

TARTU ÜLIKOOL
LOODUS- JA TÄPPISTEADUSTE VALDKOND
MATEMAATIKA JA STATISTIKA INSTITUUT

Minni-Marii Paarmets
**Primaarse ja sekundaarse ravijärgimuse
mõõtude võrdlus**

Matemaatiline statistika
Bakalaureusetöö (9 EAP)

Juhendaja: PhD Raivo Kolde

TARTU 2023

PRIMAARSE JA SEKUNDAARSE RAVIJÄRGIMUSE MÕÕTUDE VÕRDLUS

Bakalaureusetöö

Minni-Marii Paarmets

Lühikokkuvõte

Ravijärgimus näitab kui hästi isik talle määratud ravi järgib. See jaguneb kaheks: primaarne ja sekundaarne ravijärgimus. Primaarne ravijärgimus keskendub ühekordsetele retseptiravimitele, mille puhul vaadatakse, kas ravim osteti välja või mitte. See-eest sekundaarne ravijärgimus arvutatakse pikaajaliste retseptiravimite põhjal ning selle arvutusmeetodid on juba komplekssemad.

Töö esimene osa kontsentreerub töö taustale, andmetele ja kasutatavatele meetodikatele ning teises osas uuritakse täpsemalt nii primaarset kui ka sekundaarset ravijärgimust mõjutavaid tunnuseid ning võrreldakse saadud tulemusi.

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

Märksõnad: Ravijärgimus, lineaarne regressioon, logistilise regressiooni segamudel.

COMPARISON OF PRIMARY AND SECONDARY MEDICATION ADHERENCE MEASURES

Bachelor thesis

Minni-Marii Paarmets

Abstract

Medication adherence indicates how well a person follows the treatment plan.

It is divided into two categories: primary and secondary medication adherence. Primary medication adherence focuses mainly on non-recurrent prescription medications and if a person purchases medication prescribed to them. On the contrary secondary medication adherence is calculated based on long-term prescription drugs and its calculation methods are more complex.

First part of the thesis focuses on the background, data, and methodology used. In contrast, the second part examines in more detail the characteristics that affect both primary and secondary treatment adherence and compares the results obtained.

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

Key Words: Medication adherence, mixed effects linear regression model, mixed effects logistic regression model.

Sisukord

Sissejuhatus	4
1 Ravijärgimus	5
2 Andmestik	8
2.1 Sekundaarse ravijärgimuse andmestik	9
2.2 Primaarse ravijärgimuse andmestik	10
3 Mudelid	11
3.1 Logistiline regressioonimudel	11
3.2 Mitme argumentiga lineaarne regressioon	11
3.3 Lineaarse ja logistilise regressiooni segamudel	12
3.4 Mudelite hindamine	13
4 Andmestike kirjeldav analüüs	14
4.1 Primaarse ravijärgimuse andmestiku kirjeldus	14
4.2 Sekundaarse ravijärgimuse andmestiku kirjeldus	17
5 Ravijärgimuse ennustamine	20
5.1 Primaarse ravijärgimuse mudelite analüüs	20
5.2 Sekundaarse ravijärgimuse mudelite analüüs	24
5.3 Mõõtude võrdlus	28
Kokkuvõte	32
Kasutatud allikad	33

Sissejuhatus

Enamike haiguste ravi edukus sõltub osaliselt sellest, kui hästi patsiendid endale määratud ravi järgivad. Seda nimetatakse ravijärgimuseks ehk ravisoostumiseks. Ravijärgimus on jagatud kaheks: primaarne ja sekundaarne ravijärgimus. Primaarse ravijärgimuse puhul uuritakse peamiselt ühekordsete retseptiravimite väljaostmist ning sekundaarse ravijärgimuse korral uuritakse pikaajaliste retseptiravimite väljaostmist ning retseptide uuendamist. Pikaajalisi retseptiravimeid kirjutatakse välja eelkõige krooniliste haiguste puhul. Antud töös on ravijärgimuse uurimiseks kasutatud Eesti astmapatsientide andmeid.

Logistilise regressiooni segamudeliga on uuritud sekundaarset ravijärgimust ning seda mõjutavaid faktoreid Eesti astma patsientide seas (Pähklamäe, 2022). Saadud tulemuste põhjal tekkis küsimus, kas indiviidi primaarne ravijärgimus võib ennustada sekundaarset ravijärgimust ning kui suur on selle mõju.

Töö eesmärk on uurida, kas ja kui palju ennustab ravimite väljaostmine, ehk primaarne ravijärgimus, püsivust krooniliste ravimite võtmisel, ehk sekundaarset ravijärgimust, tulevikus. Selleks moodustatakse logistilise regressiooni segamudel primaarse ravijärgimuse uurimiseks ning rakendatakse sealt saadud isikul põhinevat juhusliku mõju sekundaarse ravijärgimuse lineaarse regressiooni segamudelis.

1 Ravijärgimus

Ravijärgimus ehk ravisoostumus näitab, kas ja kui hästi järgib patsient arsti poolt määratud raviplaani. Ravi mittejärgimus on ühiskonnas probleemiks mitmel põhjusel. See mõjutab ravi efektiivsust ning on eriti kahjulik pikaajalist ravi vajavate krooniliste haigusseisundite korral. Ravijärgimust võivad mõjutada mitmed sotsiaalsed ja majanduslikud faktorid nagu näiteks sissetulek, haridus, tervis ning ka indiviidi spetsiifilised ravi ja haigusega seotud tegurid. Selleks, et ravijärgimus tõuseks, tuleks enne olla teadlik teguritest, mis langetavad seda ning tegeleda nende faktorite elimineerimisega (Gast ja Mathes, 2019).

Ravi mittejärgimise üks põhjustest on patsiendi vähene arusaam ravist ja raviplaanist. Suur roll selle parandamisel on apteekritel ja arstidel. Kui teada ette, millistel isikutel on suurem tõenäosus olla ravi mittejärgiv, siis on võimalik pöörata rohkem tähelepanu nende raviplaanile ning kasutada erinevaid meetmeid olukorra parandamiseks. Kui tegu on pikaajalise raviga, siis üks nendest meetmetest on korduvate vastusvõttude tegemine (Jimmy ja Jose, 2011).

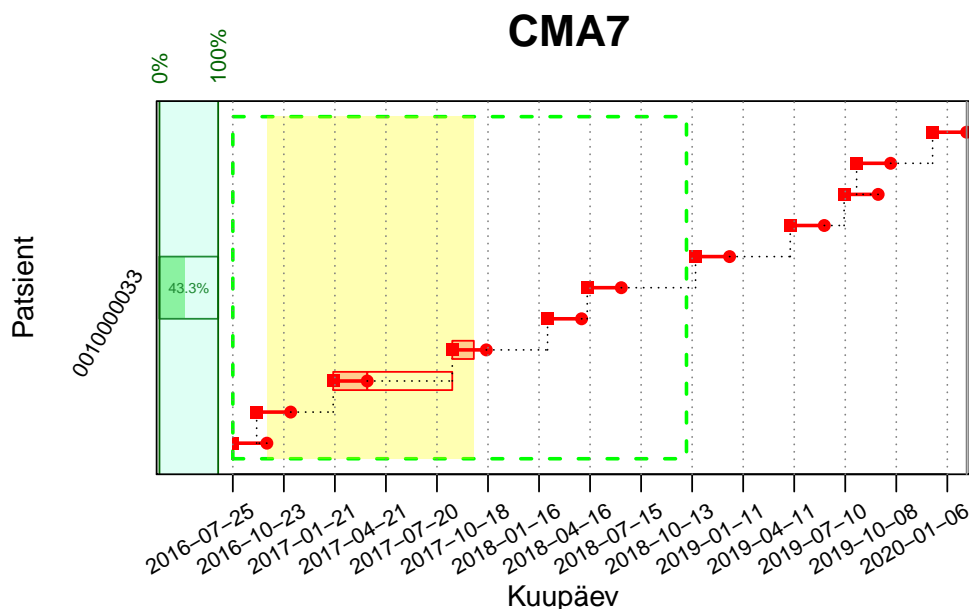
Ravijärgimus on jagatud kaheks: primaarne ja sekundaarne ravijärgimus. Primaarne ravijärgimus on binaarne näitaja, mis vaatab kas isik on talle kirjutatud ravimi välja ostnud. Sekundaarne ravijärgimus on näitaja, mis järgib, kas inimene võtab ravimit ette nähtud viisil. Kui ravim jääb välja ostmata, on tegu primaarse ravi mittejärgimisega. See-eest kui ravimit võetakse katkendlikult või mitte ettenähtud aja jooksul, siis on tegu sekundaarse ravi mittejärgimisega (Lam ja Fresco, 2015). Ravijärgivaks peetakse isikut, kelle ravijärgimus on vähemalt 80%. Kõige levinum viis selle arvutamiseks on välja kirjutatud ravimikoguse ja ravimi võtmiseks ettenähtud päevade arvu suhte leidmine (Raebel *et al.*, 2013).

Sekundaarse ravijärgimuse mõodu leidmiseks kasutatakse peamiselt statistikas kasutatava programmeerimiskeele R paketti *AdhereR*. CMA ehk *continuous multiple-interval measures of medication availability/gaps* arvutamiseks on vaja vähemalt patsiendi identifikaatorit, ravimisündmuse kuupäeva ja ravimi kestus. Ravimisündmus on ravimi väljaostmise või -kirjutamise kuupäev ja ravimi kestuse all mõeldakse päevade arvu kauaks ravimit jätkub, kui võtta seda nii nagu ettenähtud.

