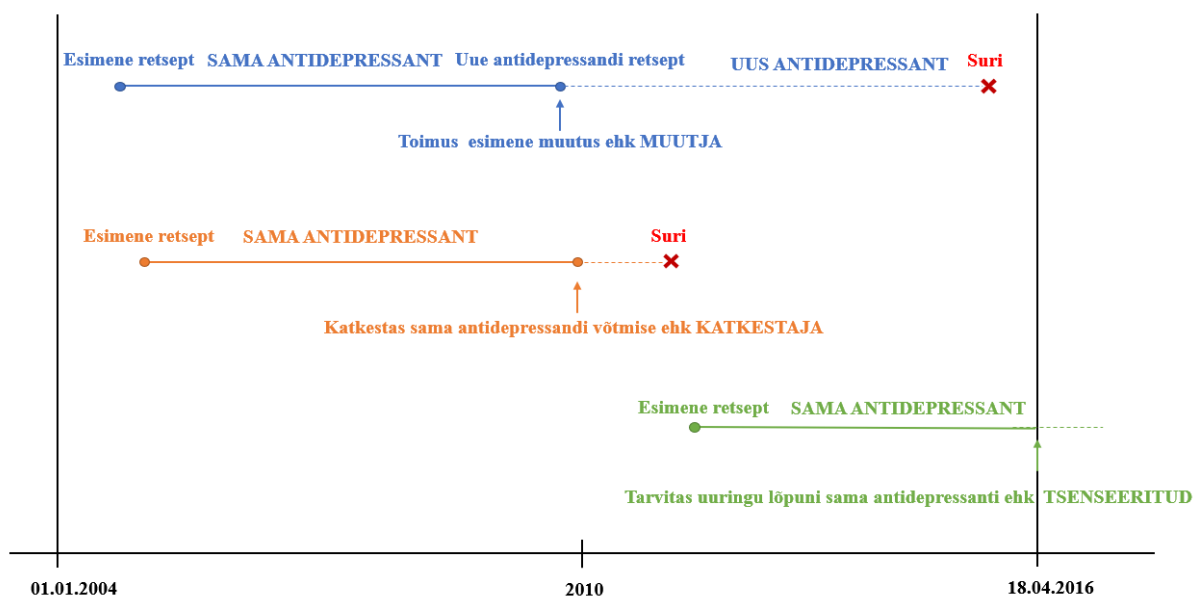
Teiseks loodi ravimi vahetamise tunnus, mis oli 0/1 tunnus, kus üks tähistas seda, et indiid muutis mingi ravimi pealt teisele, ning null seda, et muutust ei toimunud.

Analüüsi jaoks oli vaja ka aega, mis esimese vahetuse ehk muutuseni kulus. Muutuse hetke leidmiseks järjestati individide välja ostetud retseptid kuupäevade järgi ning kui ridade võrdluses tuli toimeaine osas muutus sisse ja kuupäev ei olnud sama, registreeriti ravimi muutus. Aeg muutuseni defineeriti kui uue ravimi ostmise kuupäeva ja esimese retsepti välja ostmise kuupäeva vahe.

Aja tunnustena leiti veel aeg ravimi katkestamiseni ja aeg uuringu lõpuni. Aeg katkestamiseni defineeriti kui viimase ostetud retsepti kuupäeva ja esimese ostetud retsepti kuupäeva vahe päevades. Aeg uuringu lõpuni defineeriti kui kuupäeva 18. aprill 2016 ja esimese ostetud retsepti kuupäeva vahe.

Veel lisati vanuse tunnus, mis defineeriti kui indiviidi vanus ravi alustamise hetkel. See leiti nii, et lahutati esimese retsepti välja ostmise aastast maha indiviidi sünniaasta.

Iga indiviidi korral leiti ka jälgimisaeg ehk aeg päevades esimese retsepti ostust kuni uuringu lõpuni. Uuringu lõpuks võeti viimase linkimise kuupäev 18.04.2016 või surmakuupäev, kui indiid enne uuringu lõppu suri.



Joonis 2. Ülevaade võimalikest sündmustest käesoleva uuringu kontekstis

Joonisel 2 on näha, mille põhjal tekivad analüüsid erinevad tähtsal kohal olevad sündmused. Indiviid loetakse sellesse sündmuste gruppi, mis tal kõige varem esineb. Selleks sündmuseks saab olla ravimi vahetamine ehk muutmine – indiviid on muutja, katkestamine – indiviid on katkestaja või siis sama antidepressandi kasutamine uuringu lõpuni – indiviid on tsenseeritud. Surm toimub alati pärast uuritavaid sündmusi, seega antud juhul surma eraldi sündmusena ei käsitleta.

4.3 Kirjeldav analüüs

Analüüsiks sobivatele kriteeriumitele vastavaid indiviide oli lõplikus andmestikus kokku 5709. Leidsid indiviide, kellel oli pikkuse, kaalu ning kehamassiindeksi kohta mitu tulemust, kuna nad olid mitmel visiidil käinud ning ühtluse mõttes võeti selliste inimeste korral arvesse esimesed mõõtmised.

Kokku oli uuritavate seas 1381 meest ja 4328 naist. Ravi alustamise ajal oli noorim geenidoonor 9 aastat vana ning vanim geenidoonor 98. Indiviide, kes olid 54-aastased oli kõige rohkem – kokku 142. Keskmine vanus oli 47 aastat.

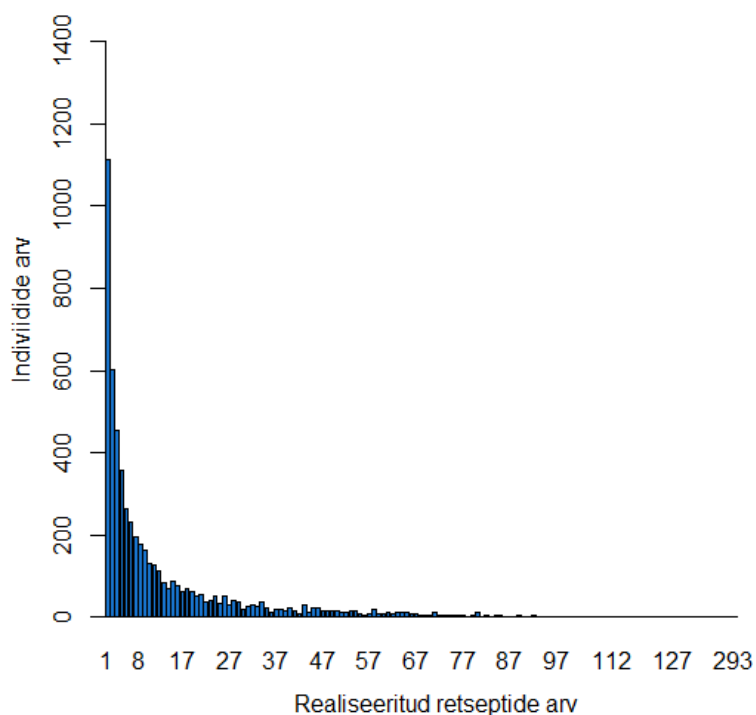
Andmestiku põhjal suri 371 indiviidi enne 18. aprilli 2016, sh 132 meest ning 239 naist.