CMA-l on üheksa erinevat arvutusviisi. Selles tõos on kasutatume CMA7, mis erinevalt osadest CMA arvutusviisidest arvestab vaatlusaknale eelneva ravimi varuga ning kannab need edasi. Lisaks on CMA7 piiratud 0 ja 1 vahel, mis elimineerib olukorra, kus isik on ostnud välja uue retsepti enne eelmise varu otsa saamist ning saab seetõttu värtuse mis on suurem kui 1 (Dima, Dediu ja Allemann, 2022).

CMA7 arvutusvalem:

$$\text{CMA7} = \frac{\text{ravi järgivate päevade arv}}{\text{päevade arv vaatlusakna algusest lõpuni}}$$



Joonis 1: Identifikaatoriga "0010000033" isiku ravijärgimus.

Kollane osa joonisel 1 on vaatlusaken. Vaatlusaknas oleva perioodi põhjal arvutatakse CMA. Rohelisega on piiritletud jälgimisaken, mis on terve periood, mille kohta ravimisündmuse andmed on kogutud ja punased pidevad jooned näitavad välja ostetud ravimi kestvust.

2 Andmestik

Andmed pärinevad Eesti digiretsepti andmebaasist. Tegemist on ajavahemikust 01.01.2011-03.02.2020 kogutud andmetega. Kõik andmestikus esinevad inimesed on astmaatikud, kes olid 2011. aastaks vähemalt 18-aastased. Nad on astma diagnoosi saanud peale 01.01.2015 ning neile on välja kirjutatud vähemalt 10 retsepti astmaravimeid. Algses andmestikus on 1 077 051 rida ning 11 tunnust. Iga rida on üks välja kirjutatud retsept. Kokku on andmestikus 5673 inimest ning 942 erinevat ATC koodi.

Esialgul olid andmestikus järgnevad tunnused:

- isiku ID;
- ravimi ATC kood;
- sugu;
- sünniaasta;
- ravimi välja kirjutamise kuupäev;
- ravimi välja ostmise kuupäev;
- ravimi toimeaine ATC koodi järgi;
- ravimise annustamise kogus;
- ravimi annustamise koguse ühik ravimvormiga;
- ravimi annustamise kordi päevas;
- ravimi annustamise päevade arv

2.1 Sekundaarse ravijärgimuse andmestik

Kõige pealt rakendame algsel andmestikul bakalaureusetöö „Ravimisoostumise enustamine kroonilistel haigetel eelneva käitumise põhjal” andmete puhastamise koodi, kuna antud töö uuris sekundaarset ravijärgimust astma patsientide seas andmestikul, mis oli analoogsel kujul (Kuus, 2021). Selle abil filtreeriti andmestikust välja astmaravimid, mille abil on võimalik arvutada sekundaarne ravijärgimus. Andmestikku jäid ravimid, mille ATC kood algas „R03”-ga. Tegu on inhaleeritavate aerosoolidega, mida kasutatakse obstruktiivsete hingamisteede haiguste raviks (WHO, 2023b).

Sekundaarse ravijärgimuse andmestik:

- isiku ID;
- ravimi ATC kood;
- sugu;
- sünnikuupäev;
- CMA;
- ravimi võtmise pikkus aastates

Tunnus „ravimi võtmise pikkus aastates” on leitud grupeerides andmestik isiku ja ATC-koodi järgi ning leides iga isiku ja tema võetud ravimi viimase ja esimese väljaostmise kuupäeva vahe.

2.2 Primaarse ravijärgimuse andmestik

Primaarse ravijärgimuse andmestiku saamiseks filtreerisime andmestikust välja kõik ravimid, mida ei kasutata astma raviks. Seejärel grupeerisime andmestiku isiku ja ATC-koodi järgi ning võtsime välja kõigi isikute iga ravimi esimese retsepti. Järgmiseks filtreerisime andmestiku nii, et sinna jääks ravimid, mis olid välja kirjutatud peale 01.01.2012 ja enne 01.01.2015. Ajavahemik on valitud selline, et filtreerida välja võimalikult palju retseptiravimeid, mida inimesed on juba pikaajaliselt kasutanud. Kuna primaarse ravijärgimuse puhul uuritakse, kas indiviid on retseptiravimi välja ostnud ning seda vaadatakse peamiselt ühekordsete retseptiravimite pealt, siis tegime uue binaarse tunnuse „väljaostmine” tunnuse „ravimi välja ostmise kuupäev” põhjal.

Primaarse ravijärgimuse andmestik:

- isiku ID;
- ravimi ATC kood;
- sugu;
- sünnikuupäev;
- väljaostmine (1 - ravim osteti välja, 0 - ravimit ei ostetud välja)

3 Mudelid

3.1 Logistiline regressioonimudel

Alapeatükk on kirjutatud (Hosmer ja Lemeshow, 2000) põhjal.

Logistilist regressioonimudelit kasutatakse kui uuritav tunnus on binaarne ehk tal on ainult kaks võimalikku väärtust, mis on tavaliselt kodeeritud nulli ja ühega. Sündmuse esinemise tõenäosust tähistatakse $\pi = P(Y = 1)$ ning mitteesinemise tõenäosust $1 - \pi = P(Y = 0)$

Logistilise regressiooni üldkuju:

$$\ln \frac{\pi}{1 - \pi} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k,$$

kus $\beta_0, \beta_1, \dots, \beta_k$ on parameetrid ja x_1, x_2, \dots, x_k on argumenttunnused. Mudelis kasutatakse Logit seosefunktsiooni $\ln \frac{\pi}{1 - \pi} = \text{logit}(\pi)$, kuna tegu on binaarse uuritava tunnusega.

3.2 Mitme argumendiga lineaarne regressioon

Alapeatükk on kirjutatud (Rencher ja Schaalje, 2008) põhjal.

Mitme argumendiga regressioonimudeli üldkuju:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_k x_{ik} + \epsilon_i,$$

milles β_0 on vabaliige, β_j on regressioonikordajad, x_{ij} on i -nda indiviidi j -nda tunnuse väärtus ning ϵ_i on juhuslik viga.

Antud mudelit tuleks kasutada kui argumenttunnused on arvulised ja sõltuv ehk uuritav tunnus on pidev. Seda kasutatakse peamiselt andmete kirjeldamiseks, parameetrite hindamiseks ning tunnuste prognoosimiseks. Parameetrite hindamiseks

kasutatakse vähimruutude meetodit, mille idee on minimaliseerida mudeli vigade ruute.

Selleks, et mudelit kasutada peavad täidetud olema eeldused:

- juhuslikud vead on normaaljaotusega
- juhuslikud vead on sõltumatud
- juhuslikud vead on konstantse hajuvusega
- Juhuslike vigade keskväärtus on null.

3.3 Lineaarse ja logistilise regressiooni segamudel

Alapeatükk on kirjutatud (Fox, 2016) põhjal.

Logistilise regressiooni segamudel sisaldab nii fikseeritud kui ka juhuslike mõjusid. Juhuslike mõjusid kasutatakse indiviidi spetsiifiliste omaduste mõju hindamiseks, kui faktoril on palju tasemeid.

Logistilise regressiooni segamudeli üldkuju:

$$\ln \frac{\pi}{1 - \pi} = X\beta + Zu + \epsilon,$$

kus $X : n \times p$ on fikseeritud mõjude disainimaatriks, $Z : n \times q$ on juhuslike mõjude disainimaatriks, β on $p \times 1$ fikseeritud mõjude parameetrite vektor, u on juhuslike mõjude vektor pikkusega q ning ϵ on mudeli viga.

Lineaarse regressiooni segamudeli üldkuju on analoogsel kujul:

$$y = X\beta + Zu + \epsilon,$$

Juhuslik mõju võib erineda sõltuvalt vaatlusest ning on oluline korduvate mõõtmiste puhul, kuna see võimaldab mudelil arvestada sõltuvate vaatluste mitmekesisusega. Töös kasutatavas andmestikus võib olla ühe inimese kohta mitu mõõtmist kui ta

on tarvitanud erinevaid ravimeid ja segamudeli rakendamine võimaldab arvestada ja hinnata isiku spetsiifilist mõju ravijärgimusele.

3.4 Mudelite hindamine

Kui mudeli uuritav tunnus on binaarne, siis mudeli prognoosimisvõimet saab hinnata eksimismatriksiga (Kohavi ja Provost, 1998) (Tabel 1).

Tabel 1: Mudeli eksimismatriks

tegelik\ennustus	sündmus ei toimunud	sündmus toimus
Sündmus ei toimunud	õige negatiivne	vale positiivne
Sündmus toimus	vale negatiivne	õige positiivne

Eksimismatriksi põhjal on võimalik leida mudeli klassifitseerimistäpsus valemiga:

$$\text{täpsus} = \frac{\text{õige positiivne} + \text{õige negatiivne}}{\text{kõik vaatlused}}$$

Lisaks on võimalik eksimismatriksi põhjal arvutada ka tundlikkus ehk sensitiivsus ja spetsiifilisus:

$$\text{sensitiivsus} = \frac{\text{õige positiivsed}}{\text{vale negatiivne} + \text{õige positiivne}}$$

$$\text{spetsiifilisus} = \frac{\text{õige negatiivne}}{\text{õige negatiivne} + \text{vale positiivne}}$$

Sensitiivsus näitab kui suure osa uuritava sündmuse toimumistest ennustab kasutatud mudel õigesti ning spetsiifilisus näitab, kui suure osa uuritava sündmuse mittetoimumistest ennustab kasutatud mudel õigesti (Kaart, 2023a).

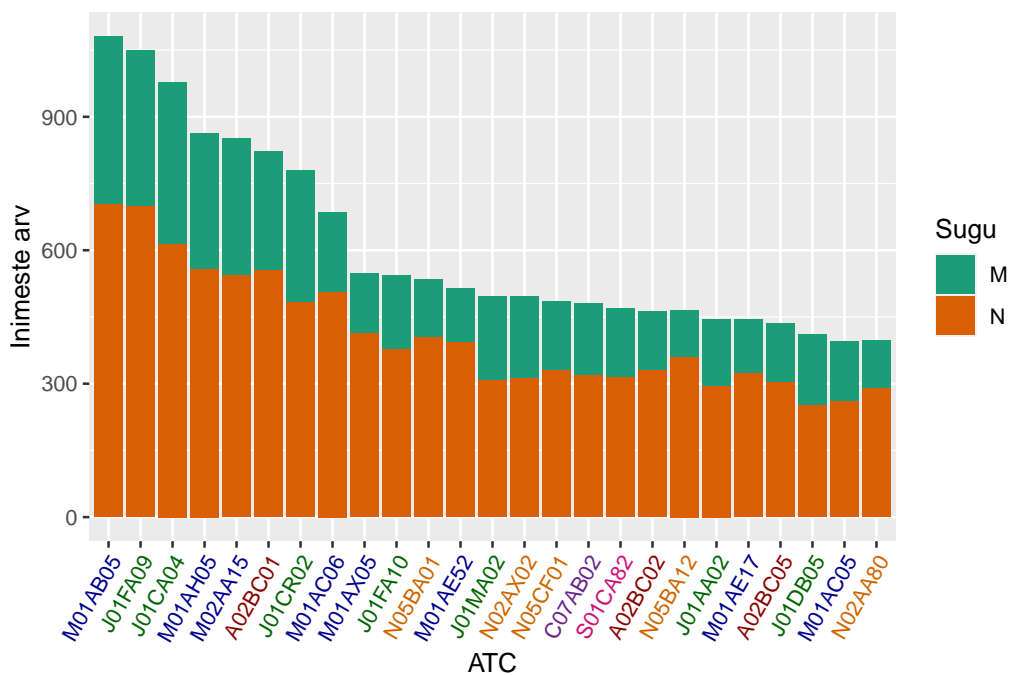
4 Andmestike kirjeldav analüüs

4.1 Primaarse ravijärgimuse andmestiku kirjeldus

Primaarse ravijärgimuse andmestikus on 33430 rida ning 123 erinevat ATC koodiga ravimit. Kokku on selles andmestikus 5135 inimest.

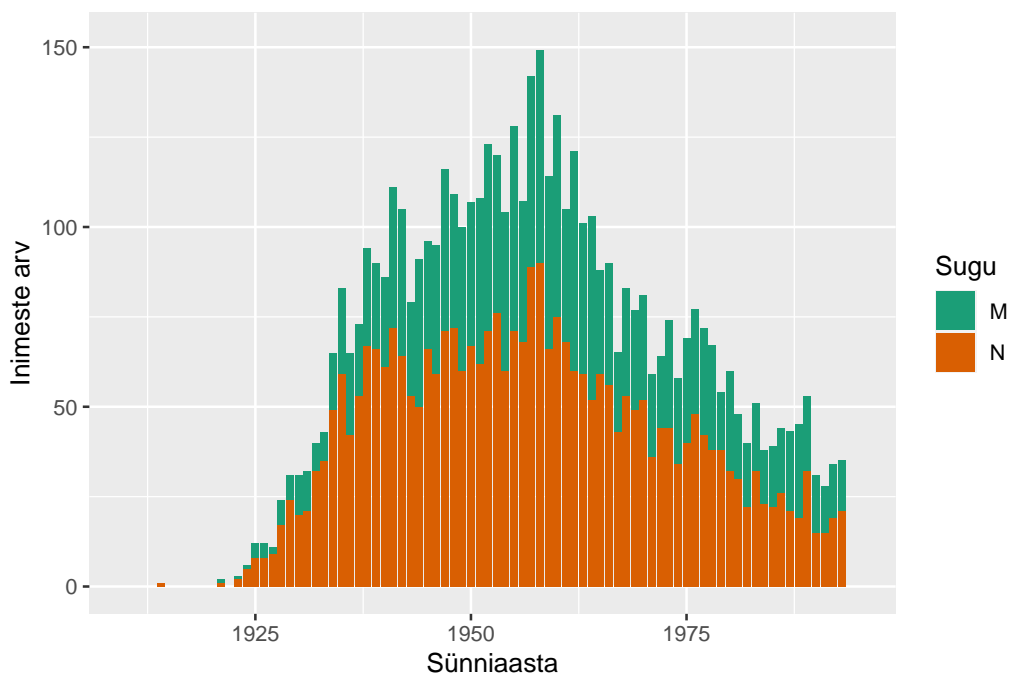
Anatoomilis-terapeutiline keemiline kood ehk ATC kood on raviainete klassifikatsioonisüsteem. Raviained jagatakse erinevatesse klassidesse olenevalt sellest, millist elundkonda nad mõjutavad. ATC koodide klassifitseerimisel on viis erinevat taset. Kõige olulisem on koodi esimene täht, mis annab teada, millisele elundkonnale raviaine mõjub (vt Lisa 1). Ülejäänud tasemed täpsustavad terapeutilisi või keemilisi alamgrupe (WHO, 2023a).

Primaarse ravijärgimuse andmestikus esineb kõige rohkem ATC koode, mis algavad tähega „M” ja „J” (vt Lisa 2). „M” tähega algavad ATC koodid on suunatud skeleti- ja lihassüsteemile ning „J” tähega algavad ATC koodid on infektsioonivastased ravimid. Veel esineb andmestikus palju „C” tähega algavaid ATC koode. Nende raviaine on suunatud kardiovaskulaarsüsteemile. Kõige rohkem andmestikus esinenud ATC koodid on välja toodud joonisel 2. Andmestikus puuduvad ATC koodid, mis sisaldavad kasvajatevastaseid ja immunomoduleerivaid aineid ning variat esindavad ATC koodid. Varia alla lähevad näiteks allergeenid, diagnostika ja toitained sisaldavad ravimid (WHO, 2023c).



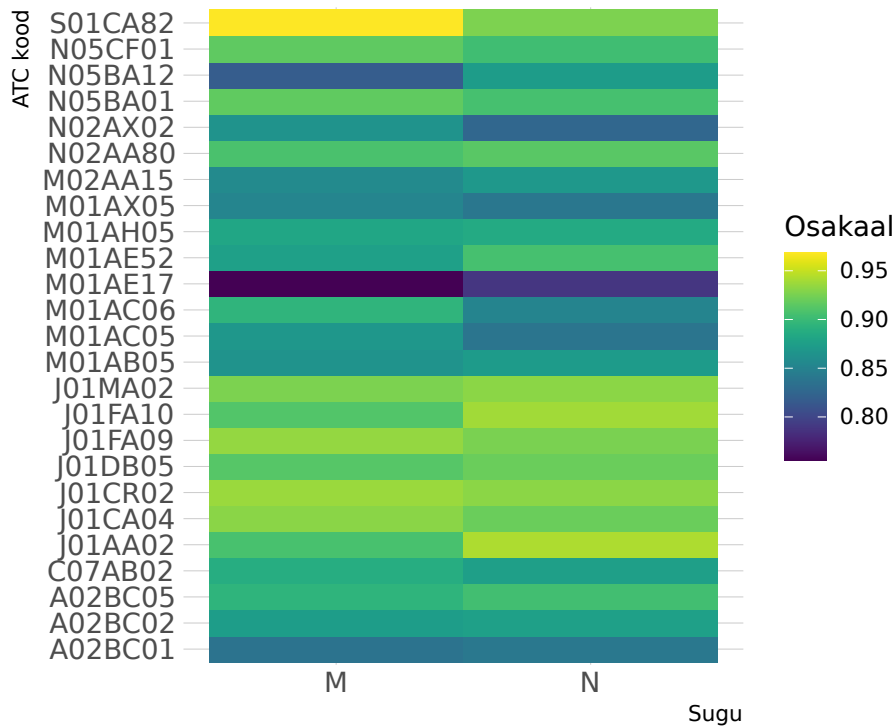
Joonis 2: 25 enim välja kirjutatud ravimit.

Joonisel 2 on näha, et spetsiifilistest ATC koodidest esineb kõige rohkem „M01AB05” ATC koodi. „M01” algusega ATC kood on põletiku- ja reumavastase raviainega. Antud koodi puhul on tegu diklofenakiga, mis on põletikuvastase ja valuvaigistava toimega aine (WHO, 2023d). Veel esineb palju „J01FA09” ATC koodi, mis on antibiootikum (WHO, 2023e). Andmestikus esinevate inimeste sünniaasta jääb vahemikku 1914-1993. 5156st inimesest 1950 on mehed ja 3185 naised. Joonisel 3 on näha täpne sünniaasta ja soo jaotus.



Joonis 3: Soo ja sünniaasta jaotus .