Tabel 1. Indiviidide pikkuse, kaalu ja kehamassiindeksi andmed

	Näitaja	Tunnus		
		Pikkus	Kaal	Kehamassiindeks
Naised	Miinumum	129	37	15
	Maksimum	197	170	60
	Keskmine	165	72	27
Mehed	Miinumum	142	35	13
	Maksimum	200	165	52
	Keskmine	178	86	27
Kokku	Miinumum	129	35	13
	Maksimum	200	170	60
	Keskmine	168	75	27

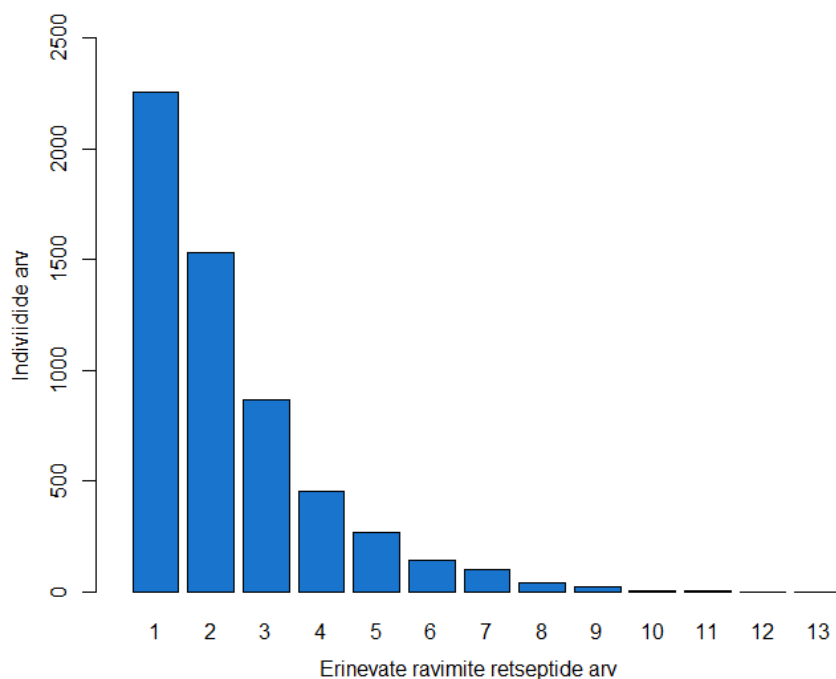
Tabelist 1 selgub, et geenidonorite keskmine pikkus on 168 cm, keskmine kaal 75 kg ning keskmine kehamassiindeks 27.

Vaadates retseptide realiseerimist, selgub, et esineb küllaltki suur varieeruvus. Kõige rohkem oli neid indiviide, kes realiseerisid ainult ühe antidepressandi retsepti ning neid oli kokku 1113. Järgmisel kohal oli 601 indiviidi, kes ostsid kaks retsepti. Kolmandal kohal oli 456 indiviidi kolme välja ostetud retseptiga. Joonisel 3 (lk 24) on näha ka tulpdiagramm välja ostetud retseptide jaotuse kohta. Maksimaalne realiseeritud retseptide arv oli 293 ning nii palju retsepte ostis välja ainult üks indiviid. Suurusjärgult järgmised välja ostetud retseptide arvud olid 186 ja 171 ning neid esindab samamoodi vaid üks geenidonor mõlemas grupis. Indiviide, kes ostsid välja üle 100 retsepti, oli kokku 58.



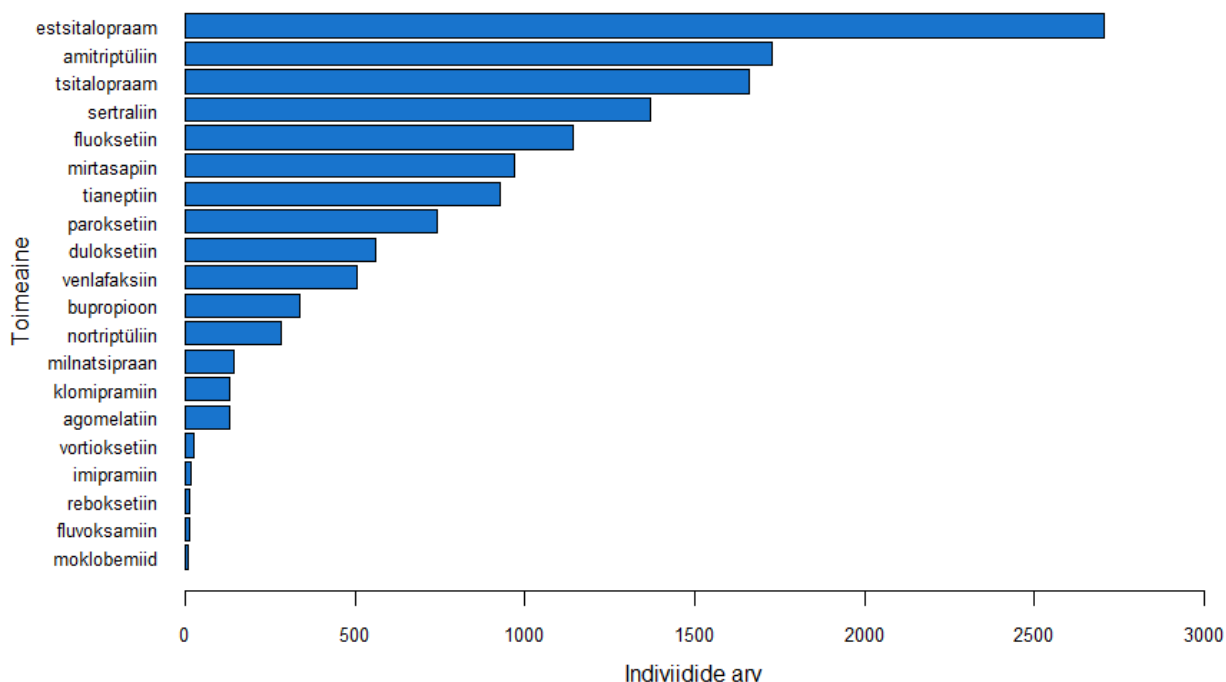
Joonis 3. Välja ostetud retseptide arvud indiviidi kohta

Neid indiviide, kes ostsid välja ainult ühe kindla ravimi retsepte, oli kokku 2257. Kahte erinevat ravimit ostis välja 1529 ning kolme erinevat 867 indiviidi. Leidus üks indiviid, kes ostis välja 13 erinevat ravimit ning kaks inimest, kes ostsid 12 erinevat ravimit. Täpsem erinevate välja ostetud ravimite retseptide jaotus on näha joonisel 4.



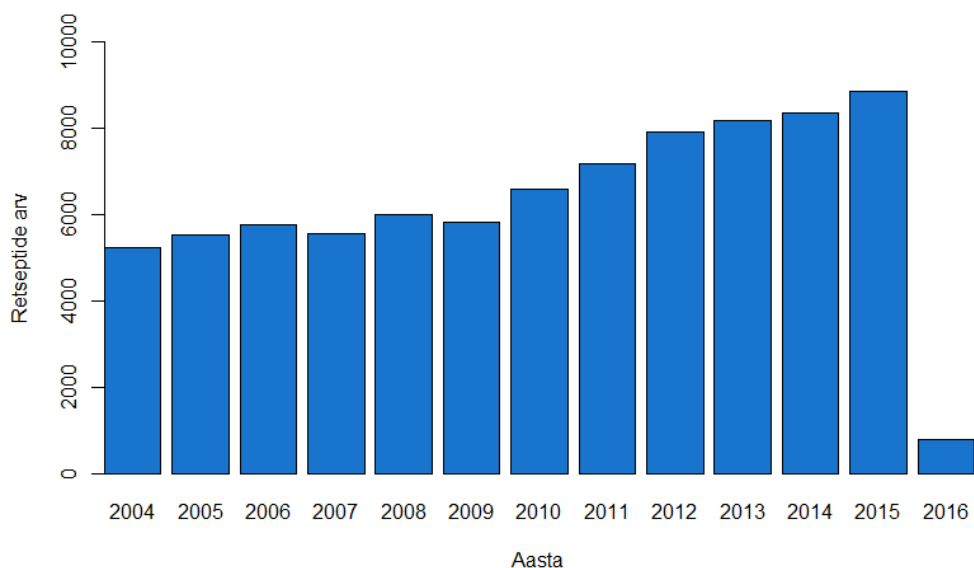
Joonis 4. Erinevate ravimite retseptide arvud indiviidi kohta