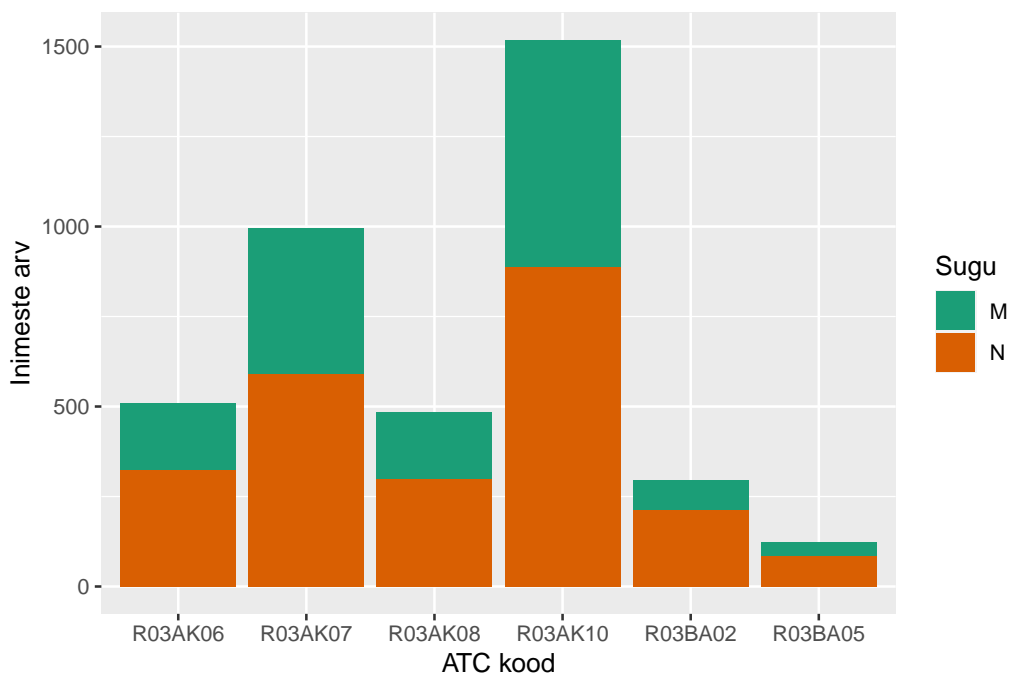
Üleüldiselt on kõige kõrgem välja ostmise osakaal B klassi ravimitel (0,95) ning kõige madalam M ja A klassi ravimitel (0,86). Kui uurida välja ostmise osakaalu erinevust meeste ja naiste hulgas, siis tuleb välja, et üldine erinevus puudub ning mõlema puhul on välja ostmise osakaal 0,89. Primaarse ravijärgimuse andmestiku pealt on leitud andestikus kõige rohkem esinevate ravimite välja ostmise osakaal soo järgi. Joonisel 4 on näha, et kõige väiksem ravimi välja ostmise osakaal 25 kõige enam välja kirjutatud ATC koodi hulgas on nii naistel kui ka meestel ravimil ATC koodiga „M01AE17“. Tegu on valuvaigistiga. Kõige kõrgem ravimi välja ostmise osakaal meestel on ravimil „S01CA82“. Selle ATC koodiga ravim on suunatud meeleelunditele, täpsemalt silmapõletike raviks (Ravimiregister, 2022). Siiski on 25 kõige rohkem kasutatava ravimi välja ostmise osakaal üle 80% iga ATC koodi korral peale „M01AE17“.



Joonis 4: 25 enim välja kirjutatud ravimi välja ostmise osakaal sooti.

4.2 Sekundaarse ravijärgimuse andmestiku kirjeldus

Sekundaarse ravijärgimuse andmestikus on 3400 inimest, kellest 1272 on mehed ning 2128 on naised. Kokku on andmestikus 6075 rida ja 6 erinevat ATC koodi. Andmestikust on välja filtreeritud read, kus isik on võtnud ravimit vähem kui aasta. Sekundaarse ravijärgimuse andmestikus olevad ATC koodid algavad kõik „R03”-ga. Kõige enam kasutavad andmestikus esinevad inimesed ravimit ATC koodiga „R03AK10”, mida on drastiliselt rohkem võrreldes teise koodidega. Lisaks on ka palju „R03AK07” ATC koodiga ravimeid, kuid ülejäänute andmestikus esinemine jääb alla 1000. ATC koodi „R03BA05” esineb andmestikus umbkaudu vaid 125 korral (Joonis 5).



Joonis 5: Välja ostetud ravimite arv soo kaupa.

Kõige kõrgema keskmise CMA-ga on ravim ATC koodiga „R03AK10”. See-eest kõrgeima mediaaniga on ravim ATC koodiga „R03AK08”. Andmestikus kõige vähem esineval ATC koodil on väikseim keskmine CMA, ent ainult kahe ravimi ATC koodide keskmine CMA jääb natukene alla 80% piiri (Tabel 2).

Tabel 2: CMA ravimite kaupa

ATC kood	keskmine	mediaan	dispersioon
R03AK06	0.8095150	0.9123288	0.05301477
R03AK07	0.7867478	0.8575342	0.05371194
R03AK08	0.8513715	0.9863014	0.04621539
R03AK10	0.8704888	0.9780822	0.03808589
R03BA02	0.8455989	0.9753425	0.04914979
R03BA05	0.7746612	0.8657534	0.06488894

Uurides CMA-de väärtusi olenevalt sellest kui kaua on isik ravimit võtnud, tuleb välja, et mida kauem on isik ravimit võtnud, seda kõrgem on tema CMA. Andmes-
 tikus olevatest isikutest oli ravimit vähemalt 4 aastat järjest võtnud 215 inimest
 ning 5 aastat järjest kõigest 5 inimest (Tabel 3).

Tabel 3: CMA aastate kaupa

Mitmes aasta	keskmine	mediaan	dispersioon
1.	0.8102394	0.9041096	0.049150864
2.	0.8364405	0.9698630	0.053432716
3.	0.9193975	1.0000000	0.023008691
4.	0.9786939	1.0000000	0.003874525
5.	0.9857534	1.0000000	0.001014825

5 Ravijärgimuse ennustamine

5.1 Primaarse ravijärgimuse mudelite analüüs

Primaarse ravijärgimuse uurimiseks tegime kõige pealt mudeli, kus ATC koodid on kategoriseeritud selle järgi, millisele elundkonnale nad suunatud on. Selle põhjal on võimalik näha, kas ja kui palju ennustavad demograafilised faktorid ning konkreetne ravimiklass primaarset ravijärgimust. Tegu on logistilise regressiooni mudeliga, mis on kujul:

$$\begin{aligned} \text{logit}(P(\text{on ravimi välja ostnud})) = & 30,699 + 0,933 \cdot (\text{B klass}) + 0,343 \cdot (\text{C klass}) \\ & + 0,434 \cdot (\text{D klass}) + 0,374 \cdot (\text{G klass}) + 0,656 \cdot (\text{H klass}) \\ & + 0,757 \cdot (\text{J klass}) + 0,036 \cdot (\text{M klass}) + 0,192 \cdot (\text{N klass}) \\ & + 0,417 \cdot (\text{P klass}) + 0,225 \cdot (\text{R klass}) + 0,940 \cdot (\text{S klass}) \\ & - 0,015 \cdot (\text{sünniaasta}) - 0,058 \cdot (\text{sugu=naine}), \end{aligned}$$

Olulisuse nivool 0,05 on kõik mudelis esinevad tunnused olulised peale soo. Mudelist näeme, et kui isikud on samal aastal sündinud ja samast soost, siis S klassi ravimi võtjatel on kõige suurem šanss ravim välja osta.

Nüüd moodustame mudeli, kus on samad demograafilised faktorid, kuid ravimite ATC koodid pole mitte esimese tasandi järgi kategoriseeritud, vaid eraldi välja toodud. Kuna andmestikus on erinevaid ATC koode 123, siis pole mudelis neid ja nende kordajad välja kirjutatud.

$$\begin{aligned} \text{logit}(P(\text{on ravimi välja ostnud})) = & 30,529 + \text{ATC} - 0,015 \cdot (\text{sünniaasta}) \\ & - 0,042 \cdot (\text{sugu=naine}), \end{aligned} \tag{1}$$

Saadud mudelis on jällegi kõik tunnused peale soo statistiliselt olulised olulisusnivool 0,05.

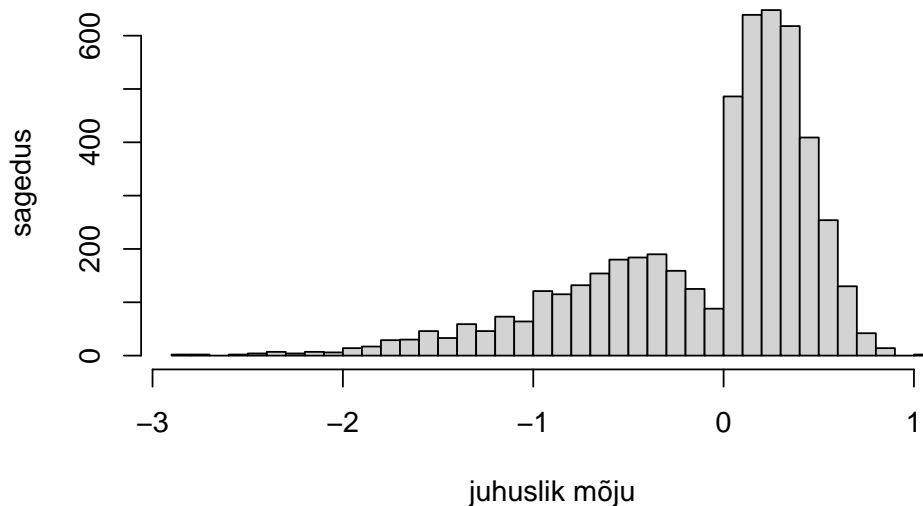
Kuna ühele inimesele võib olla välja kirjutatud mitu ravimit siis võib andmestikus olla ühe inimese kohta ka mitu rida. Selleks, et uurida isiku mõju ravimi väljaostmisele ja andmete kirjeldamisele, tegime logistilise regressiooni segamudeli.

$$\begin{aligned} \text{logit}(P(\text{on ravimi välja ostnud})) &= 32,015 + \alpha_{0,j[1]} + \text{ATC} & (2) \\ &- 0,015 \cdot (\text{sünniaasta}) - 0,008 \cdot (\text{sugu=naine}), \end{aligned}$$

Selle mudeli põhjal on saadud igale isikule juhuslik mõju ehk $\alpha_{0,j[1]}$ on i-nda isiku juhuslik mõju, mis kirjeldab indiviidi erinevust keskmisest isikust primaarse ravi järgimuse põhjal.

Juhusliku mõjuga logistilise regressiooni segamudel tehti ka sellisel kujul, kus ATC koodid ei olnud ükshaaval välja toodud, vaid klasside kaupa, kuid võrreldes seda mudeliga (2), tuli välja, et mudel (2) kirjeldab uuritavat tunnust siiski paremini.

Mudelist (2) saadud indiviidist tingitud juhuslik mõju on visualiseeritud joonisel 6. On näha, et mida suurem on isikul välja ostetud ravimite osakaal, seda suurem on tema juhuslik mõju. Enamiku inimeste juhuslik mõju tundub jäävat vahemikku 0-1. Näiteks isikul, kelle juhuslik mõju on -2,87, on välja kirjutatud kümme retsepti ning vaid üks neist on välja ostetud. Samas isikul, kelle juhuslik mõju on 1,07, on välja kirjutatud 19 retsepti, mis on kõik välja ostetud. Kui isikul on kõik välja kirjutatud ravimid välja ostetud, siis tema juhuslik mõju on üle nulli. See-eest kui tal on kõik ravimid välja ostetud, kuid neid on ainult paar tükki, siis ta juhuslik mõju on väiksem, võrreldes isikuga, kellel on ka kõik ravimid välja ostetud, kuid on neid rohkem.



Joonis 6: Primaarse ravijärgimuse indiviidi juhuslik mõju.

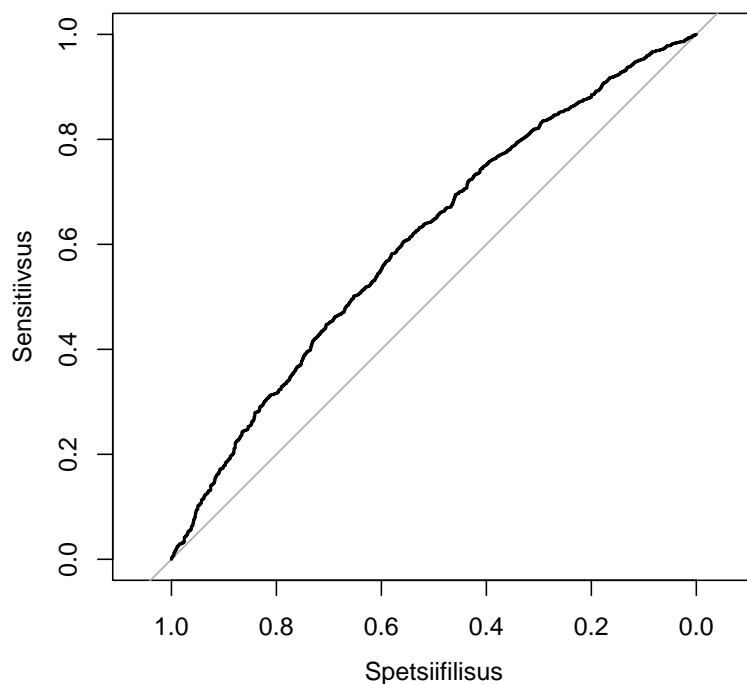
Selleks, et näha kui hästi primaarse ravijärgimuse mudel ennustab ravimite väljastmist, teeme eksimismatriksi mudelile (1). Kui valida ka primaarse ravijärgimuse puhul isiku ravijärgimuse piiriks 0,8, siis mudel klassifitseerib 86,8% inimestest õigesti (tabel 4).

Tabel 4: Juhusliku mõjuta mudeli eksimismatriks

	Ennustatud mittesoostuv	Ennustatud soostuv
Mittesoostuv	81	775
Soostuv	311	7073

Uurime ka mudeli ROC kõverat (*receiver operating characteristic curve*). Kui ROC kõver on joonisel 7 diagonaalil, siis on sensitiivsus ja spetsiifilisus 0,5 ning tegu on juhusliku pakkumisega. ROC kõvera pealt on võimalik valida otsustuspiir soovitud sensitiivsuse ja spetsiifilisuse jaoks (Kaart, 2023b). ROC kõvera alune pindala (*area*

under the curve ehk AUC) on üks mudeli headuse näitajatest. Hetkel on see 0,61 ehk mudeli ennustamisvõime on 61%, mis pole just kõige parem tulemus (Kaart, 2023c).



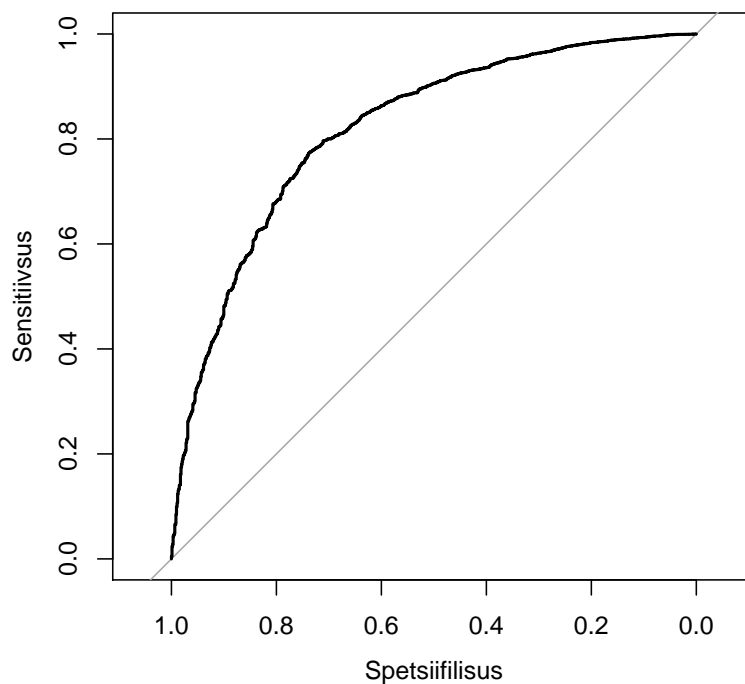
Joonis 7: Juhusliku mõjuta mudeli ROC kõver.

Nüüd teeme eksimismatriksi mudelile (2), et võrrelda, kas isiku juhuslik mõju parandab mudeli klassifitseerimisvõimet. Näeme eksimismatriksi põhjal, et mudel klassifitseerib 88,3% isikutest õigesti, mis on umbkaudu 2% parem tulemus kui mudelil ilma indiviidi juhusliku mõjuta (tabel 5).

Tabel 5: Juhusliku mõjuga mudeli eksimismatriks

	Ennustatud mittesoostuv	Ennustatud soostuv
Mittesoostuv	342	590
Soostuv	377	6931

Teeme jällegi mudelile ROC kõvera (joonis 8). Selle kõvera alune pindala on 0,82 ehk mudeli ennustamistäpsus on 82%, mis on drastiliselt parem tulemus võrreldes mudeli (1) ROC kõveraga. Jooniselt on näha, et on võimalik valida selline otsustuspiir, kus sensitiivsus ja spetsiifilisus on juba peaaegu 0,8.



Joonis 8: Juhusliku mõjuga mudeli ROC kõver.

5.2 Sekundaarse ravijärgimuse mudelite analüüs

Selleks, et uurida, kas primaarne ravijärgimus võib ennustada isiku sekundaarset ravijärgimust, moodustame kõige pealt lineaarse regressimudeli CMA prognoosimiseks. Sekundaarse ravijärgimuse prognoosimiseks tehti lineaarse regressiooni

model, mis on kujul:

$$\begin{aligned} \text{CMA} = & 1,712 - 0,018 \cdot (\text{R03AK07}) + 0,045 \cdot (\text{R03AK08}) + 0,067 \cdot (\text{R03AK10}) \\ & + 0,048 \cdot (\text{R03BA02}) - 0,011 \cdot (\text{R03BA05}) \\ & - 0,001 \cdot (\text{sünniaasta}) - 0,002 \cdot (\text{sugu=naine}) + 0,048 \cdot (\text{mitmes aasta}), \end{aligned}$$

Antud mudelis on kõik tunnused peale soo statistiliselt olulised olulisuse nivool 0,05.