Andmetest tuli välja, et inividid kasutasid 20 erinevat ravimit. Joonisel 5 on naha, kuidas erinevate toimeainetega depressiooniravimeid vija osteti. Ulekaalukalt on koge populaarsem toimeaine *estsitalopraam*, mida sisaldavaid ravimeid ostis 2704 indiviidi. Sellele jargnesid *amitriptuulin* ja *tsitalopraam* vastavalt 1726 ja 1660 indiviidiga. Ainult uheksa inimest ostis ravimeid toimeainega *moklobemiid*.



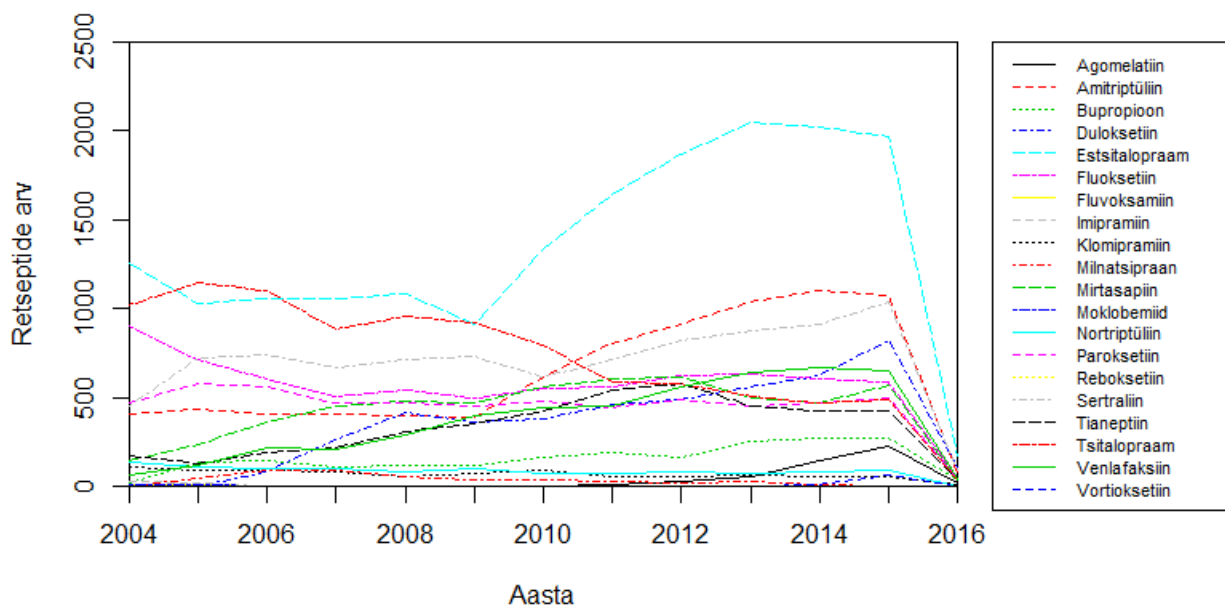
Joonis 5. Indiviidide jaotus erinevate toimeainete korral

Aastate loikes on naha, et koge rohkem retsepte osteti vija aastal 2015, kui retseptide koguarv oli 8848. Ka aastatel 2013 ja 2014 osteti vija ule 8000 retsepti, teistel aastatel olid koguarvud vaiksemad. Joonisel 6 (lk 26) on naha, kuidas aastate jooksul vija ostetud retseptide arv on muutunud. Andmed parinevad ajavahemikust 2004. aasta algus kuni 2016. aasta 18. aprill.



Joonis 6. Välja ostetud retseptide koguarvud aastate lõikes

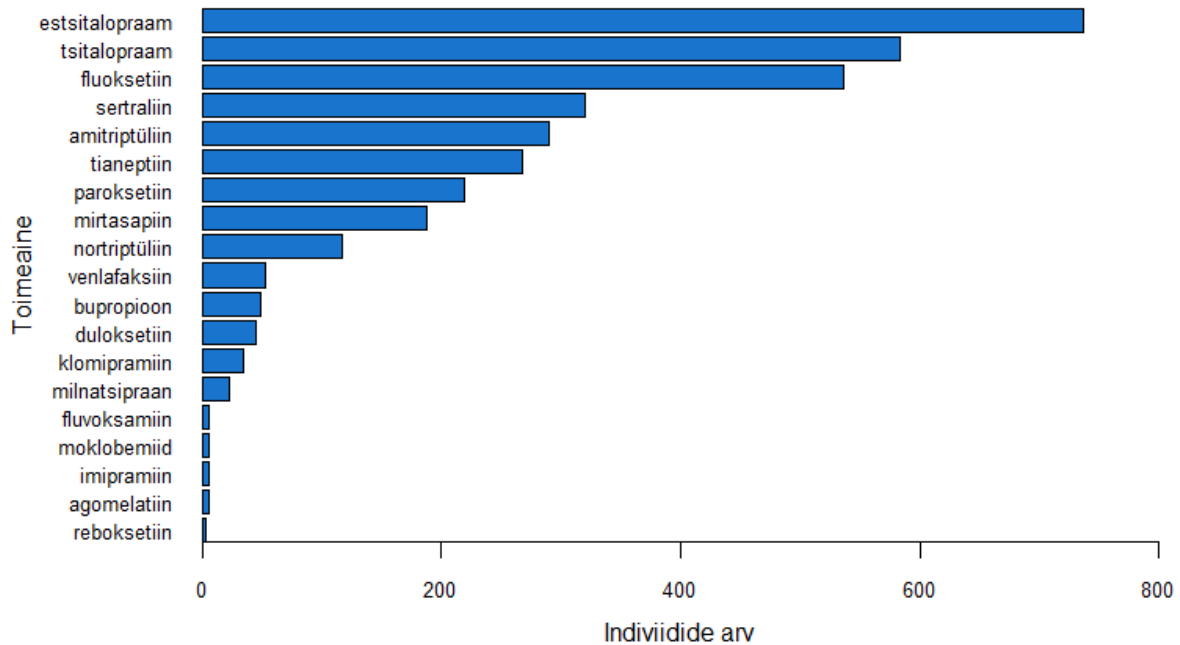
Joonisel 7 on kajastatud retseptide arvu muutused läbi aasta toimeainete gruppide kaupa. Koheselt hakkab silma, et *estsitalopraam* on peaaegu terve aja vältel kõige populaarsem toimeaine. *Tsitalopraam* oli 2010. aastani kahe populaarseima toimeaine seas, kuid pärast seda aastat populaarsus langes.



Joonis 7. Välja ostetud retseptide arvu muutused aastate lõikes toimeainete gruppide kaupa

Tehti ülevaade ka sellest, kui mitme erineva toimeainega ravimilt esimene kord ära vahetati. Kokku vahetas ravimit 3434 inimest. Selgus, et vahetus on üldse tehtud 19 erineva toimeaine puhul. *Vortiooksetiin* oli toimeaine, millelt kordagi teisele ravimile ära ei muudetud.

Joonis 8 annab ülevaate, milliseid toimeaineid ja kui sagedasti esimese muutuse käigus mõne teise toimeaine vastu vahetati. Kõige rohkem vahetati ära *estsitalopraamilt*, 737 korda. Sellele järgnes *tsitalopraam* 583 korruga, edasi *fluoksetiin* 536 ja *sertraliin* 320 korruga.



Joonis 8. Toimeained, millelt esimesel muutusel ära vahetati

Vaadates uuritavate jagunemist vaatluse all olevate ensüümide aktiivsusgruppidesse selgus, et CYP2D6 korral oli aeglaseid metaboliseerijaid 238, normaalseid 5452 ning kiiremaid 19. Ensüümi CYP2C19 korral oli aeglaseid metaboliseerijaid 188, normaalseid 2766 ja kiireid 2755. Geeni *CYP2C19* korral oli 107 inimest, kellel esines deletsioon, ning *CYP2D6* korral 130 inimest, kellel oli deletsioon.

4.4 Analüüs mudelite abil

Käesolevas peatükis on ensüümide aktiivsusgruppide tunnuste korral baastasemeks normaalsete metaboliseerijate grupp.

4.4.1 Loendusandmete mudelid retseptide arvule

Esiteks pakkus huvi, kas mõni tunnus võiks oluliselt mõjutada seda, kui palju erinevate ravimite retsepte geenidonor välja ostis ning milline oli kõigi välja ostetud retseptide koguarv. Saab loendada, kui mitu retsepti keegi välja ostis, seega sobilik oleks kasutada Poissoni regressioonimudelit. Poissoni jaotuse korral on lubatud ka väärtus null. Töös kasutatavate andmete korral on alati välja ostetud retseptide arv vähemalt üks. Selleks, et Poissoni

regressioonimudel paremini sobiks, otsustati retseptide koguarv defineerida kui retseptide arv, mis pärast esimest retsepti välja osteti, sest sellisel juhul on võimalik ka väärtus null, kui indiviid ostiski ainult ühe retsepti. Erinevate ravimite retseptide korral kasutatakse sama loogikat, et tunnuseks on erinevate ravimite retseptide arv, mis pärast esimese ravimi retsepti välja osteti.

Välja ostetud retseptide koguarvule hinnati esialgu Poissoni regressioonimudel. Argumenttunnusteks olid tausttunnustena sugu, vanus ravi alustamise hetkel, pikkus ja kaal ning lisaks mõlema uurimise all oleva ensüümi metaboliseerimise kiirusgrupid ja geenide deletsioonigrupid. Kuna argumenttunnuseks valiti vanus ravi alustamise hetkel, siis sünniaasta jäeti välja, kuna need tunnused on seotud. Samamoodi jäi välja kehamassiindeks, kuna see on seotud pikkuse ja kaaluga. Ka kõigi järgnevate mudelite korral on tausttunnused sama loogika põhjal valitud. Mudelisse lisati ka *offset* tunnus, milleks oli jälgimisaeg ning mudeli kuju väljakirjutatuna on järgnev:

$$\ln \frac{\text{välja ostetud retseptide arv}}{\text{jälgimisaeg}} = \beta_0 + \beta_1(\text{sugu} = \text{naine}) + \beta_2\text{vanus} + \beta_3\text{pikkus} + \\ + \beta_4\text{kaal} + \beta_5(\text{CYP2C19 aktiivsus} = \text{aeglane}) + \beta_6(\text{CYP2C19 aktiivsus} = \text{kiire}) + \\ + \beta_7(\text{CYP2D6 aktiivsus} = \text{aeglane}) + \beta_8(\text{CYP2D6 aktiivsus} = \text{kiire}) + \\ + \beta_9(\text{CYP2C19 deletsioon} = 1) + \beta_{10}(\text{CYP2D6 deletsioon} = 1).$$

Olgu mainitud, et $\ln(\text{jälgimisaeg})$ on eespool mainitud *offset*.

Selgub, et esialgses mudelis esineb suur ülehajuvus, kuna $\frac{D \text{ ehk hälbimusjääk}}{df \text{ ehk vabadusastmed}} = \frac{106\,692}{5698} = 19$ ja antud Poissoni regressioonimudel ei ole andmete jaoks sobilik ning seetõttu sobitati järgmise sammuna negatiivne binoomimudel.

Tabel 2. Retseptide koguarvu mudeli (negatiivne binoommudel) parameetrite hinnangud (β) koos 95%-usaldusintervalli (UI) ja olulisuse tõenäosuse ehk p -väärtusega (p)

Tunnus	Esialgne mudel	Lõplik mudel	
	β (UI)	β (UI)	p
Sugu = naine	0,2328 (0,1171; 0,3478)	0,2047 (0,1136; 0,2950)	8,48e-06
Vanus	0,0008 (-0,0017; 0,0033)		
Kaal	0,0023 (0,0002; 0,0048)	0,0028 (0,0005; 0,0050)	0,019
Pikkus	0,0028 (-0,0033; 0,0089)		
CYP2C19 aktiivsus = aeglane	-0,1092 (-0,3091; 0,1015)		
CYP2C19 aktiivsus = kiire	0,0398 (-0,0336; 0,1131)		
CYP2C19 deletsioon = 1	0,0083 (-0,2501; 0,2869)		
CYP2D6 aktiivsus = aeglane	0,0791 (-0,0961; 0,2631)		
CYP2D6 aktiivsus = kiire	-0,1841 (-0,7618; 0,5070)		
CYP2D6 deletsioon = 1	0,2211 (-0,0137; 0,4721)	0,2375 (0,0034; 0,4878)	0,0541

Negatiivse binoommudeli tulemused on toodud tabelis 2. Ei esialgse ega ka lõpliku mudeli korral ei ole ülehajuvuse probleemi, sest hälbimusjäägi ja vabadusastmete jagatis on mõlemal juhul ligikaudu üks. Lõplikus mudeli järgi saab öelda, naistel on võrreldes meestega retseptide arv $e^{0,2047} = 1,2272$ korda suurem ning tulemustest võib järeldada, et kaalu kasvades retseptide arv suureneb, kuigi selle tunnuse mõju ei ole väga tugev.

Lisaks hinnati Poissoni mudel ka erinevate välja ostetud ravimite retseptide arvule. Argumenttunnused ja *offset* olid samad nagu retseptide koguarvu mudelis eespool. Ka siinkohal pöörati tähelepanu sellele, kas tegu võib olla ülehajuvusega, mida tuleb arvesse võtta. Selgus, et esialgses ja lõplikus mudelis ülehajuvusega märgatavat probleemi ei olnud. Mudeli tulemused on nähtavad tabelis 3 (lk 30). Selgub, et lõplikus mudelis on olulised sugu ja CYP2C19 deletsioon. Naistel on erinevate ravimite retseptide arv $e^{0,1567} = 1,1696$ korda suurem kui meestel. Lisaks selgub, et geeni CYP2C19 korral on deletsiooniga indiviididel retseptide arv $e^{0,1892} = 1,2083$ korda suurem.