Eemaldame statistiliselt ebaolulise tunnuse mudelist ning saame mudeli kujul:

$$\begin{aligned} \text{CMA} = & 1,697 - 0,017 \cdot (\text{R03AK07}) + 0,045 \cdot (\text{R03AK08}) + 0,068 \cdot (\text{R03AK10}) \\ & + 0,048 \cdot (\text{R03BA02}) - 0,011 \cdot (\text{R03BA05}) \tag{3} \\ & - 0,001 \cdot (\text{sünniaasta}) + 0,048 \cdot (\text{mitmes aasta}), \end{aligned}$$

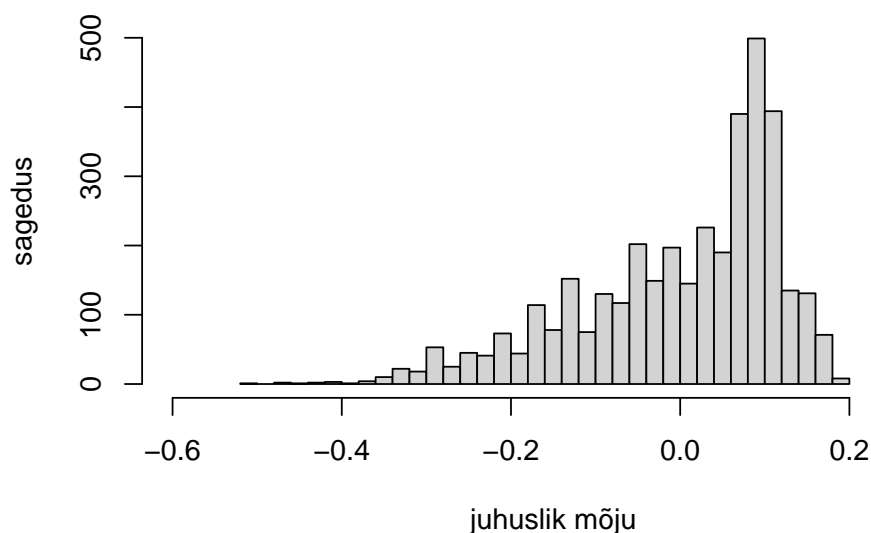
Mudelist (3) näeme, et kui kõik muud tunnused on samad, siis inimestel, kes võtavad ravimit ATC koodiga „R03AK10” on kõige kõrgem CMA võrreldes inimestega, kes võtavad teisi mudelis esinevaid ravimeid.

Nüüd moodustame lineaarse regressiooni segamudeli, kuna analoogselt primaarse ravijärgimuse andmestikule, võib ka sekundaarse ravijärgimuse andmestikus olla mitu rida ühe inimese kohta kui nad on võtnud erinevaid astma ravimeid.

$$\begin{aligned} \text{CMA} = & 1,810 + \alpha_{0,j[1]} - 0,023 \cdot (\text{R03AK07}) + 0,053 \cdot (\text{R03AK08}) \\ & + 0,071 \cdot (\text{R03AK10}) + 0,063 \cdot (\text{R03BA02}) - 0,016 \cdot (\text{R03BA05}) \tag{4} \\ & - 0,001 \cdot (\text{sünniaasta}) + 0,022 \cdot (\text{mitmes aasta}), \end{aligned}$$

Mudelis (4) on kõik tunnused statistiliselt olulised olulisunivool 0,05. Isiku mõju sekundaarsele ravijärgimusele on välja toodud joonisel 9. Kõik inimesed, kelle juhuslik mõju on üle nulli, on ravijärgivad. Kõige madalam juhuslik mõju -0,51 oli isikul, kes oli sama ravimit võtnud kaks aastat järjest. Esimesel aastal oli ta CMA 0,16 ja teisel 0. Kõige kõrgem juhuslik mõju 0,19 oli isikul, kes oli sama ravimit

neli aastat järjest ning kõigil nendel aastatel oli ta CMA 1.



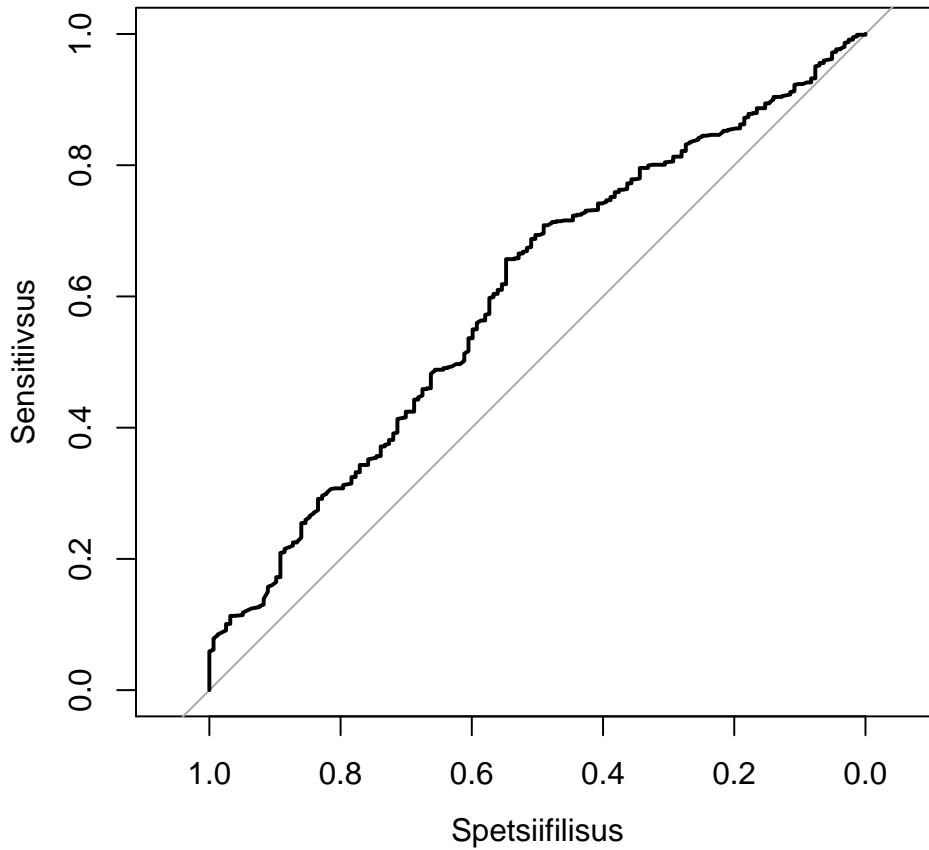
Joonis 9: Juhuslik mõju.

Uurides mudeli (3) eksimismaatriksit tuleb välja, et see mudel klassifitseerib õigesti 82,58% isikutest õigesti, kui ravijärgimuse piiriks on 0,8% (Tabel 6).

Tabel 6: Juhusliku mõjuta mudeli eksimismaatriks

	Ennustatud mittesoostuv	Ennustatud soostuv
Mittesoostuv	7	150
Soostuv	19	794

Mudeli (3) ROC kõvera alune pindala on kõigest 0,60 ning jooniselt on näha, et ROC kõvera parim võimalik otsustuspiir on selline, kus sensitiivsus ja spetsiifilisus on natuke alla 0,6 (Joonis 10). Tegu on päris kasina tulemusega, sest $AUC = 0,5$ korral on tegemist juhusliku pakkumisega.



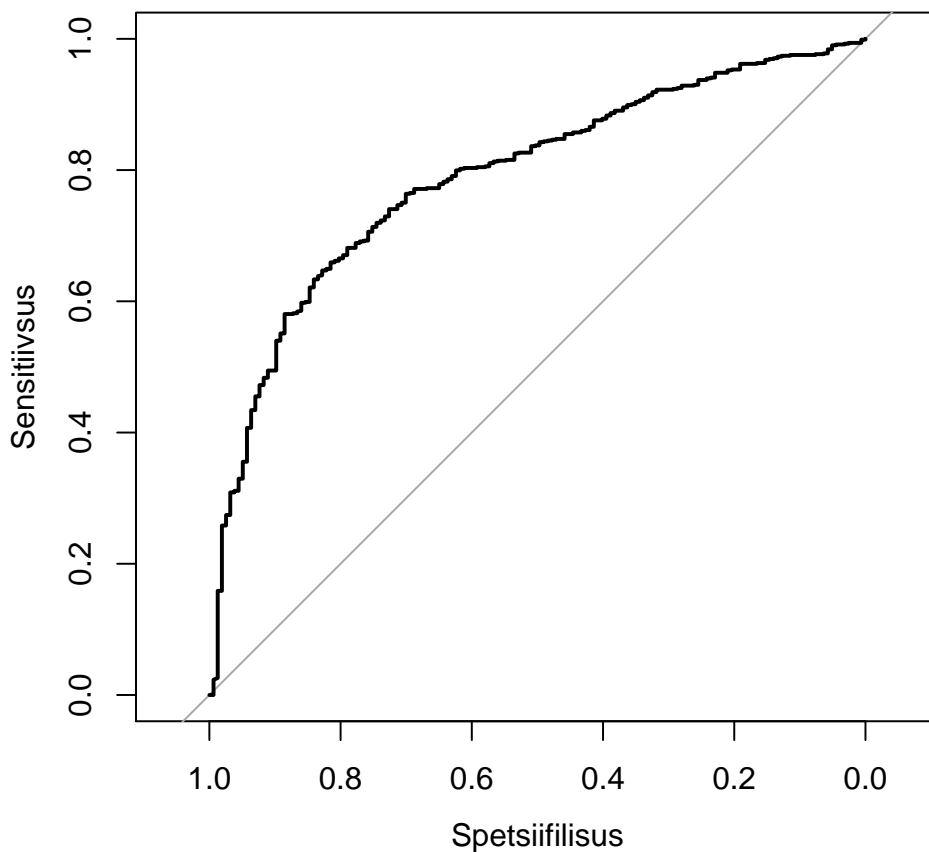
Joonis 10: Juhusliku mõjuta.

Vaatame, kas lineaarne regressiooni segamudel (4) ennustab sekundaarset ravijärgimust täpsemalt. Tabelist 7 on näha, et kui ravijärgimuse piiriks on 0,8%, siis mudel (4) klassifitseerib 81,86% inimestest õigesti.