Tabel 3. Erinevate ravimite retseptide arvu mudeli (Poissoni regressioonimudel) parameetrite hinnangud (β) koos 95%-usaldusintervalli (UI) ja p -väärtusega (p)

Tunnus	Esialgne mudel		Lõplik mudel	
	β (UI)		β (UI)	p
Sugu = naine	0,1549 (0,0808; 0,2294)		0,1567 (0,0840; 0,2298)	2,51e-05
Vanus	0,0004 (-0,0012; 0,0020)			
Kaal	-0,0008 (-0,0024; 0,0007)			
Pikkus	0,0042 (0,0003; 0,0080)		0,0034 (0,0000; 0,0068)	0,0526
CYP2C19 aktiivsus = aeglane	0,0313 (-0,0970; 0,1552)			
CYP2C19 aktiivsus = kiire	0,0240 (-0,0211; 0,0692)			
<i>CYP2C19</i> deletsioon = 1	0,1903 (0,0395; 0,3343)		0,1892 (0,0390; 0,3325)	0,0115
CYP2D6 aktiivsus = aeglane	0,0387 (-0,0695; 0,1435)			
CYP2D6 aktiivsus = kiire	-0,0397 (-0,4793; 0,3442)			
<i>CYP2D6</i> deletsioon = 1	-0,0232 (-0,1820; 0,1280)			

4.4.2 Logistiline regressioonimudel ravimi vahetamisele

Selgitamaks välja, kas antidepressandi vahetamine sõltub uurimise all olevate geenide deletsioonist, nende poolt kodeeritud ensüümide aktiivsusgrupist või tausttunnustest, hinnati esmalt logistiline regressioonimudel, kus uuritavaks tunnuseks on ravimi vahetamise tunnus.

Seletavateks tunnusteks olid ensüümi CYP2C19 aktiivsusgrupp, CYP2D6 aktiivsusgrupp, *CYP2C19* ja *CYP2D6* deletsioon ning tausttunnused (sugu, vanus ravi alustamise hetkel, kaal, pikkus).

Välja otsustati jätta ka ühekordsed ostjad, kuna neil ei saanud muutust toimuda. Arvesse võeti 4596 indiviidi. Ensüümi CYP2C19 korral oli aeglaseid metaboliseerijaid 140, normaalseid 2229 ja kiireid 2227. Ensüümi CYP2D6 korral oli aeglaseid metaboliseerijaid 198, normaalseid 4382 ja kiireid 16. Esialgse ja lõpliku mudeli tulemused on kajastatud tabelis 4 (lk 31).

Tabel 4. Logistilise regressioonimudeli hinnangud (β) ravimi vahetamisele koos 95%-usaldusintervalli (UI) ja p -väärtusega (p)

Tunnus	Esialgne mudel		Lõplik mudel	
	β (UI)		β (UI)	p
Sugu = naine	0,4474 (0,2357; 0,6588)		0,4543 (0,2904; 0,6174)	5,11e-08
Vanus	-0,0078 (-0,0125; -0,0032)		-0,0078 (-0,0121; -0,0036)	0,000296
Kaal	0,0042 (-0,0004; 0,0089)		0,0042 (-0,0001; 0,0086)	0,058279
Pikkus	-0,0002 (-0,0116; 0,0113)			
CYP2C19 aktiivsus = aeglane	0,2548 (-0,1492; 0,6881)			
CYP2C19 aktiivsus = kiire	0,0501 (-0,0858; 0,1860)			
CYP2C19 deletsioon = 1	0,3500 (-0,1750; 0,9353)			
CYP2D6 aktiivsus = aeglane	0,2144 (-0,1253; 0,5753)			
CYP2D6 aktiivsus = kiire	-0,7377 (-1,7347; 0,3005)			
CYP2D6 deletsioon = 1	-0,0411 (-0,4795; 0,4288)			

Sellel juhul jäid lõplikus mudelis olulised sugu ja vanus. Saab öelda, et meestega võrreldes on naistel šanss ravimi vahetamiseks ligikaudu $e^{0,4543} = 1,5751$ korda suurem. Kui võrrelda kahte inimest, kellest üks on aasta vanem kui teine ning ülejäänud tunnuste väärtused on samad, siis saab öelda, et vanemal inimesel on šansid vahetamiseks $e^{-0,0078} = 0,9922$ korda suuremad ehk $e^{0,0078} = 1,0078$ korda väiksemad.

4.4.3 Coxi võrdeliste riskide mudel ning konkureerivate riskide mudel ravimi vahetamisele

Kokku vahetas antidepressanti 3434 indiviidi ning 2275 oli neid, kes ei vahetanud. Neist, kes ei vahetanud, olid 2245 sellised, kes katkestasid ravi enne jälgimisaja lõppu. Arvestades, et pärast retsepti realiseerimist tarvitatakse veel viimase retsepti ravimeid, siis võeti kuupäev 18. aprill 2016 – 2 kuud = 18. veebruar 2016 tsenseerituks määramise piiriks ehk kui indiviid ostis ravimi pärast seda kuupäeva, siis loeti ta tsenseerituks. Selliseid indiviide oli 30.

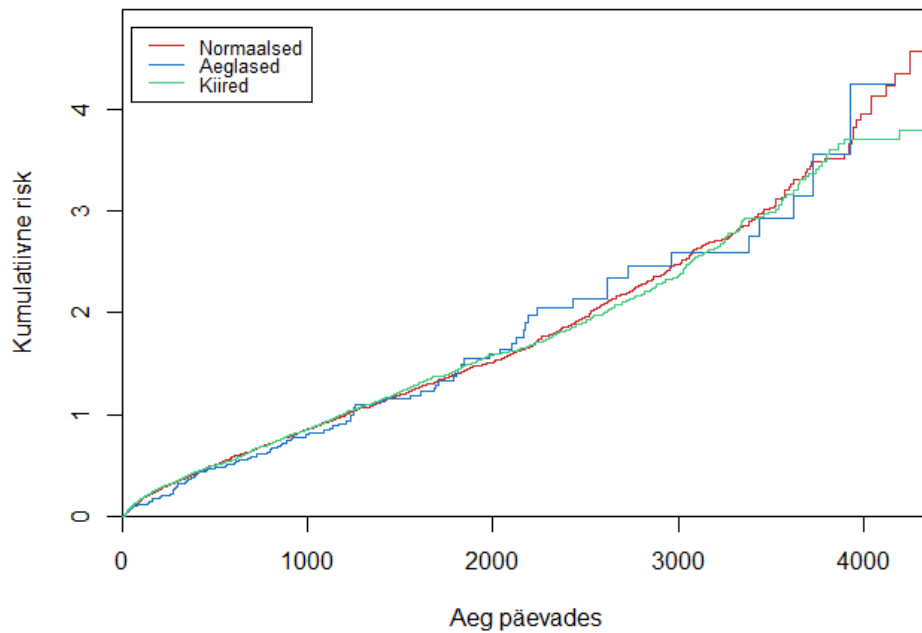
Coxi võrdeliste riskide mudel hinnati nii, et sündmuse väärtus oli üks, kui indiviid vahetas antidepressanti ning null siis, kui indiviid ravimit ei muutnud (ehk siia kuulusid need, kes võtsid lõpuni või katkestamise hetkeni sama ravimit, millega ravi alustati). Ajatunnuseks oli siis vastavalt aeg vahetuseni, aeg lõpuni või aeg katkestamiseni. Antud juhul jäeti erinevalt logistilisest regressioonimudelist ühekordsed ostjad sisse, kuna Coxi võrdeliste riskide mudel võtab arvesse ka aja. Argumenttunnustena lisati mudelisse sugu, vanus ravi alustamise ajal, pikkus, kaal ning mõlema ensüümi aktiivsusegruppide ja geenide deletsioonide tunnused.

Mudeli hindamise tulemused kajastuvad tabelis 5. Samuti leiti, et Coxi võrdeliste riskide mudeli eeldus riskide võrdelisuse kohta on täidetud.

Tabel 5. Coxi võrdeliste riskide mudeli parameetrite hinnangud (e^{β}) koos 95%-usaldusintervalli (UI) ja p -väärtusega (p)

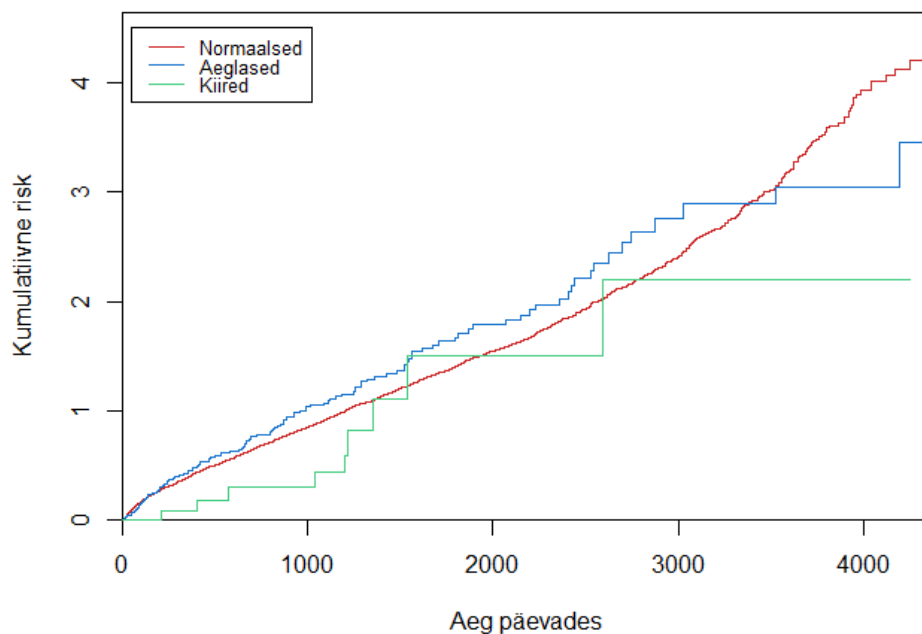
Tunnus	Esialgne mudel	Lõplik mudel	
	e^{β} (UI)	e^{β} (UI)	p
Sugu = naine	1,0891 (0,9742; 1,2175)	1,1034 (0,9898; 1,2301)	0,07595
Vanus	1,0006 (0,9982; 1,0030)		
Kaal	0,9981 (0,9958; 1,0005)		
Pikkus	1,0097 (1,0039; 1,0155)	1,0085 (1,0034; 1,0136)	0,00103
CYP2C19 aktiivsus = aeglane	0,9609 (0,7919; 1,1660)		
CYP2C19 aktiivsus = kiire	1,0085 (0,9419; 1,0797)		
CYP2C19 deletsioon = 1	1,1558 (0,9094; 1,4688)		
CYP2D6 aktiivsus = aeglane	1,0883 (0,9264; 1,2785)		
CYP2D6 aktiivsus = kiire	0,6586 (0,3416; 1,2699)		
CYP2D6 deletsioon = 1	0,9952 (0,7934; 1,2483)		

Tabelist 5 selgub, et vahetamise risk on pikematel inimestel suurem. Kui võrrelda kahte inimest, kellest esimene on teisest ühe sentimeetri võrra pikem, siis pikemal inimesel on risk vahetada ravimit 1,0085 korda suurem kui lühemal inimesel. Pikkuse mõju on seega pigem väike.



Joonis 9. Ensüümi CYP2C19 aktiivsusgruppide kumulatiivse riski graafik muutmisele

Mudeli põhjal ei olnud kummagi ensüümi puhul üheski grupis normaalsete metaboliseerijate grupiga võrreldes statistiliselt olulisi erinevusi. Joonis 9 annab ülevaate sellest, kuidas ensüümi CYP2C19 erinevate aktiivsusgruppide korral ravimi muutmise risk aja jooksul muutub. On näha, et kuni päevani 2000 on grupid üsna sarnased. Pärast seda hakkab silma, et aeglase grupis on vahepeal risk suurem kui teistes gruppides ning mingil hetkel väiksem kui teistes gruppides.



Joonis 10. Ensüümi CYP2D6 aktiivsusgruppide kumulatiivse riski graafik muutmisele

Joonis 10 (lk 33) kirjeldab CYP2D6 aktiivsusgruppide võimalikku mõju ravimi vahetamise riskile. On näha, et aeglastel metaboliseerijatel on alguses pikka aega ravimi muutmise risk suurem kui teistes gruppides. Kiirete metaboliseerijate grupp on väga väike, kuid tundub, et seal on alguses risk ravimit muuta väiksem kui teistes gruppides.

Lisaks hinnati Coxi võrdeliste riskide mudel ka nii, et määrati ravimi vahetamisele ajapiirang. Esialgu võeti ajapiiranguks 5 aastat ehk 1825 päeva. Kui indiviid selle aja jooksul ravimit vahetas, siis kuulus ta nn muutjate ehk vahetajate hulka, vastasel juhul loeti tema ravimi vahetamise aeg tsenseerituks. Sellisel juhul oli muutjaid 2876. Lõplikus mudelis oli oluline vaid pikkus (ka see mõju oli nõrk).

Võrdluseks otsustati võtta ka lühem periood, täpsemalt aasta ehk 365 päeva. Sellisel juhul oli muutjaid 1394 ning mittemuutjaid 4315. Lõplikus mudelis olid olulised pikkus (p -väärtus 0,00184 ja $e^{\beta} = 1,0103$) ning kaal (p -väärtus 0,04985 ja $e^{\beta} = 0,9965$).

Nii üheaastase kui ka viieaastase perioodi mudeli korral on võrdeliste riskide eeldusega kõik korras.

Järgmisena hinnati konkureerivate riskide mudel, kus sündmuse väärtus oli üks, kui geenidoonor muutis ravimit. Neil, kes ei muutnud ravimit, aga katkestasid enne uuringu lõppu, oli sündmuse väärtus kaks ja neil, kes ei muutnud ravimit ja tarbisid seda uuringu lõpuni, oli sündmuse väärtus null. Ajatunnuseks oli vastavalt aeg muutmiseni, aeg katkestamiseni või aeg uuringu lõpuni. Seletavad tunnused olid sugu, vanus ravi alustamise ajal, pikkus, kaal ning mõlema ensüümi aktiivsusgruppide ja geenide deletsioonide tunnused. Sündmus kodeeriti väärtusega null 30 inimese puhul, väärtusega üks 3434 ja väärtusega kaks 2245 inimese puhul. Indiviid kuulus selle sündmuse gruppi, mis kõige varem toimus.

Tabel 6 annab ülevaate mudeli hindamise tulemustest. Sealt selgub, et naistel on meestega võrreldes risk muutmisele 1,2823 korda suurem, samuti on rohkem kaaluvatel indiviididel risk muutmisele suurem võrreldes väiksema kaaluga indiviididega. Vaadates kahte inimest, kellest üks on aasta vanem kui teine ning teiste tunnuste väärtused neil ei erine, siis saab öelda, et vanemal inimesel on risk muutmisele 1,0072 korda väiksem.