Tabel 7: Juhusliku mõjuga mudeli eksimismatriks

	Ennustatud mittesoostuv	Ennustatud soostuv
Mittesoostuv	51	106
Soostuv	70	743

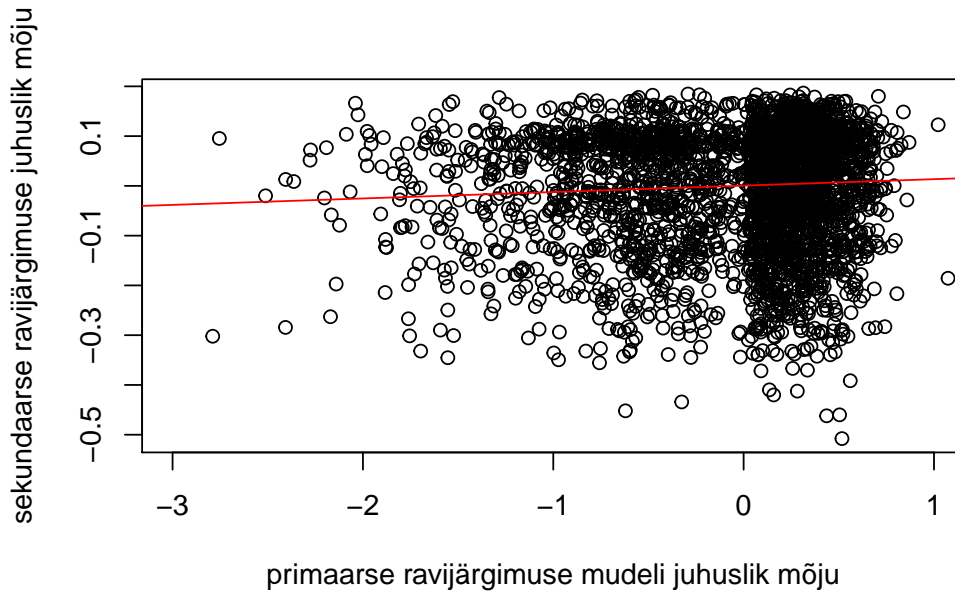
Mudeli (4) ROC kõvera alune pindala on 0,79 (Joonis 11). Tulemus on küll parem võrreldes mudeli (3) ROC kõveraga, kuid parima võimaliku otsustuspiiri korral on sensitiivsus ja spetsiifilisus siiski umbkaudu 0,7.



Joonis 11: Juhusliku mõjuga.

5.3 Mõõtude võrdlus

Visualiseerides isiku juhusliku mõju nii primaarsele kui ka sekundaarsele ravijärgimusele, tuleb välja, et korrelatsioon nende vahel on kõigest 0,1 (Joonis 12). See tähendab, et tunnuste vaheline statistiline seos on väga nõrk.



Joonis 12: Juhuslike mõjude võrdlus.

Selleks, et uurida, kas primaarne ravijärgimus mõjutab või aitab ennustada sekundaarset ravijärgimust, võtame primaarse ravijärgimuse logistilise regressiooni segamudelid välja isikule vastava juhusliku mõju ja lisame selle fikseeritud mõju sekundaarse ravijärgimuse lineaarse regressiooni segamudelisse. Saame mudeli kujul:

$$\begin{aligned}
 \text{CMA} = & 1,799 - 0,023 \cdot (\text{R03AK07}) + 0,053 \cdot (\text{R03AK08}) + 0,072 \cdot (\text{R03AK10}) \\
 & + 0,064 \cdot (\text{R03BA02}) - 0,017 \cdot (\text{R03BA05}) \\
 & - 0,001 \cdot (\text{sünniaasta}) + \alpha_{0,j[1]} + 0,021 \cdot (\text{mitmes aasta}) \\
 & + 0,022 \cdot (\text{indiviidi juhuslik mõju primaarsele ravijärgimusele}),
 \end{aligned} \tag{5}$$

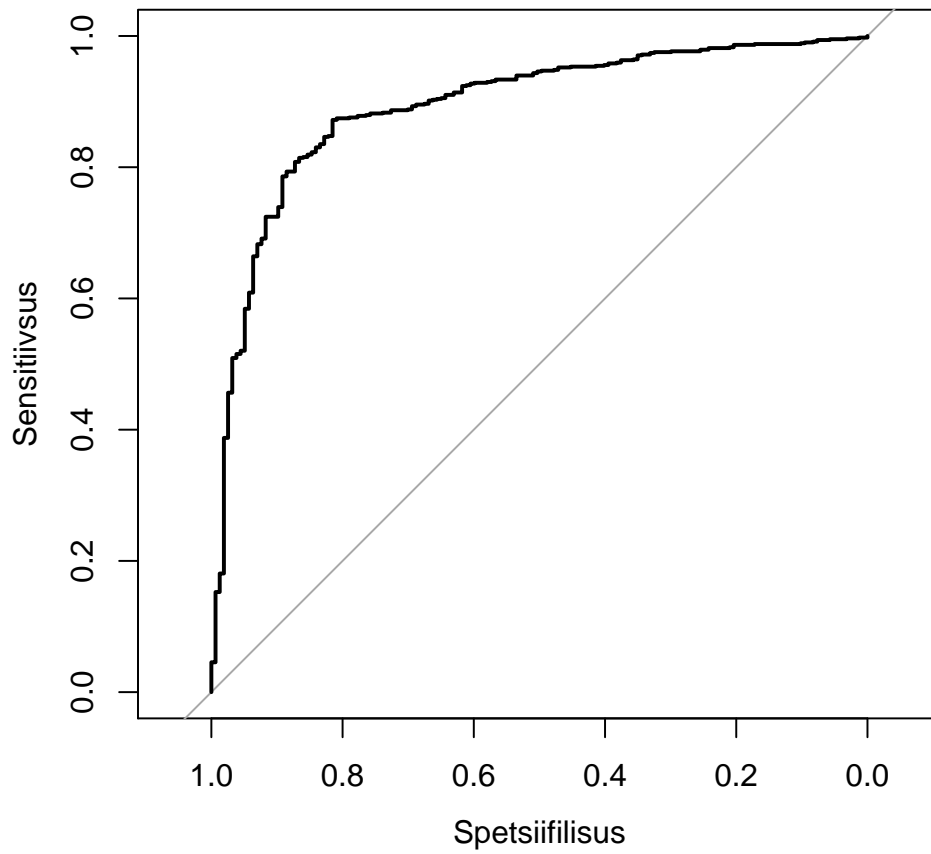
Selles mudelis on kõik tunnused statistiliselt olulised olulisusnivool 0,05. Võrreldes mudeleid (4) ja (5) tuleb välja, et mudel, millesse on lisatud primaarse ravijärgimuse mudelist saadud indiviidi juhuslik mõju on parem, kui mudel ilma selleta.

Moodustades mudelile (5) eksimismaatriksi, saame teada, et sekundaarse ravijärgimuse mudel koos juhusliku mõju ja isiku mõjuga primaarsele ravijärgimusele klassifitseerib 87,11%, mis on siiani kõige parem tulemus (Tabel 8).

Tabel 8: Sekundaarse ravijärgimuse mudeli eksimismaatriks

	Ennustatud mittesoostuv	Ennustatud soostuv
Mittesoostuv	74	83
Soostuv	42	771

Joonisel 13 on näha mudeli (5) ROC kõver, mille alune pindala on 0,89. Tegu on väga hea tulemusega. Saadud ROC kõvera puhul on võimalik valida selline otsustuspiir, kus sensitiivsuse ja spetsiifilisus on 0,8 või isegi natukene rohkem. Saadud mudel (5) on nii ROC kõvera aluse pindala kui ka eksimismaatriksi põhjal kõige paremate tulemustega ning kokkuvõttes ennustab sekundaarset ravijärgimust väga hästi.



Joonis 13: Juhuslik mõju.

Kokkuvõte

Töö eesmärk oli uurida, kas isiku primaarse ravijärgimuse põhjal on võimalik ennustada tema sekundaarset ravijärgimust. Kõige pealt moodustati logistilise regressiooni mudelid primaarse ravijärgimuse prognoosimiseks. Tuli välja, et ainult demograafilistest näitajatest ja ATC koodidest ei piisa hea mudeli saamiseks. Kuigi saadud mudeli klassifitseerimistäpsus oli 86,8%, siis selle ROC kõvera alune pindala oli kõigest 0,61. Seejärel moodustati primaarse ravijärgimuse ennustamiseks logistilise regressiooni segamudel, mille eksimismatriksi ja ROC kõvera tulemused olid juba paremad.

Sekundaarse ravijärgimuse ennustamisel kasutati kõige pealt lineaarse regressiooni mudelit. Mudeli ROC kõvera alune pindala oli jällegi päris väike, seega moodustati lineaarse regressiooni segamudel, et vaadata, kas indiviidi juhuslik mõju parandaks saadud tulemusi. Tuli välja, et analoogselt primaarset ravijärgimust ennustatavale logistilise regressiooni segamudelile, parandab indiviidi juhusliku mõju lisamine saadud mudeli ROC kõverat märkimisväärselt.

Sekundaarset ravijärgimust ennustatavale segamudelile lisati isiku juhuslik mõju, mis saadi primaarset ravijärgimust ennustatavast segamudelist, fikseeritud mõjuna ning võrreldes kõigi teiste tehtud mudelitega, olid sellel mudelil parimad tulemused. See tähendab, et primaarset ja sekundaarset ravijärgimust ei ole võimalik hästi ennustada vaid demograafiliste faktoridega ja isiku juhuslikul mõjul on suur roll ravijärgimuse korrektse ennustamises. Kui teada isiku juhuslikku mõju primaarsele ravijärgimusele, siis on saadud lõpliku mudeli kõrge klassifitseerimistäpsuse tõttu võimalik ennustada sekundaarset ravijärgimust üpris hästi.