Tabel 6. Konkureerivate riskide mudeli parameetrite hinnangud (e^{β}) koos 95%-usaldusintervalli (UI) ja p -väärtusega (p)

Tunnus	Esialgne mudel	Lõplik mudel	
	e^{β} (UI)	e^{β} (UI)	p
Sugu = naine	1,2963 (1,1581; 1,4510)	1,2823 (1,1738; 1,4008)	3,6e-08
Vanus	0,9930 (0,9907; 0,9953)	0,9929 (0,9908; 0,9949)	2,5e-11
Kaal	1,0029 (1,0006; 1,0052)	1,0031 (1,0009; 1,0052)	5,8e-03
Pikkus	1,0011 (0,9954; 1,0068)		
CYP2C19 aktiivsus = aeglane	0,9401 (0,7800; 1,1331)		
CYP2C19 aktiivsus = kiire	1,0314 (0,9636; 1,1039)		
CYP2C19 deletsioon = 1	1,1250 (0,8847; 1,4304)		
CYP2D6 aktiivsus = aeglane	1,1604 (0,9892; 1,3614)	1,1587 (0,9877; 1,3592)	7,1e-02
CYP2D6 aktiivsus = kiire	0,6816 (0,3697; 1,2569)	0,6802 (0,3683; 1,2560)	2,2e-01
CYP2D6 deletsioon = 1	0,9497 (0,7628; 1,1824)		

Lisaks eelnevale konkureerivate riskide mudelile otsustati proovida ka selliseid mudeleid, kus vaatluse all on kindel ajaperiood. Otsustati, et vaadatakse sarnaselt Coxi võrdeliste riskide mudelile perioodi pikkusega viis aastat ja lisaks ka lühemat üheaastalist perioodi.

Esialgu võeti perioodiks 365 päeva ehk aasta. Sündmuse väärtus oli null, kui patsiendid võtsid ühte ja sama ravimit vähemalt 365 päeva ning üks, kui patsiendid vahetasid ravimit 365 päeva jooksul. Tunnus sai väärtuseks kaks kui katkestati ühe ja sama ravimi võtmine vähem kui 365 päeva jooksul. Väärtuse null sai 2554 indiviidi, väärtuse üks 1394, väärtuse kaks 1761 indiviidi.

Selgub, et 365-päevase perioodi puhul on lõplikus mudelis oluline ainult vanus. Sarnaselt eelmise mudeliga on ka siin nii, et vanuse kasvades riskid vähenevad. Täpsemalt on risk aasta vanemal inimesel võrreldes aasta noorema inimesega 0,9943 korda suurem ehk 1,0057 korda väiksem.

Järgmisena hinnati eelneva mudeliga analoogne mudel, kus perioodiks oli viis aastat ehk 1825 päeva. Selgus, et 723 inimest võtsid 1825 päeva sama ravimit, 2876 inimest vahetasid ravimit

selle aja jooksul ning 2110 inimest katkestasid ühe ja sama ravimi võtmise. Tulemused on näha tabelis 7.

Tabel 7. Konkureerivate riskide mudeli parameetrite hinnangud (e^β) koos 95%-usaldusintervalli (UI) ja p -väärtusega (p) 1825-päevase perioodi korral

Tunnus	Esialgne mudel	Lõplik mudel	
	e^β (UI)	e^β (UI)	p
Sugu = naine	1,2693 (1,1233; 1,4342)	1,2132 (1,1026; 1,3349)	7,4e-05
Vanus	0,9948 (0,9923; 0,9973)	0,9942 (0,9919; 0,9964)	5,2e-07
Kaal	1,0018 (0,9993; 1,0044)	1,0024 (1,0000; 1,0048)	5,1e-02
Pikkus	1,0040 (0,9978; 1,0103)		
CYP2C19 aktiivsus = aeglane	0,9182 (0,7451; 1,1316)		
CYP2C19 aktiivsus = kiire	1,0456 (0,9706; 1,1264)		
CYP2C19 deletsioon = 1	1,2225 (0,9616; 1,5541)		
CYP2D6 aktiivsus = aeglane	1,1981 (1,0125; 1,4177)	1,1965 (1,0111; 1,4159)	3,7e-02
CYP2D6 aktiivsus = kiire	0,7298 (0,3944; 1,3503)	0,7244 (0,3909; 1,3425)	3,1e-01
CYP2D6 deletsioon = 1	0,9567 (0,7539; 1,2138)		

Saab öelda, et naistel on risk ravimit vahetada 1,2132 korda kõrgem kui meestel ning vanuse kasvades risk väheneb. Ensüümi CYP2D6 korral on aeglaste grupis võrreldes normaalsete grupiga risk ravimi vahetamiseks 1,1965 korda suurem.

Kokkuvõte

Magistritöö eesmärk oli välja selgitada, kuidas geneetiline varieeruvus *CYP2C19* ja *CYP2D6* geenides, mis põhjustab antidepressantide metabolismis oluliste ensüümide ravimilagundamise kiiruse erinevusi, mõjutab ühe antidepressandi vahetamist teise vastu ning üldist antidepressantide kasutust. Täpsemalt uuriti, millist mõju avaldavad geenide *CYP2C19* ja *CYP2D6* deletsioonid ja nende geenide poolt kodeeritud ensüümide metaboliseerimise kiirusgrupid ning erinevad tausttunnused (sugu, vanus ravi alustamise hetkel, kaal ja pikkus) antidepressantide vahetamisele ja kasutamisele. Analüüsis kasutati Poissoni regressioonimudelit ja negatiivset binoommudelit retseptide arvu hindamiseks ning logistilist regressioonimudelit, Coxi võrdeliste riskide mudelit ja konkureerivate riskide mudelit depressiooniravimi vahetamise jaoks. Analüüsis kasutatavad andmed olid TÜ Eesti Geenivaramu geenidonorite geeni- ja retseptiandmed.

Selgus, et naistel on meestega võrreldes retseptide koguarv 1,2272 korda suurem. Kaalu kasvades retseptide arv suureneb.

Erinevate välja ostetud retseptid arv on naistel 1,1696 korda suurem kui meestel. Neil indiviididel, kellel esineb geen *CYP2C19* deletsioon, on erinevate ravimite retseptide välja ostmise arv 1,2083 korda suurem võrreldes indiviididega, kellel antud geen deletsiooni ei esine.

Logistilise regressiooni korral selgus, et naistel on šanss ravimi vahetamiseks 1,5751 korda suurem kui meestel ning võrreldes kahte indiviidi, kellest üks on aasta noorem, saab öelda, et vanemal on šansid ravimi vahetamiseks 1,0078 korda väiksemad.

Coxi võrdeliste riskide mudeli korral selgus, et sõltuvalt ravimi vahetamise perioodist võivad vahetust mõjutada kas ainult pikkus või lisaks pikkusele ka kaal.

Konkureerivate riskide korral kitsendusteta perioodi põhjal mõjutavad ravimi vahetamist vanus ja kaal ning naistel on risk ravimi vahetamiseks 1,2823 korda suurem kui meestel. Aastase perioodi puhul on oluline ainult vanus ning 5-aastase perioodi puhul saab öelda, et võrreldes normaalsete metaboliseerijatega on aeglase metaboliseerijate puhul ravimi vahetamise risk 1,1965 korda suurem. Samuti saab öelda, et naistel on risk 1,2132 korda suurem kui meestel ning vanuse kasvades risk väheneb.

Käesolev töö on sobiv alus edasiarenduse tegemiseks. Antud töös vaadeldi geenide *CYP2C19* ja *CYP2D6* varieeruvuse mõju kõikide antidepressantide korral. Osade antidepressantide puhul on näidatud, et vähemalt üks ensüümidest *CYP2C19* ja *CYP2D6* on peamine lagundaja. Täpsemalt tasuks moodustada alagrupid antidepressantidest selle järgi, milline neist ensüümidest on peamine lagundaja ning vaadelda siis aktiivsusgruppide mõju. Lisaks oleks võimalik täpsemaid tulemusi saada, kui suuremal osal indiviididel saaks määrata täpse aktiivsusgrupi ning selle abil tõsta uuritavate indiviidide arvu. Veel aitaks kaasa, kui iga välja ostetud retsepti korral oleks teada, kui palju tablette see retsept sisaldas, kuna nii oleks võimalik täpsemalt välja arvutada reaalne aeg, mil inimene ravimeid tarbis.

Kirjanduse loetelu

- [1] McKenzie, K. (2010). *Mis on depressioon*. Tallinn: Eesti Ajalehed AS.
- [2] Souter, K. (2015). *Mis on depressioon ja kuidas sellega toime tulla*. Tallinn: Tammerraamat.
- [3] *Perearst: antidepressandid töötavad, kuid tulemusi ei näe ühe päevaga*. (2018). <https://novaator.err.ee/691090/perearst-antidepressandid-tootavad-kuid-tulemusi-ei-nae-uhe-paevaga> (14.05.2019).
- [4] Wasserman, D. (2011). *Depression*. Second Edition. New York: Oxford University Press.
- [5] Marandi, T. (1999). *Drug Treatment of Depression in Estonia*. Tartu: Tartu Ülikooli Kirjastus.
- [6] Heinaru, A. (2012). *Geneetika*. Õpik kõrgkoolile. Tartu: Tartu Ülikooli Kirjastus.
- [7] Digiõppevaramu. (2017). Nukleiinhapped. *Nukleiinhapete ehitus ja ülesanded*. <https://vara.e-koolikott.ee/taxonomy/term/3538> (14.05.2019).
- [8] *Help Me Understand Genetics*. (2017). https://courses.cs.ut.ee/MTAT.03.242/2017_fall/uploads/Main/Genetics.pdf (14.05.2019).
- [9] Simmons, H. (2019). Medical Science News. *What are Cytochrome P450 Enzymes?* <https://www.news-medical.net/life-sciences/What-are-Cytochrome-P450-Enzymes.aspx> (14.05.2019).
- [10] Genetics Home Reference. *Help Me Understand Genetics. Cytochrome p450*. <https://ghr.nlm.nih.gov/primer/genefamily/cytochromep450> (14.05.2019).
- [11] Lynch, T., Price, A. (2007). *The effect of cytochrome P450 metabolism on drug response, interactions, and adverse effects*. <https://www.ncbi.nlm.nih.gov/pubmed/17708140> (14.05.2019).
- [12] Li, J., Bluth M. H. (2011). *Pharmacogenomics of drug metabolizing enzymes and transporters: implications for cancer therapy*. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3513217> (14.05.2019).

- [13] Kalman, L.V., Agúndez, J. A. G., Appell, M. L., Black, J. L., Bell, G. C., Boukouvala S., Bruckner, C., Bruford, E., Caudle, K., Coulthard, S., Daly, A. K., Del Tredici, A. L., et al. (2015). *Pharmacogenetic Allele Nomenclature: International Workgroup Recommendations for Test Result Reporting*. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4724253/> (14.05.2019).
- [14] Robarge, JD., Li, L., Desta, Z., Nguyen, A., Flockhart, DA. (2007). *The star-allele nomenclature: retooling for translational genomics*. <https://www.ncbi.nlm.nih.gov/pubmed/17700589> (14.05.2019).
- [15] Aldrich, S. L., Poweleit, E. A., Prows, C.A., Martin, L. J., Strawn, J. R., Ramsley, L. B. (2019). *Influence of CYP2C19 Metabolizer Status on Escitalopram/Citalopram Tolerability and Response in Youth With Anxiety and Depressive Disorders*. <https://www.ncbi.nlm.nih.gov/pubmed/30837874> (14.05.2019).
- [16] Hicks, J. K., Sangkuhl, K., Swen, J. J., Ellingrod, V. L., Müller, D. J., Shimoda, K., Bishop, J. R., Kharasch, E. D., Skaar, T. C., Gaedigk, A., Dunnenberger, H. M., Klein, T. E., Caudle, K. E., Stingl, J. C. (2017). *Clinical Pharmacogenetics Implementation Consortium Guideline (CPIC) for CYP2D6 and CYP2C19 Genotypes and Dosing of Tricyclic Antidepressants: 2016 Update*. *Clinical Pharmacology & Therapeutics*, juuli 2017. https://cpicpgx.org/content/guideline/publication/TCA/2016/TCA_2016.pdf (14.05.2019).
- [17] Käärik, E. (2013). *Andmeanalüüs II. Loengukonspekt*. Tartu Ülikool, matemaatilise statistika instituut. <http://dspace.ut.ee/bitstream/handle/10062/35401/AndmeanalüüsII.pdf> (14.05.2019).
- [18] Võrk, A. (2015). *Mikroökonomieetria konspekt*. Tartu Ülikool, rahvamajanduse instituut. https://www.researchgate.net/publication/280012070_Mikroökonomieetria_konspekt (14.05.2019).
- [19] Hilbe, J. M. (2011). *Negative Binomial Regression*. Second Edition. Cambridge: Cambridge University Press.
- [20] Käärik, M. (2018). *Üldistatud lineaarsed mudelid*. Loenguslaidid. Tartu Ülikool, matemaatika ja statistika instituut. <https://courses.ms.ut.ee/2018/glm/spring/Main/Lectures> (14.05.2019).

- [21] Collett, D. (2015). *Modelling Survival Data in Medical Research*. Third Edition. Boca Raton: CRC Press.
- [22] Klein, J. P., Van Houwelingen H. C., Ibrahim, J. G., Scheike, T. H. (2014). *Handbook of Survival Analysis*. Boca Raton: CRC Press. https://books.google.ee/books?id=t1vOBQAAQBAJ&printsec=frontcover&redir_esc=y#v=onepage&q&f=false (14.05.2019).
- [23] Rdocumentation. Package „stats“. *Fitting Generalized Linear Models*. <https://www.rdocumentation.org/packages/stats/versions/3.5.3/topics/glm> (14.05.2019).
- [24] Rdocumentation. Package „stats“. *Family Objects For Models*. <https://www.rdocumentation.org/packages/stats/versions/3.5.3/topics/family> (14.05.2019).
- [25] Rdocumentation. Package „MASS“. *Fit A Negative Binomial Generalized Linear Model*. <https://www.rdocumentation.org/packages/MASS/versions/7.3-51.3/topics/glm.nb> (14.05.2019).
- [26] Therneau, T. M., Lumley, T. (2019). Package „survival“. *Survival Analysis*. <https://cran.r-project.org/web/packages/survival/survival.pdf> (14.05.2019).
- [27] Gray, B. (2015). Package „cmprsk“. *Subdistribution Analysis of Competing Risks*. <https://cran.r-project.org/web/packages/cmprsk/cmprsk.pdf> (14.05.2019).
- [28] Reisberg, S., Krebs, K., Lepamets, M., Kals, M., Mägi, R., Metsalu, K., Lauschke, V. M., Vilo, J., Milani, L. (2018). *Translating genotype data of 44,000 biobank participants into clinical pharmacogenetic recommendations: challenges and solutions*. *Genetics in Medicine*, 2018. <https://www.nature.com/articles/s41436-018-0337-5> (14.05.2019)

Lihtlitsents lõputöö reprodutseerimiseks ja üldsusele kättesaadavaks tegemiseks

Mina, Pille-Riin Vare,

1. annan Tartu Ülikoolile tasuta loa (lihtlitsentsi) minu loodud teose „Geneetilise varieeruvuse mõju antidepressantide tarvitamisele TÜ Eesti Geenivaramu andmete põhjal“, mille juhendajad on Krista Fischer ja Kristi Krebs,

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.

Pille-Riin Vare

15.05.2019