Kasutatud allikad

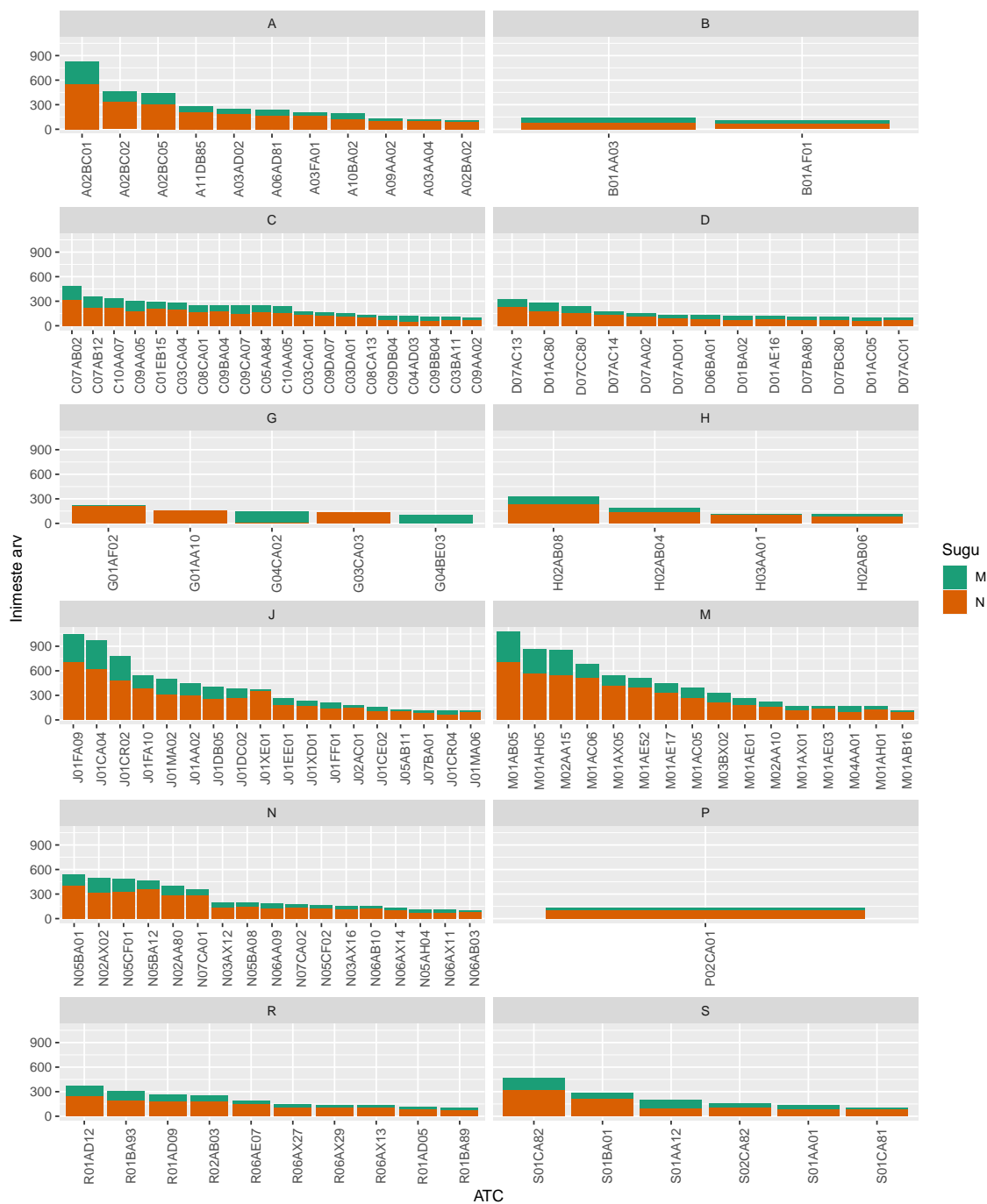
- Dima, Alexandra L., Dan Dediu ja Samuel Allemann (2022). *AdhereR: Adherence to Medications*. URL: <https://cran.r-project.org/web/packages/AdhereR/vignettes/AdhereR-overview.html> (vaadatud 02.05.2023).
- Fox, John (2016). *Applied regression analysis and generalized linear models*. 3. väljaanne. Sage.
- Gast, Alina ja Tim Mathes (2019). “Medication adherence influencing factors—an overview of systematic reviews”. *Systemic Reviews*. DOI: <https://doi.org/10.1186/s13643-019-1014-8>.
- Hosmer, David W. ja Stanley Lemeshow (2000). *Applied Logistic Regression*. 2. väljaanne. Pearson Addison Wesley.
- Jimmy, Beena ja Jimmy Jose (2011). “Patient Medication Adherence: Measures in Daily Practice”. *Oman Medical Journal*. DOI: [10.5001/omj.2011.38](https://doi.org/10.5001/omj.2011.38).
- Kaart, Tanel (2023a). *Binaarsete tunnuste analüüsimeetodid*. URL: http://ph.emu.ee/~ktanel/bin_tunnuste_analyys/pt35.php (vaadatud 29.04.2023).
- (2023b). *Binaarsete tunnuste analüüsimeetodid*. URL: http://ph.emu.ee/~ktanel/bin_tunnuste_analyys/pt36.php (vaadatud 29.04.2023).
- (2023c). *Binaarsete tunnuste analüüsimeetodid*. URL: http://ph.emu.ee/~ktanel/bin_tunnuste_analyys/pt38.php (vaadatud 29.04.2023).
- Kohavi, Ron ja Foster Provost (1998). “Glossary of terms. Special issue of applications of machine learning and the knowledge discovery process”. *Machine Learning* 30. DOI: <https://doi.org/10.1023/A:1017181826899>.
- Kuus, Kaari (2021). “Ravimisoostumise ennustamine kroonilistel haigetel eelneva käitumise põhjal”. Bakalaureusetöö. Tartu Ülikool.

- Lam, Wai Yin ja Paula Fresco (2015). “Medication Adherence Measures: An Overview”. *Biomed Res Int*. DOI: [10.1155/2015/217047](https://doi.org/10.1155/2015/217047).
- Pähklamäe, Jarmo (2022). “Ravisoostumuse ennustamine patsiendi isiku põhjal”. Bakalaureusetöö. Tartu Ülikool.
- Raebel, Marsha A., Julie Schmitt diel, Andrew J. Karter, Jennifer L. Konieczny ja John F. Steiner (2013). “Standardizing Terminology and Definitions of Medication Adherence and Persistence in Research Employing Electronic Databases”. *Medical Care* 51. DOI: [10.1097/MLR.0b013e31829b1d2a](https://doi.org/10.1097/MLR.0b013e31829b1d2a).
- Ravimiregister (2022). *OFTAN DEXA-CHLORA*. URL: <https://www.ravimiregister.ee/en/Print.aspx?pv=Pakendid.Pakend&vid=5743f310-0adc-42de-9ea1-83e12be726e8> (vaadatud 07.05.2023).
- Rencher, Alvin C. ja G. Bruce Schaalje (2008). *Linear models in statistics*. 2. väljaanne. John Wiley ja Sons, Inc.
- WHO (2023a). *Anatomical Therapeutic Chemical (ATC) Classification*. URL: <https://www.who.int/tools/atc-ddd-toolkit/atc-classification> (vaadatud 29.04.2023).
- (2023b). *ATC/DDD Index*. URL: https://www.whocc.no/atc_ddd_index/?code=r03 (vaadatud 07.05.2023).
- (2023c). *ATC/DDD Index*. URL: https://www.whocc.no/atc_ddd_index/?code=V&showdescription=yes (vaadatud 29.04.2023).
- (2023d). *ATC/DDD Index*. URL: https://www.whocc.no/atc_ddd_index/?code=M01AB05 (vaadatud 07.05.2023).
- (2023e). *ATC/DDD Index*. URL: https://www.whocc.no/atc_ddd_index/?code=J01FA09 (vaadatud 07.05.2023).

Lisa 1. ATC koodide esimene tasand

Esimene täht	Elundkond
A	seedetrakt ja ainevahetus
B	veri ja vereloomeelundid
C	kardiovaskulaarsüsteem
D	dermatoloogilised preparaadid
G	urogenitaalsüsteem ja suguhormoonid
H	süsteemsed hormoonpreparaadid (v.a suguhormoonid)
J	infektsioonivastased ravimid
L	kasvajatevastased ja immunomoduleerivad ained
M	skeleti- ja lihassüsteem
N	kesknärvisüsteem
P	parasiitide vastased ained
R	hingamissüsteem
S	meeleelundid
V	varia

Lisa 2. ATC koodid



Lihtlitsents lõputöö reprodutseerimiseks ja üldsusele kättesaadavaks tegemiseks

Mina, Minni-Marii Paarmets,

1. annan Tartu Ülikoolile tasuta loa (lihtlitsentsi) minu loodud teose „Primaarse ja sekundaarsse ravijärgimuse mõõtude võrdlus“, mille juhendaja on Raivo Kolde, 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 4.0, mis lubab autorile viidates teost reprodutseerida, levitada ja üldsusele suunata ning keelab luua tuletatud teost ja kasutada teost ärieesmärgil, kuni autoriõiguse kehtivuse lõppemiseni.
3. Olen teadlik, et punktides 1 ja 2 nimetatud õigused jäävad alles ka autorile.
4. Kinnitan, et lihtlitsentsi andmisega ei riku ma teiste isikute intellektuaalomandi ega isikuandmete kaitse õigusaktidest tulenevaid õigusi.

Minni-Marii Paarmets

06.06.2023